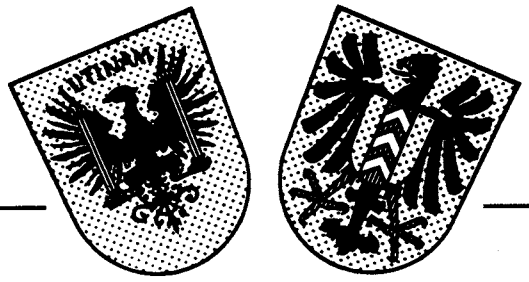


E F T F 95

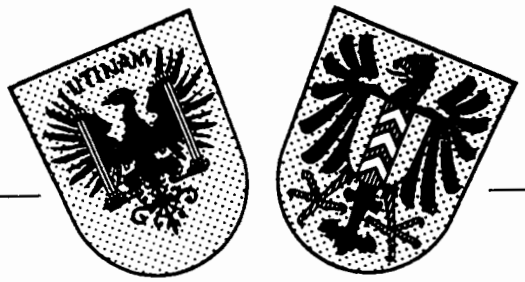
Neuvième Forum Européen Temps-Fréquence

Ninth European and Time Forum

ACTES
PROCEEDINGS

8, 9, 10 MARS 1995 **8, 9, 10 MARCH 1995**

BESANÇON
FRANCE



E F T F 95

Neuvième Forum Européen Temps-Fréquence

Ninth European Frequency and Time Forum

ACTES

PROCEEDINGS

8, 9, 10 MARS 1995

8, 9, 10 MARCH 1995

BESANÇON
FRANCE

9e FORUM EUROPEEN
TEMPS-FREQUENCE

9th EUROPEAN FREQUENCY
AND TIME FORUM

Organisé par

Organized by

Société Française des Microtechniques
et de Chronométrie - Besançon (France)

8, 9, 10 mars 1995

8, 9, 10 March 1995

Le KURSAAL

Place Granvelle

25000 - BESANÇON (France)

Ces actes ont été imprimés avec l'aide du CONSEIL GENERAL DU DOUBS et par les soins de son imprimerie.

These proceedings have been printed by courtesy of "CONSEIL GENERAL DU DOUBS" in its printing office.

9e FORUM EUROPEEN TEMPS-FREQUENCE
9th EUROPEAN FREQUENCY AND TIME FORUM

1995 Local Organization	R.J. BESSON
1995 Technical Program Chairman	D. HAUDEN
Organisation/Secrétariat - Organization/Secretary	I. BOURGON, C. HISLEUR

E.N.S.M.M. - L.C.E.P., 26 chemin de l'Epitaphe, 25030 BESANÇON Cedex, France
Tél : 81.40.28.20 - 81 40.28.21 - Fax : 81.88.57.14

L.P.M.O. - C.N.R.S., 32 avenue de l'Observatoire, 25044 BESANÇON Cedex, France
Tél. 81.66.69.99 - Fax : 81.66.69.98

COMITE EXECUTIF - EXECUTIVE COMMITTEE

R. BESSON
Laboratoire de Chronométrie, Electronique et Piézoélectricité, Ecole Nationale Supérieure de Mécanique et des Microtechniques, Besançon, France

M. ECABERT
Fondation Suisse pour la Recherche en Microtechnique, Neuchâtel, Switzerland

J.J. GAGNEPAIN
Département Sciences pour l'Ingénieur, CNRS, Paris, France

P. KARTASCHOFF
Cormondrèche, Switzerland

C. AUDOIN
Laboratoire de l'Horloge Atomique du CNRS, Paris, France

B. SCHLUETER
Keres Engineering Ltd., Neuchâtel, Switzerland.

Honorary member of the Executive Committee

A. REMOND
Service Chronométrie de l'Observatoire, Besançon, France

Avec le soutien : With support of :

Région de Franche-Comté, Conseil Général du Doubs, Ville de Besançon

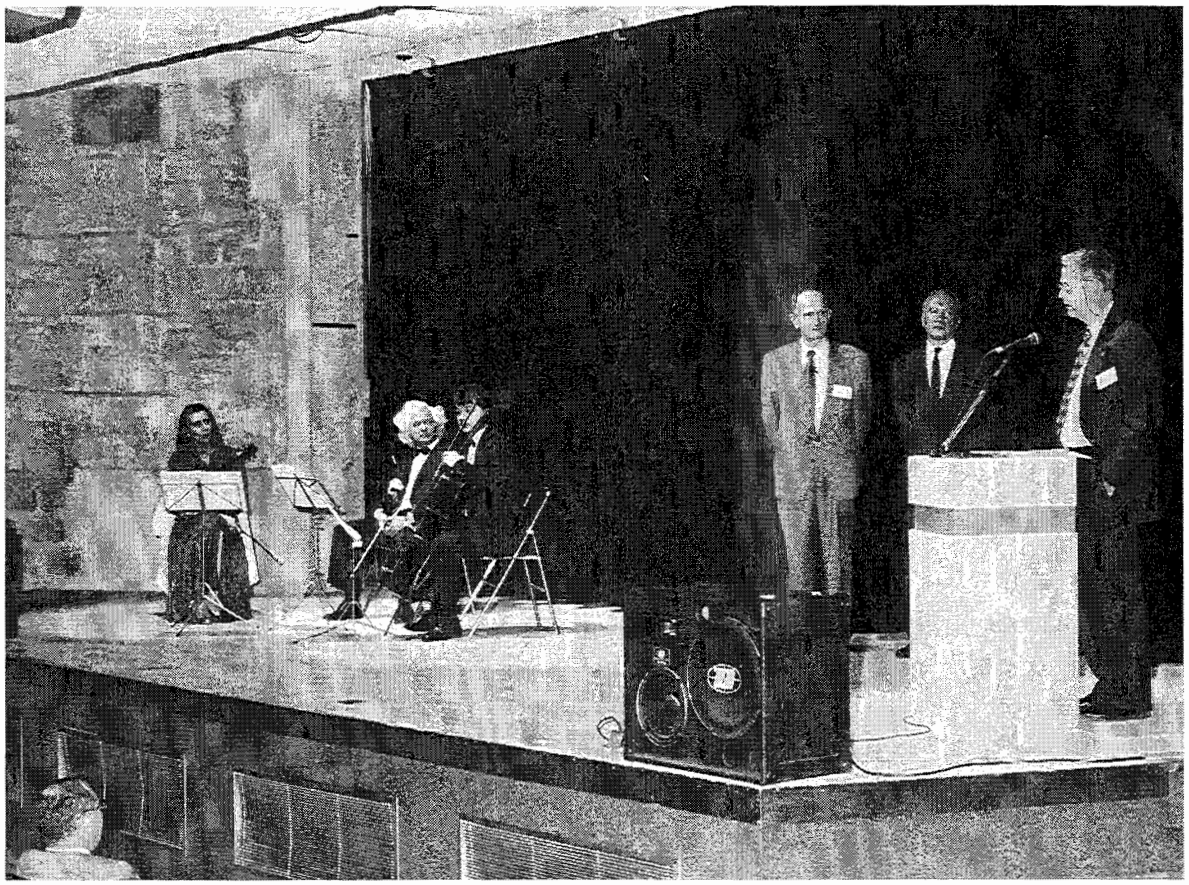
Société Française des Microtechniques et de Chronométrie,
Ecole Nationale Supérieure de Mécanique et des Microtechniques, Observatoire de Besançon, Laboratoire de Chronométrie, Electronique et Piézoélectricité (ENSMM), Laboratoire de Physique et Métrologie des Oscillateurs (CNRS).

COMITE SCIENTIFIQUE - SCIENTIFIC COMMITTEE

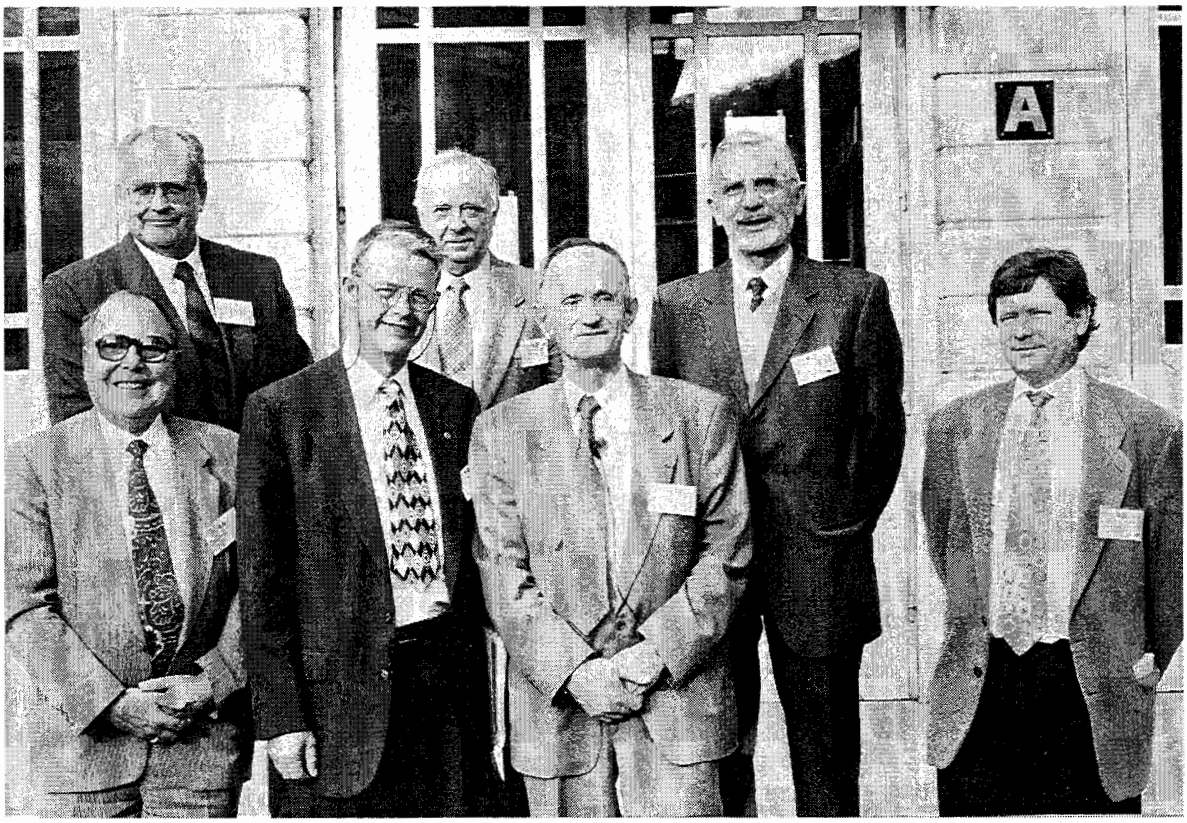
AUBRY J.P.	Thomson CEPE, Argenteuil	France
AUDOIN C.	LHA - CNRS, Paris	France
BALLATO A.	US army Research Lab., Fort Monmouth	USA
BAUCH A.	P.T.B., Braunschweig	Germany
BESSON R.J.	ENSMM - LCEP, Besançon	France
BOLOIX R.	Observatoire de San Fernando	Spain
BRUNET M.	CNES, Toulouse	France
BUSCA G.	Observatoire Cantonal, Neuchâtel	Switzerland
CUTLER L.	Hewlett-Packard Co, Palo Alto	USA
DE JONG G.	NMi Van Swinden Lab., Delft	Netherlands
DE MARCHI A.	Politechnico di Torino	Italy
DEFRANOULD P.	Thomson-Microsonics, Sophia-Antipolis	France
DEYZAC F.	ONERA, Chatillon sous Bagneux	France
DORENWENDT K.	P.T.B., Braunschweig	Germany
DOWSETT J.	C-MAC Quartz Crystals Ltd., Harlow Essex	UK
DUCHAUSSOY H.	DRET, Paris	France
ESLELINEN P.	Kotka Institute of Technology, Kotka	Finland
FELTHAM S.	ESA / ESTEC, Noordwijk	Netherlands
GODONE A.	IEN Galileo Ferraris, Torino	Italy
GRAF E.	Oscilloquartz SA, Neuchâtel	Switzerland
GRANVEAUD M.	LPTF - Observatoire de Paris	France
HARTL P.	Institut fur Navigation, Universität Stuttgart	Germany
HAUDEN D.	LPMO - CNRS, Besançon	France
HEIGHWAY J.	Advanced SAW Products SA, Bevaix	Switzerland
HILTY K.	Swiss Telecom PTT, Bern	Switzerland
HRUSKA C.K.	York University, Downsview	Canada
KALLIOMAKI K.	University of Oulu	Finland
KARTASCHOFF P.	Cormondrèche	Switzerland
KIHARA M.	NTT Transmission Systems Lab., Yokosuka	Japan
KIRCHNER D.	Technische Universität Graz	Austria
KREMPL P.W.	AVL List GmbH, Graz	Austria
LESCHIUTTA S.	Polytecnico di Torino, Torino	Italy
NAKAGIRI K.	Radio Research Laboratory, Tokyo	Japan
NAU H.	DLR-Abt Hochfrequenz-Verfahren, Welssling	Germany
NEUBIG B.W.	Tele Quartz GmbH, Neckarbischofsheim	Germany
NIEUWEJAAR P.W.	Norwegian Naval Material Command, Haaskonsvern	Norway
PROST L.	Office Fédéral de Métrologie - Wabern	Switzerland
QUINN T.J.	B.I.P.M., Sèvres	France
REID A.	British Telecom, London	UK
SCHLUTER W.	Institut fur Angewandte Geodasie, Kotzing	Germany
STEELE J.	National Physical Lab., Teddington	UK
THOMANN P.	Observatoire Cantonal, Neuchâtel	Switzerland
ULJANOV A.	Kvarz Inst. of Electronic Meas., Nizhny Novgorod	Russia
UNVERDROSS R.	Unverdross Technik, Worthsee	Germany
VANIER J.	National Research Council of Canada, Ottawa	Canada
WALLS F.L.	N.I.S.T., Boulder	USA
WINKLER G.	US Naval Observatory, Washington	USA
YANNONI N.	Hanscom AFB, Bedford	USA
ZAMPETTI G.	Telecom Solutions, San José	USA
ZINGG W.	Microcrystal Div. of ETA SA, Grenchen	Switzerland

Opening session of 1995 EFTF

1995 EFTF attendees in front of Kursaal in Besançon



Dr Fred L. WALLS just after the Award has been presented to him
by Prefet J.L. DUFEIGNEUX (center)



After the Award ceremony
The Award Winner Dr Fred L. WALLS with the Executive Committee

TABLE DES MATIERES - TABLE OF CONTENTS

INTRODUCTION

Executive Committee, Support iii
Scientific Committee iv
Illustrations v

Opening session

Trapped Ion Standards : current status and future prospects 1
Lute Maleki (invited) - Jet Propulsion Laboratory -
California Institute of Technology - Pasadena (CA) - USA

Session 1
RESONATORS I

Accelerometric and magnetic sensitivities of BVA resonators 9
J.J. Boy, F. Deyzac**, A. Berthaut**
* LCEP-ENSMM - Besançon - France
** ONERA - Chatillon sous Bagneux - France

A novel technique for measuring drive level dependence of quartz units 16
M. Koyama, Y. Watanabe**, H. Selimoto**, Y. Oomura***
* Nihon Dempa Kogyo - Saitama - Japan
** Tokyo Metropolitan Univ. - Tokyo - Japan

Characterization of losses in quartz resonators using electromagnetic waves reflection coefficient measurements 20
G.D. Mansfeld, A.D. Freik*, M. Mourey***
* Institute of Radioengineering and Electronics RAS - Moscow - Russia
** LCEP-ENSMM - Besançon - France

Study of SC-cut lateral field excited resonators features 24
A. Masiukicwicz, B. Gniewinska, W. Szulc, K. Weiss
Tele & Radio Research Institute - Warsaw - Poland

Session 2
SPACE APPLICATIONS

Automatic control system for ground- and space- based hydrogen frequency standards 29
Y.G. Gouzhva, A.G. Gevorkyan, V.Z. Gankin, A.L. Myasnikov
R.I.R.T. - St-Petersburg - Russia

Monitoring of the prare clock on-board meteor-3 : operations and performance analysis 34
S. Bedrich, A. Härting***
* Institute of Radio Frequency Technology - DLR - Oberpfaffenhofen - Germany
** GeoForschungsZentrum Postdam - Oberpfaffenhofen - Germany

Time aspects of the European complement to GPS : summary of the experimental results 40
P. Uhrich, B. Juompan*, R. Tourde*, M. Brunet**, J.-F. Dutrey***
* LPTF - Paris - France
** CNES - Toulouse - France

Frequency and time transfer with geodetic GPS receivers : first results 46
P. Baeriswyl, T. Schildknecht, J. Utzinger, G. Beutler
Astronomical Institute - Berne - Switzerland

Session 3
MATERIALS I

On the current state and problems of determination of the nonlinear constants of quartz 53
C.K. Hruska (invited)
Piezoelectricity Research Lab. - North York - Canada

Progress in crystal growth and acoustic characterization of gallium orthophosphate : $GaPO_4$ 59
D. Palmier, M. Cochez*, R. Gohier**, C. Bonjour***, G. Marianneau***, A. Zarembovitch**, E. Bigler***, E. Philippot**
* LPMS-CNRS - Montpellier - France
** LDRU-CNRS - Paris - France
*** LPMO-CNRS - Besançon - France

$GaPO_4$: a critical review of material data 66
P.W. Krempl, F. Krispel, W. Wallnofer, G. Leuprecht
AVL List GmbH - Graz - Austria

Session 4
TIME TRANSFER

Improved time and frequency transfer opportunities via GPS 71
D.W. Allan, R. Giffard**, L.S. Cutler**, J.A. Kusters***
* Allan's Time, Inc. - USA
** Hewlett Packard Company

An automated signal delay monitoring system for a two-way satellite time transfer terminal 75
D. Kirchner, H. Ressler**, R. Robnik***
* Technical University Gras - Austria
** Space Research Institute - Graz - Austria

GLONASS computer controlled receiver : first experiments on time transfer ■ 80
N. Koshelyaeusky, A. Zorin***
* GP "VNIIFTRI - Mendeleevo - Moscow Region - Russia
** Moscow State Inst. for Elect. Engng - Moscow - Russia

Closing errors observed during two-way time transfer experiments 81
J.A. Davis, P.R. Pearce and K.A. Bartle
National Physical Laboratory, Queens Road, Teddington, Middlesex, UK
P. Hetzel
Physikalisch-Technische Bundesanstalt - Braunschweig - Germany
D. Kirchner
Tecnical University Graz, Austria
H. Ressler
Space Research Institute, Graz, Austria
W. Powell, A. McKinley, J.A. De Young, W. Klepczynski
United States Naval Observatory, Washington, DC, USA
G. de Jong
NMi, Van Swinden Laboratories, Delft, The Netherlands
A. Söring
Forschungs- une Technologiezentrum, DBP Telekom, Darmstadt, Germany
F. Baumont, P. Claudon, P. Grudler
Observatoire de la Côte d'Azur, Grasse, France
C. Hackman
National Institute of Standards and Technology, Boulder, Colorado, USA
L. Veenstra
RSI, Comsat World Systems, 6560 Rock Spring Drive, Bethesda, Marylands, USA

Session 5
OSCILLATORS

Recent advances in composite heated quartz resonators intended for high stability mobile OCXOs ■ 87
I.V. Abramson
Omsk Piezoengineering Co. Ltd. - Omsk - Russia

Nonlinear simulation of quartz crystal oscillators 88
N. Ratier, R. Brendel*, T. Blin*, G. Marianneau*, P. Guillemot***
* LPMO - CNRS - Besançon - France
** CNES - Toulouse - France

New developments in capacitance switch array for digital command of crystal oscillators 93
E. Girardet, J. Marguier*, G. Marianneau**, W. Daniau***
* AR Electronique - Besançon - France
** LPMO - CNRS - Besançon - France

Voltage controlled crystal oscillators using strong piezoelectric crystal units ■ 99
V.M. Aksyonov, V.V. Bezdelkin
Joint Stock Company "Fonon", Moscow - Russia

■ Abstracts only

Session 6
ATOMIC STANDARDS I

A Cesium fountain frequency standard : present status and future improvements 101
S. Ghezali, P. Laurent, S. Lea, G. Santarelli, M. Bahoura, K. Szymaniec, A. Clairon (invited)
LPTF - Paris - France

Proposal for the generation of a continuous beam of cold atoms 111
N. Dimarcq, E. Aucouturier, P. Petit, P. Cerez
LHA - CNRS - Orsay - France

Atomic trapping and cooling in an anisotropic trap 115
N. Sagna, G. Dudle, P. Thomann
Observatoire Cantonal - Neuchâtel - Switzerland

Session 7
RESONATORS II

Finite element analysis using trigonometric interpolating functions for high overtones thickness modes resonators 121
F. Raoelijaona, B. Dulmet
LCEP - ENSMM - Besançon - France

Torsional double-ended tuning fork quartz crystal force sensor 127
P. Truchot, R. Bourquin
LCEP - ENSMM - Besançon - France

On minimising quartz crystal resonator ageing behaviour and comprehensive ageing statistics for 3rd and 5th overtone quartz crystal resonators in the frequency range 5 to 16.384 MHz 132
K.K. Tuladhar, G. Jenni
Oscilloquartz S.A., Neuchâtel - Switzerland

An electrodeless BT cut, 5th overtone quartz crystal resonator at 50 MHz 141
C. Longet, C. Robichon
bva INDUSTRIE S.A., Besançon - France

Session 8
ATOMIC FREQUENCY STANDARDS : SIGNAL GENERATION AND ANALYSIS

Reducing the effect of interrogation phase noise on the frequency stability in passive frequency standards : experimental results 147
R. Barillet, D. Venot, C. Audoin
LHA - CNRS - Orsay - France

Confidence on the modified Allan variance and the time variance 153
M.A. Weiss, F.L. Walls*, C. Greenhall**, T. Walter****
* NIST - Time & Frequency Div. - Boulder - USA
** Jet Propulsion Laboratory - Pasadena, CA - USA
*** HEPL Gravity Probe B. - Stanford Univ. - USA

Systematic errors in cesium beam frequency standards introduced by digital control of the microwave excitation 166
W.D. Lee, J.J. Shirley, F.L. Walls, R.E. Drullinger
NIST - Time & Frequency Div. - Boulder - USA

The effect of harmonic distortion on phase errors in frequency distribution and synthesis 170
F.L. Walls, F. Ascarrunz***
* NIST - Time & Frequency Div. - Boulder - USA
** Spectral Dynamics Inc. - Boulder - USA

On the dependence of spatial phase variations with the RF source in cavities for cesium fountain standards 177
A. Khursheed, G. Vecchi, A. de Marchi
Politecnico di Torino - Dip. di Elettronica - Torino - Italy

Session 9
SAW I

SAW filters for mobile communication application 183
G. Riha (invited), K. Anemogiannis, J. Bauregger, J. Machui
Siemens Matsushita Components - Munich - Germany

Equivalent circuit configurations for SAW coupled resonator filters and the experimental determination of their elements 190
D.P. Chen, C. Lambert*, M.A. Schwab*, C. Hartmann***
* Advanced SAW Products S.A. - Bevaix - Switzerland
** Hartmann Research Inc. - Dallas - USA

Determination of the equivalent circuit parameters of transversely coupled SAW resonator filters 195
J. Beier, B. Neubig***
* Technical Univ. of Ilmenau - Germany
** Tele Quarz Group - Neckarbischofsheim - Germany

Accurate analysis of pseudo-SAW devices 200
P. Ventura, J.M. Hode, M. Solal, L. Chommeloux
Thomson Microsonics - Sophia-antipolis - France

Session 10
TELECOM APLICATIONS

Telecommunication network synchronization 205
M. Wolf
Alcatel SEL - Stuttgart - Germany

Telecommunications network synchronisation and wander generated by SDH equipment 210
R. Urbansky
Philips Kommunikations Ind. - Nuremberg - Germany

Some aspects of timing in telecommunication and power systems 217
A. Dobrogowski, M. Jessa
Institute of Electronics & Telecom. - Poznan - Poland

Stability characterization of stand alone synchronisation equipment 220
R.Bonello, A.Manzalini*, A.Godone**, C. Novero**, P. Tavella***
* CSELT - Turin - Italy
** Istituto Elettrotecnico Nazionale G. Ferraris - Turin- Italy

Award session

The quest to understand and reduce 1/f noise in amplifiers and BAW quartz oscillators 227
F.L. Walls
National Institute of Standards and Technology - Boulder - USA

Session 11
MATERIALS II

A look into the crystal ball : some thoughts on the prediction of the ageing of BAW and SAW resonators 245
D. Bower, N. Churchill, J. Dowsett, D. Salt
C-Mac Quartz Crystals - Harlow - England

Insitu X-ray diffraction study of sweeping of impurity ions in quartz crystals 251
M.T. Sebastian, R.A.Becker**, H. Klapper***
* Regional Research Lab. - Trivandrum - India
** Institut fur Kristallographie - Aachen - Germany

Defects characterization in synthetic and natural quartz : the role of ionizing radiations 255
(Part II of "Nondestructive study of Al-hole concentration and frequency change of quartz resonators")
J.J. Boy, M.A. Pasquali*, B. Viard**, K. Yamni***
* LCEP - ENSMM - Besançon - France
** LCCM - UFR Sciences - Besançon - France

Study of impurities in quartz by electric measurements 261
C. Poignon, G. Jeandel, G. Morlot
LEMTA - Vandoeuvre-lès-Nancy - France

Session 12
ATOMIC STANDARDS II

First results on the high C-field Cesium resonator at the Politecnico di Torino 265
C. Costanzo, E. Rubiola, A. De Marchi
Politecnico di Torino - Dip. di Elettronica - Italy

Study of the S/N performance of passive atomic clocks using a laser pumped vapour 271
G. Mileti, P. Thomann
Observatoire Cantonal - Neuchâtel - Switzerland

New Russian national time and frequency standard and time & frequency service in Russia 277
B. Al'shin, V. Borisochkin, B. Gaigerov, G. Yolkin, N. Koshelyaevsky, S. Pushkin, V. Tatarenkov
GP "VNIIFTRI", Mendeleevo, Moscow region
G. Aprishkin
NPO "DALSTANDARD", Khabarovsk
V. Koutcherov
GOSSTANDARD, Moscow
Yu Medvedev
ATZUS, Golitzino, Moscow region
N. Tel'pukhovski
VS NIIFTRI, Irkutsk
V. Stepanov
NPO "SNIIM", Novosibirsk, Russia

The use of hydrogen masers in TAI computation 283
J. Azoubib, C. Thomas
BIPM - Sèvres - France

Session 13
RESONATORS III

Langasite plano-convexe resonators using lateral or perpendicular field excitation 289
J. Detaint, B. Capelle*, A. Zarka**, D. Cochet-Muchy****
* France Telecom CNET/PAB - Bagneux - France
** LMCP - Paris - France
*** Crismatec - Gières - France

Original mode selection for whispering gallery mode resonator 297
*V. Giordano, W. Daniau, G. Martin, F. Lardet-Vieudrin, M. Olivier, O. di Monaco, M. Chaubet***
* LPMO - CNRS - Besançon - France
** CNES - Toulouse - France

Finite element modelling of infrared-sensitive resonators used as thermal sensors 301
M.R. Hamrour, B. Dulmet, S. Galliou
LCEP - ENSMM - Besançon - France

Frequency spectrum and frequency-temperature characteristics of 3rd overtone AT and BT-cut quartz resonators 306
J. Zelenka
Technical University in Liberec - Czech Republic

Session 14
SYNCHRONIZATION IN SYSTEMS

Highly accurate realtime clocksynchronization with GPS time receivers 311
H. Nau, W. Klische***
* DLR Oberpfaffenhofen - Wessling - Germany
** K + K, Vechelde - Germany

Development of the UTC (CRL) time dissemination server for the computer network 314
T. Gotoh, K. Imamura***
* Kashima Space Research Center - Ibaraki - Japan
** Communications Research Lab. - Tokyo - Japan

SDH-based time and frequency transfer system 317
M. Kihara, A. Imaoka
NTT Optical Network Systems Lab. - Kanagawa - Japan

Relativistic theory for the realization of coordinate time scales and the syntonization of clocks in the vicinity of the earth 323
P. Wolf (1,2), G. Petit (1)
(1) BIPM - Sèvres - France
(2) Queen Mary and Westfield College - London - U.K.

Session 15
SAW II

Recent advance in ZnO piezoelectric transducer technology (film deposition, characteristics, applications) 327
S. Kulakov (invited), V. Nefedov, O. Gabaraev, O. Gusev, V. Zhuhov
St. Petersburg State Academy of Aerospace Inst. - Russia

Surface transverse waves propagating on piezoelectric plates with thin metal strip gratings on top and bottom side simultaneously 332
S. Ballandras, E. Gavignet, E. Bigler
LPMO - CNRS - Besançon - France

Green's function analysis of the multistrip couplers for narrow band low loss filters 338
V.P. Plessky, S.N. Krasnikova***
* Advanced SAW Products SA - Bevaix - Switzerland
** Dalniaia Sviaz - St Petersburg - Russia

Comparative performances of 460 MHz low-loss SAW ring and dual-track image impedance connected filters 343
S.A. Dobershtein, V.A. Malyukhov
Omsk Research Electronic Equipment Institute - Russia

Session 16
LASER FREQUENCY STABILIZATION

Comparison of calculated phase-sensitive recoverable signals using a Fabry-Perot cavity in reflection and in transmission 345
E. Bava, F. Massari
Politecnico di Milano - Milano - Italy

Ultra-stable cavity-stabilized laser diode at 729 nm 349
B. Fermigier, M. Desaintfuscien, J. Dupont, M. Houssin, F. Plumelle
LHA - CNRS - Orsay - France

An ultrastable high power Nd:YAG laser for the detection of gravitational waves 353
F. Bondu, A. Brillet, P. Fritschel, F. Perrone, F. Cleva, L. Latrach, C.N. Man,*
CNRS - Virgo-LAL - Orsay - France
* MIT - Cambridge - USA

POSTER SESSION

Oscillators

A low cost ovenized crystal oscillator 357
P. Eskelinen, J. Frilander*, J. Engelberg*, M. Matola***
* Kotka Institute of Technology - Dept of Electrical Eng. - Finland
** RF Consulting Co - Finland

Time series analysis : an application to the stability of a crystal controlled oscillator 361
C. Eckert, J.A. Miehé, B. Cunin
Groupe d'Optique Appliquée - Centre de Recherches Nucléaires
Strasbourg - France

Analog ASIC design for miniature temperature compensated crystal oscillators 362
Y. Richard, J.M. Tarot
SOREP - Eurowave Div. - Chateaubourg - France

Frequency standards and noise

A low noise 100 MHz distribution amplifier for precision metrology 365
M. Siccardi, S. Römisch*, , F.L. Walls**, A. De Marchi**
* Dip. di Elettronica, Politecnico di Torino - Torino - Italy
** NIST - Time and Frequency Division - Boulder - USA

Low frequency noise reduction in DMTD systems 369
M.P. Dinca
University of Bucharest - Bucharest-Magurele - Romania

Cut-off frequencies and noise power law model of spectral density : adaptation of the multi-variance method using the structure function approach 373
F. Vernotte, E. Lantz**, M. McHugh*, G. Zalamansky**
* Observatoire de Besançon - France
** Lab. d'Optique P.M. Duffieux - Besançon - France

Is the frequency noise of an oscillator of deterministic origin ? 377
M. Planat, V. Giordano*, G. Marianneau*, F. Vernotte**,*
*M. Mourey***, C. Eckert****, J. Miehe*****
* LPMO - CNRS - Besançon - France
** Observatoire de Besançon
*** LCEP - ENSMM - Besançon - France
**** Centre de Recherches Nucléaires - Strasbourg - France

A miniaturized rubidium physics package of atomic clock for space, air and industrial applications 380
A.G. Gevorkyan, V.I. Khutorshchikov, G.M. Smirnova
The Russian Inst. of Radionavigation & Time - St-Petersburg - Russia

Operation of optically pumped clocks with high spectral purity laser diodes 384
G. Selo, B. Boussert, P. Cérez, N. Dimarcq
LHA - CNRS - Orsay - France

Signal-to-noise considerations in trapped-atom spectroscopy 387
V. Rieger, J.H. Müller**, D. Bettermann**, F. Ruschewitz*, J.L. Peng*, K. Sengstock*, W. Ertmer**
* Inst. f. quantenoptik der Universität, Hannover - Germany
** Instit. f. Angewandte Physik der Universität Bonn - Germany

Spectroscopic study of the CaII long-lived 3D-doublet in view of a frequency standard application 390
M. Knoop, J. Rocher, M. Vedel, F. Vedel
PI2M - Univ. de Provence - Marseille - France

Influence of spurious microwave fields in cesium clocks 394
B. Boussert, G. Théobald, N. Dimarcq
LHA - CNRS - Orsay - France

The M8 and M9 hydrogen masers as the national frequency standard at the national institute of metrology, Bucharest 397
O.C. Gheorgiu, L.C. Giurgiu**, B.M. Mihalcea*, D.M. Cacicovschi*, A. Niculescu****
* Inst. of Physics & Techno of Radiation Devices - Bucharest - Romania
** University of Bucharest - Faculty of Physics - Romania
*** National Institute of Metrology - Bucharest - Romania

Czech national time and frequency group standard 400
O. Buzek, J. Cermak, B. Cemusova
Institute of Radio Engineering and Electronics - Csech Acad. of Sciences - Praha - Czech Republic

Test results of KVARZ CH1-75 active and CH1-76 passive hydrogen masers 404
E.D. Powers, E.C. Jones*, F. Danzy*, W.M. Golding**, A.H. Frank***
Naval Research Laboratory, Washington, D.C. 20375 - USA
SFA, Inc.

Phase coherent frequency measurement chain to the visible 408
H. Schnatz, B. Lipphardt, J. Helmcke, G. Kramer, F. Riehle
Physikalisch-Technische Bundesanstalt - Braunschweig - Germany

Materials

Hydroxyl defects in germanium-doped quartz : defect dynamics and radiation effects 413
H. Bahadur
National Physical Laboratory - New Delhi - India

A continuum theory of piezoelectric and ferroelectric dielectrics including singular surfaces and interfacial properties 417
N. Daher
LPMO - CNRS - Besançon - France

Studies of crystalline quality and vibration modes in langasite single crystals 422
A. Zarka, B. Capelle*, J. Detaint**, D. Cochet-Muchy****
* Lab. de Minéralogie-Cristallographie - Paris - France
** CNET - Bagneux - France
*** Crismatec - Gières - France

Sweeping of very high purity synthetic quartz under vacuum and air atmosphere 423
J.P. Bachheimer
Lab. de Spectrométrie Physique - St-Martin-d'Hères - France

Micro fabricated silicon structures : experimental and theoretical investigations of 3-D etched shapes 426
C. Tellier, S. Durand
LCEP - ENSMM - Besançon - France

Resonators

$Ba_{5-x}Sr_xNb_4O_{15}$: a new ceramic dielectric resonator for application in communication systems 431
H. Sreemoolanathan, M.T. Sebastian*, P. Mohanan***
* Regional Research Laboratory - Trivandrum - India
** Dept of Electronics - CUSAT - Kochi - India

Microminiature crystal units and temperature sensors at 150 - 1 000 kHz 434
M.E. Bogoush, V.B. Grouzinenko
Scientific Research Institute "Phonon" - Moscow - Russia

Influence of crystallographic orientation on ferrobielastic twinning in quartz 438
P.L. Guzzo, J.J. Boy
LCEP - ENSMM - Besançon - France

The effect of mass-loading on the frequency-temperature characteristics of fundamental AT-cut quartz resonators 443
I. Mateescu, J. Zelenka**, I. Mateescu**
* Inst. of Physics & Technology of Materials - Bucharest - Romania
** Technical Univ. of Liberec - Czech Republic

Intermodulation products of quartz resonators and their measurement 447
J. Nosek
Technical University of Liberec - Czech Republic

The influence of surface structure on amplitude deviation and electrical parameters of quartz crystals 450
E. Seydel, A. Landmann
Kristall-Verarbeitung Neckarbischofsheim GmbH - Germany

Specific problems for quartz sensors in liquid media 454
C. Filiâtre, A. Foissy**, F. Membrey**, J.-P. Valentin*, M. Valentin**
* Equipe de Physique et Electronique des Capteurs - ENSMM - Besançon
** Lab. d'Electrochimie des Solides - Besançon

Six-port technique in reference measurements of piezoelectric resonators parameters 458
A. Lisowiec, M. Wojcicki
Tele-and-Radio Research Institute - Warsaw - Poland

Computer arranged crystal sets by evaluation of spurious resonances and motional parameters, enabling professional filter production with resonators of wider tolerances 462
M. Frey
Tele Quarz GmbH - Neckarbischofsheim - Germany

Recent advances in the measurement of spurious resonances of crystal units 469
B. Neubig
Tele Quarz GmbH - Neckarbischofsheim - Germany

SAW

Rotated Y-cut SAW devices immune to the thermal shock due to intermittent laser beam irradiation 473
N. Oura, M. Yamakita*, M. Takagi**, E. Momosaki***
* Precision and Intelligence lab. - Yokohama - Japan
** R&D Division, Seiko Epson Corp. - Suwa - Japan

Influence of the aluminum layer thickness on the properties of Bleustein-Gulyaev waves in $Li_2B_4O_7$ crystal 477
W. Soluch
Institute of Electronic Materials Technology - Warsaw - Poland

Design and optimization of a low cost SAW stabilised fixed frequency oscillator 480
U. Stehr, B. Adler**, H. Kostner***
* Technical University of Ilmenau - Institute of Solid - Germany
** Tele Quarz GmbH - Neckarbischofsheim - Germany

Space Applications / Synchronization

Precise two-way microwave time transfer to space-borne atomic clocks 485
S. Bedrich, J. Hahn
DLR Oberpfaffenhoffen - Wessling - Germany

Custom turnkey time & frequency systems. A structured, expandable approach 492
D.F. Wright
Radiocode Clocks Ltd - Cornwall - U.K.

Stability of a Cs-clock based 2048 kHz optical ISDN synchronization signal 498
K. Kalliomäki, T. Mansten***
* Dept of Electrical Engineering - University of Oulu - Finland
** VTT Automation, Measurement Technology - VTT - Finland

A protected GPS time / frequency reference source 502
Y. Bourrague, V. Pelletier, E. Benoist, V. Mourier
Tekelec-Systèmes - Les Ulis - France

Telecom Applications

Impact measurement configurations on the behaviour of time stability quantities 509
*R. Bonello**, M. Carbonelli*, D. De Seta*, A. Manzalini**, D. Perucchini**
* Fondazione Ugo Bordoni - Rome - Italy
** CSELT - Turin - Italy

Robust control of G.812 clock's 513
A. Dobrogowski
Institute of Electronics and Telecommunication - Poznan - Poland

Absolute frequency measurement of the 3-0 R(21) transition of CO at 1.5605 µm for optical communication standards 516
D.J.E. Knight, K.I. Pharaoh** and D.A. Humphreys***
* DK Research - Twickenham, Middx. - U.K.
** Div. of Electrical Science - National Physical Lab. - Teddington - U.K.

Time Applications

Phase-stability of XR5 GPS-receivers at VTT and University of Oulu 521
J. Mannermaa, K. Kalliomäki
Dept of Electrical Engineering - University of Oulu - Finland

The development of a prototypic time synchronized speaking clock 525
C.C. Lin, C.S. Lia, J.L. Shi, C.M. Sheu, H.D. Lin
Telecommunication Laboratories -MOTC - ROC

Digital correlator for tracking pseudo-random coded signals 529
N. Khalifa, O. Bazzi, R. Torguet
IEMN - Dept OAE - Univ. of Valenciennes - France

The accuracy of Loran-C based time scales in Finland during 1991-1994 533
K. Kalliomäki, T. Mansten***
* Dept of Electrical Engineering - University of Oulu - Finland
** VTT Automation, Measurement Technology - VTT - Finland

Theoretical treatment of closing errors occuring in two-way satellite time transfers 537
J.A. Davis
National Physical Laboratory - Teddington - U.K.

Time reference system of the ESO Very Large Telescope 541
W.R. Lange, M. Ravensbergen***
* Lange-Electronic GmbH - Gernlinden - Germany
** European Southern Observatory - Garching bei Muenchen - Germany

List of exhibitors 547

Session chairmen 548

Index of authors 549

OPENING SESSION

Trapped ion standards: Current status and future prospects *

Lute Maleki

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

April 6, 1995

Abstract

The interest in the trapped ion standards is based on the physical characteristics of ion traps, and their ability to support atoms in their unperturbed energy states. Recent developments in trapped ion standards have led to stability performance exceeding all other atomic frequency standards. Current research in this area involves the study of various ion species for the development of microwave and optical standards. Progress in lasers together with new trap configurations hold the promise of ultra-stable microwave and optical frequency standards based on laser cooled ions confined in rf traps.

1 Introduction

The first proposal for the development of a high stability frequency standard based on an ion trap was put forth by F. Major in 1969.[1] Since that time, trapped ion frequency standards have attracted much attention, and continue to be the focal point of research in a number of laboratories around the world. The interest in ion traps as a basis for the development of a frequency reference source stems from the general requirements of ultra- stable atomic standards. In such standards the achievable stability depends on the energy level structure of the active atom, as well as the environment in which the atom undergoes the clock transition. The first condition determines the inherently achievable stability, and the sensitivity of the atomic transition to the environmental perturbations. The second determines the degree to which the environmental perturbations can be reduced with respect to the atomic energy level sensitivity. The trapped ion frequency standard evidently fulfills, or shows the potential to fulfill, both these conditions, since ions with desirable energy level structures are readily confined in traps which provide an essentially benign environment.

The current research in the field of trapped ion frequency standards includes work towards both the enhancement of the stability and the performance of the microwave standards based on trapped mercury and ytterbium ions [3, 4, 7, 16, 21, 20, 26, 28], and the development of optical frequency standards based on mercury, calcium, ytterbium, and strontium ions [2, 12, 9, 13, 10, 19]. This research has already produced impressive results with the devlopment of a standard with stability corresponding to $7 \times 10^{-14}/\sqrt{\tau}$ [26], and has demonstrated the highest measured stability for averaging intervals longer than 10^5 s. The future work holds the promise for standards with even higher stability in the microwave and optical regions of the spectrum which are suitable for deployment on board spacecraft for the highest precisision metrological applications.

In this paper first a brief history of the development of the trapped ion standards will be presented, together with the physical background of these instruments. In section two a review of the current status of research on trapped ion standards will be given, followed by a perspective on future developments in section IV. The paper will conclude with a summary section.

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

2 Background

Electrodynamic ion traps used in frequency standards applications utilize an oscillating electromagnetic field to confine the ions in a small region of space. The oscillating field has a quadrupolar geometry traditionally generated by applying a potential to a set of three hyperbolic electrodes. Owing to the quadrupolar geometry, the field within the trap has a position dependence, which includes a node at the center. The position dependent field exerts a ponderomotive force on a charged particle, which for a rapidly oscillating field, always points to the center. A quadrupole field may be generated in a region od space with the potential ϕ in a region of space (x, y, z) , given as:

$$\phi = \frac{\phi_0}{r_0^2}(\lambda x^2 + \sigma y^2 + \gamma z^2). \qquad (1)$$

Here λ, σ, γ are constants, and r_0 is an appropriate dimension of the region of space in which the field is generated. ϕ_0 is of the form

$$\phi_0 = U - V\cos\Omega t \qquad (2)$$

with U a dc potential, and V the amplitude of a time varying (rf) potential of angular frequency Ω. When ϕ is used in Laplace's equation one readily obtains the condition $\lambda + \sigma + \gamma = 0$.

To obtain the equation of motion for a charged particle placed in the quadrupole potential we set up the force equation in the x-direction:

$$m\frac{d^2x}{dt^2} = -e\frac{\partial\phi}{\partial x}, \qquad (3)$$

and obtain:

$$\frac{d^2x}{dt^2} + \frac{2\lambda x}{r_0^2(U - V\cos\Omega t)})x = 0. \qquad (4)$$

Similar equations are found for y and z. Equation 4 above can be readily converted to a standard type of second order differential equation known as the Mathieu equation:

$$\frac{d^2u}{d\xi^2} + (a_u - 2q_u\cos 2\xi)u = 0 \qquad (5)$$

by setting $\xi = \omega t/2$ and defining a_x and q_x as:

$$a_x = \frac{8\lambda eU}{mr_0^2\Omega^2} \text{ and } q_x = \frac{4\lambda eV}{mr_0^2\Omega^2}, \qquad (6)$$

with similar valus for a_y,a_z, q_y, and q_z. Parameters a_u and q_u define regions in which the solutions to Mathieu equations are bounded. This allows selection of values for a_u and q_u with an applied potential where the motion of a charged particle with a particular charge to mass ratio is stable in all directions, thus confining or trapping the particle.

W. Paul made the first demonstration of ion confinement with rf quadrupole fields using a hyperbolic electrode structure. In this "Paul trap" configuration, hyperbloids of revolution form a set of three electrodes consisting of a central or ring electode and two complemantary end caps. The ring electode has a characteristic dimension r_0 and the end caps are separated from the center of the ring by the characteristic dimension z_0, as shown in figure 1.

The apparent x-y symmetry of the Paul trap, and the condition $\lambda+\sigma+\gamma = 0$ fix the coefficints of x, y and z in equation 2 as $\lambda = \sigma = 1$, and $\gamma = -2$. This leads to a particularly simple expression in polar coordinates for ϕ in the case of the Paul trap:

$$\phi(x, y, z) = \frac{\phi_0}{r_0^2}(r^2 - 2z^2) \qquad (7)$$

from which we simply obtain:

$$a_z = -2a_r = \frac{-16eU}{m\Omega^2 r_0^2}, \qquad (8)$$

and

$$q_z = -2q_r = \frac{-8eV}{m\Omega^2 r_0^2}. \qquad (9)$$

For the above expressions it is assumed that $r_0^2 = 2z_0^2$. In the general case for an aribtrary hperbola we should replace r_0^2 everywhere with $r_0^2 + 2z_0^2$.

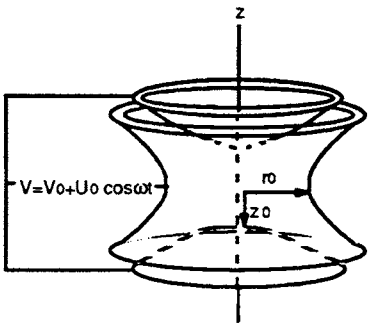


Figure 1. The Paul Trap

With the geometry described above ions may be confined by fields whose frequency Ω is in the range of fractions to tens of Mhz. Thus the perturbation of the trapping field through the

quadratic Stark effect on the hyperfine energy levels (whose frequencies are in the range of ten Ghz) of the confined ions is extremely small, and contributes negligibly to instability. Furthermore, the electrically charged ions interact only through elastic Coulomb collisions, thus remaining immune from collisions which can affect their energy level population. Finally, the long confinement times in combination with the interrogation of the clock transition with the Ramsey technique [18] allow extremely large atomic line Q's.

The confining fields of the rf trap do, however, contribute to the second order Doppler shift of the transition frequency. This is because for an aggregate of ions in the trap, the position of any ion with respect to the central node of the field determines the amplitude of the rf confining field exprienced by that ion, and thus the ion's kinetic energy due to its driven motion. The second order Doppler shift produced by of this effect results in a shift of frequency, which in turn fluctuates with the fluctuating ion number in the trap. For a large cloud of ions the size of the second order Doppler shift due to the driven motion can exceed the second order Doppler shift resulting from the thermal motion of the ions at room temperature. This is the largest source of instability caused by trapping ions in an rf trap.

An effective way to reduce the influence of the second order Doppler shift for a large cloud of ions is to utilize a trapping geometry wherein the region of the trap with zero field is larger than the singular node of the hyperbolic trap. This can be accomplished with an axially symmetric geometry where the spherical symmetry with one node is reduced to cylindrical symmetry with a line of nodes on the axis. Referring to the equation 1 above, we set $\lambda = -\sigma$ and $\gamma = 0$ to obtain:

$$\phi(x, y, z) = \frac{\phi_0}{r_0^2}(x^2 - y^2). \qquad (10)$$

This geometry is easily obtained with four parallel rods arranged in a cylindrical geometry such that the distance R between the axis of the cylinder and the surface of a rod is related to r_0 by $r_0 = \sqrt{2}R$. Ions may readily be confined with such a geometry if a small dc potential is placed on the axis at either end of the cylinder to keep them from running out (see figure 2). The field on the axis of such a trap will be zero, except near the ends, and thus a larger number of ions may be confined in the zero field region. Similarly, for a fixed number of ions, the second order Doppler shift due to the driven motion in such a trap is smaller than that produced by the single node geometry of a hyperbolic trap. Such a linear trap was constructed at JPL and was demonstrated to hold a larger number of ions with smaller second order Doppler shift [15].

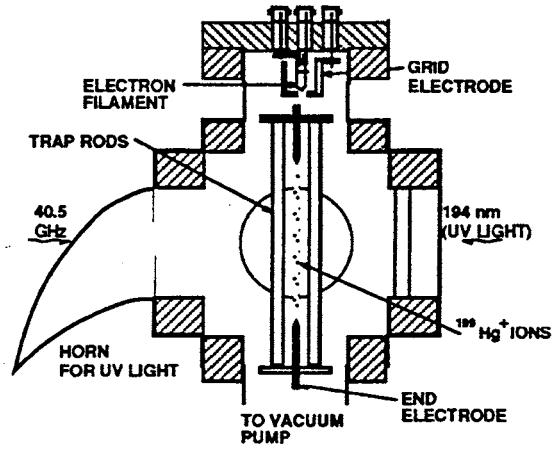


FIGURE 2. JPL'S LINEAR TRAP AND VACUUM CUBE

The ions may be cooled to near room temperature by drag cooling through collision with a buffer gas to further reduce the overall effect of the second order Doppler shift [5].

3 Operation of the Standard

The process of converting the rf ion trap to an atomic frequency standard is accomplished by first preparing the ions in the "clock state" of interest and then observing the "clock transition" as the ions interact with photons at the clock frequency. The state preparation step is based on optical pumping, in which light from a laser or a resonant lamp connects one of the ground hyperfine states of the ion to a short lived($\sim 10^{-9}$ s) excited state. The process is depicted for Hg^+ ion in Fig.[3]. The applied light at 194 nm excites a transition out of the ground F=1 state to the excited $^2P_{1/2}$ state. The latter decays after a few nano-seconds populating both F=1 and F=0 hyperfine sublevels of the ground state. After many cycles, the population of the F=1 level is transferred to the F=0 level via this optical pumping action. At this point the 194 nm radiation is no longer absorbed by the ions, and the fluorescence associated with the decay of

the $^2P_{1/2}$ state ceases; the ions no longer scatter the 194 nm pumping radiation. The excitation of the clock transition (F=0, m_f=0 → F=1, m_f=0) is the subsequent step, where radiation with a frequency corresponding to the clock transition frequency is generated with an auxiliary slaved oscillator and applied to the ions. When the frequency of the applied radiation match the clock transition (40.5 Ghz for isotope 199 Hg^+ ions) photons are absorbed, promoting the ion to the higher energy. This event may be observed since the ions can once again scatter the pumping light (194 nm for Hg^+). The intensity of the scattered light is now a measure of how closely the frequency generated by the slaved oscillator matches the clock transition. This information is used to close the loop by steering the frequency of the auxiliary oscillator to exactly match the frequency associated with the clock transition. The state preparation and the observation of the clock transition outlined above do not rely on resonant cavities, as required with other standards such as the hydrogen masers. The long observation times afforded with the confined ions in rf traps also enable long interrogation of the clock transition, thus leading to exceedingly narrow lines.

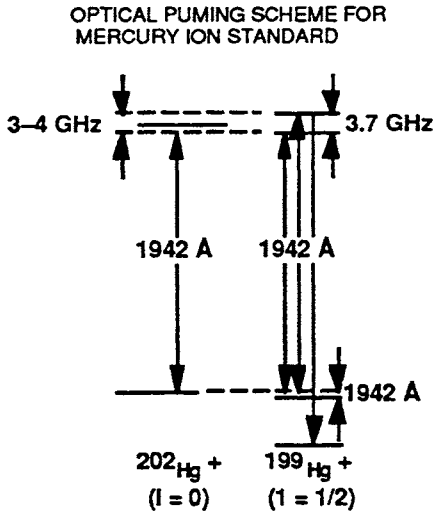


FIGURE 3

This fact, and the correspondingly larger energy separation of hyperfine sublevels in ions, as compared to atoms, lead to line Q's that can be as large as 10^{13} for an operating microwave standard and potentially orders of magnitude larger for optical standards.

The requirements of optical pumping and the observation of the clock transition by scattering of optical photons from the ions naturally lead to the consideration for the use of lasers with trapped ion standards. The stability of the trapped ion standard is given by the relation :

$$\sigma_y(\tau) = \frac{k}{SNR\, Q_\ell \sqrt{\tau}}, \qquad (11)$$

where SNR is the signal-to-noise ratio, Q_ℓ is the atomic line Q, and $k \approx 1$. Lasers can yield large signal-to-noise ratios, since they are brighter than lamps and laser light can be readily collimated. Furthermore, laser light may also be used to cool ions in the trap, thus reducing the deleterious second order Doppler effect.[14] The advent of semiconductor and solid state lasers with operational simplicity and moderate power requirements promise to further increase the stability of standards operating with lamps.

The physical simplicity of the ion trap, together with the relative insensitivity of ions to ambient field fluctuations in comparison with neutral atoms, render the task of environmental control much more simple than most other types of atomic frequency standards. For example, according to the Breit-Rabi equation the magnetic field shift produced by an external field on an atomic hyperfine level in the low field regime is given by:

$$\Delta\nu = \frac{1}{2\nu_0}\left(\frac{(g_I + g_J)\mu_B}{h}\right)^2 B_0^2. \qquad (12)$$

here B_0 is the applied field, and g_I and g_J are the nuclear and the electron g factors. Since the (field-free) hyperfine transition frequency ν_0 is larger for ions of interest than that of neutral atoms, the corresponding fractional frequency shift is smaller for trapped ions. Thus the requirement of magnetic shielding is generally easier to fulfill. Likewise, the influence of ambient electric fields is also smaller on ions. Furthermore, the absence of the need for resonant cavities in trapped ion standards eliminates the associated temperature sensitivity and frequency shifts.

4 Present Status

The first demonstration of a trapped ion frequency standard was made at the Laboratoire de l'Horloge Atomique in Orsay in 1981 [8], using mercury as the active atom . Subsequently three other mercury ion standards were fabricated at the Hewlett-Packard laboratory and installed in

the US Naval Observatory [6]. The stability of these instruments was in the range of $10^{-12}/\sqrt{\tau}$ to a few times $10^{-13}/\sqrt{\tau}$. All these instruments used an rf lamp to produce the 194 nm light required for the optical pumping , and also operated with a background pressure of helium which was used as a buffer gas. Demonstration of the clock transition with Be^+ ions was also accomplished at NIST in 1985 where laser light was used for optical pumping as well as cooling the ions.

The first demonstration of a trapped ion frequency standard with stability exceeding $10^{-13}/\sqrt{\tau}$ was made at JPL in 1993 [25, 26], using mercury ions in a linear trap. Since then other groups in Australia and Germany have developed microwave frequency standards based on laser excited ytterbium ions [7, 3]. The ytterbium ion requires light at 369 nm for optical pumping through the excitation of the $^2S_{1/2} \rightarrow ^2P_{1/2}$ transition. This wavelength is readily obtained by simply doubling the frequency of a red solid state or semiconductor laser [24, 27]. Because of this feature the ytterbium ion has attracted much attention in connection with frequency standards applications. The Australian group has utilized a linear trap to demonstrate the operation of a room temperature laser excited microwave standard to achieve stability of $2 \times 10^{-13}/\sqrt{\tau}$ out to over 10,000 s averaging intervals [7]. A similar work at JPL using the $^2S_{1/2} \rightarrow ^2 P_{3/2}$ transition (λ=329 nm) and nitrogen buffer gas points to the potential of the ytterbium ion standard with stability exceeding the best number demonstrated by the mercury ion standard ($7 \times 10^{-14}/\sqrt{\tau}$) [22].

5 Future Prospects

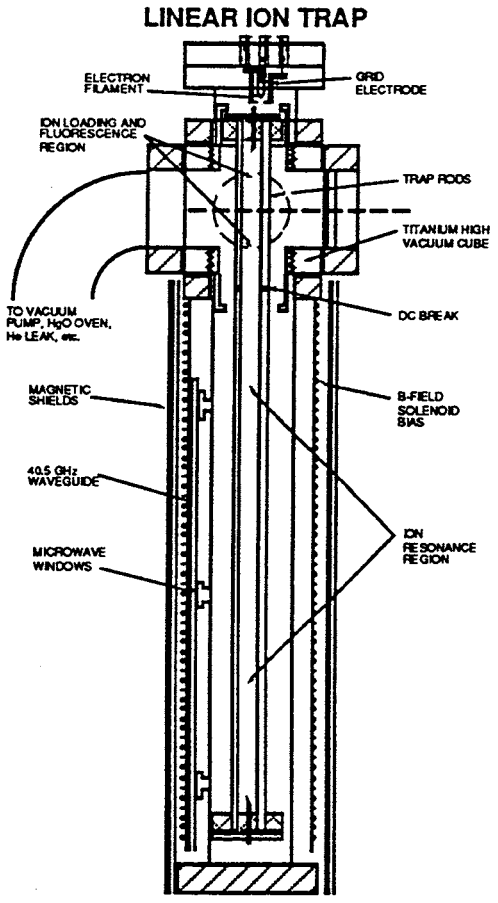
Several new technological developments point to future improvement of the performance of trapped ion standards, as well as the extension of their frequency range to the optical region. The improved stability may be obtained by increasing the Q and the SNR. Both these parameters may be increased by use of lasers, which can improve the SNR and allow development of optical standards with significantly larger Q_l's. Furthermore, lasers can be used to cool ions as mentioned before for reduced Doppler contributions to the frequency instability. Several laboratories in Canada, France, Germany, Japan, UK, and US [11, 13, 30, 20, 23, 2, 9] are actively pursuing the development of an optical frequency standard based on laser excitation and cooling of trapped ions. Long lived states of Ca^+ and Sr^+ are being exploited with semiconductor lasers for frequency standards applications.

The interest in ytterbium as the active atom in a frequency standard extends to the optical region of the spectrum as well. The existence of long-lived D and F metastable states in the energy level structure of singly ionized ytterbium brings this ion in close competition with Ca^+ and Sr^+ as candidates for optical frequency sources based on simple laser systems. Groups in Australia, Germany, and UK are active in research for the development of a optical ytterbium ion standard. Despite their attractiveness, these ions have to compete with mercury, which is the candidate of choice, save for the need for a 194 nm source. This need may however be remedied in the near future as suitable nonlinear materials with high second harmonic generation efficiency in this wavelength regime become available. Currently, work at NIST constitutes the most extensive effort towards the development of an optical standard based on laser cooled trapped mercury ions. The NIST group plans to develop a clock ensemble consisting of 10 or so trapped ions in a miniature linear trap, each of which will individually undergo the clock transition.[29] This approach promises the development of ultra-precise clocks, as well as highly stable frequency standards.

Improvements of stability are also anticipated for microwave standards using a new technique proposed at JPL [17]. The technique of shuttling attempts to bring to the trap ion standard the same configuration which most other microwave standards have enjoyed, viz. the separation of the state preparation region from the microwave interrogation region. This scheme is depicted in figure 4 where a long trap assembly actually consists from two separate traps. The first trap is where the ions are created, and prepared in the desired clock state. Subsequent to this step, the ions are "shuttled" to the second trap by applying a bias between the two trap segments. In this second region of the trap the ions interact with the applied microwave Ramsey pulses. At the conclusion of this step, the ions are shuttled back to the first trapping region where they are interrogated for the clock transition. Since shielding is required for the ions only during their interaction time with the microwave fields, the task of magnetic shielding becomes considerably less demanding for the

shuttle. This is becuase unlike the conventional linear trap where some of the input and signal collection optics must also be included in the shielded region, the shuttle geometry only requires shielding the small trapping region where the microwave field is applied. A shuttle trap such as the one depicted in the figure has been constructed at JPL and is currently undergoing tests.

The combination of laser excitation, laser cooling and the shuttle trap allow development of ultra-stable microwave frequency standards suitable for spacecraft applications. This application is important from two points of views.



First, the benign space environment is ideal for achieving ultra-stable performance of frequency standards. Once the spacecraft is in the cruise mode an onboard standard will be free of vibrations experienced on earth, and will operate in a thermally stable environment. Ambient fields are also significantly diminished in space and large reductions in their fluctuations are usually achieved. Thus, a space-borne frequency standard may operate with stability commensurate with its inherent properties and unperturbed by outside influences.

Such an ultra-stable standard may be used for important tests of relativity and gravitation physics, quantum mechanics, and cosmology. These tests are crucial to extending our understanding of physical theories, and the physical world. They are indeed the most exciting ones in the domain of ultra-stable frequency standards, where the highest achievable precision for physical measurements are accomplished.

6 Summary

Ion traps provide a very attractive basis for the development of ultra-stable frequency standards. Recent developments in the technology of ion traps have led to the development of ultra-stable mercury ion standards which now hold the world record as the most stable standards for averaging intervals longer than a few thousand seconds. Ions confined in rf traps may be excited with laser radiation to obtain improved stability through increased signal-to-noise ratio, and through the reduction of the second order Doppler effect by laser cooling. Currently a number of laboratories around the world are engaged in the development of both microwave and optical frequency standards based on laser excited ion traps. New techniques, such as shuttling of the ions between two traps, and the application of lasers will make the development of a small, low mass trapped ion standard for space applications possible. The development of optical standards and spacecraft standards based on ion traps are anticipated within the next five years.

References

[1] F. Major. NASA internal document, Goddard Space Flight Center, 1969.

[2] G.P. Barwood, C.S. Edwards, P. Gill, H.A. Klein, and W.R.C. Rowley. Development of diode and fibre laser sources for a Sr^+ trapped ion optical frequency standard. *Proceedings of the SPIE - The International Society for Optical Engineering*, 1837:271–277, 1993.

[3] R. Blatt, R. Casdorff, V. Enders, W. Neuhauser, and P.E. Toschek. New frequency standard based on Yb^+. In A. DeMarchi, editor, *in Frequency Standards and Metrology*, pages 306–311. Springer-Verlag, Berlin, 1989.

[4] R. Casdorff, V. Enders, R. Blatt, W. Neuhauser, and P.E. Toschek. A 12-GHz standard clock on trapped ytterbium ions. *Ann. Physik*, 48:41–55, 1991.

[5] L.S. Cutler, R.P. Giffard, and M.D. McGuire. Thermalization of ^{199}Hg ion macromotion by a light background gas in an rf quadrupole trap. *Appl. Phys. B*, B36(3):137–142, March 1985.

[6] L.S. Cutler, R.P. Giffard, P.J. Wheeler, and G.M.R. Winkler. Initial operational experience with a mercury ion storage frequency standard. In *Proc. 41st Annual IEEE International Frequency Control Symposium*, pages 12–17. IEEE Publication CH2427-3, 1987.

[7] Peter T.H. Fisk, Matthew J. Sellars, Malcolm A. Lawn, and Colin Coles. Performance of a prototype microwave frequency standard based on trapped $^{171}Yb+$ ions. In *in Proc. 48th Annual IEEE International Frequency Control Symposium*, pages 731–738. IEEE Publication 94CH3446-2, 1994.

[8] M. Jardino, M. Desaintfuscien, and F. Plumelle. Prospects for a mercury ion frequency standard. *Journal de Physique*, c8(12):327–338, 1981.

[9] S.R. Jefferts, C. Monroe, A.S. Barton, and D.J. Wineland. Rf-trap apparatus for optical frequency standards. In *1994 Conference on Precision Electromagnetic Measurements Digest (Cat. No.94CH3449-6)*, 1994.

[10] H. A. Klein, A. S. Bell, G. P. Barwood, and P. Gill. Laser cooling of trappped Yb^+. *Appl. Phys. B*, 50:13–17, 1990.

[11] A.A. Madej and J.D. Sankey. Quantum jumps and the single trapped barium ion: Determination of collisional quenching rates for the $5d^2D_{5/2}$ level. *Phys. Rev. A*, 41(5):2621–2630, March 1990.

[12] A. Morinaga nd F. Riehle, J. Ishikawa, and J. Helmcke. A ca optical frequency standard: frequency stabilization by means of nonlinear ramsey resonances. *Applied Physics B*, B48(2):165–171, Feb. 1989.

[13] F. Plumelle, M. Desaintfuscien, and M. Houssin. Prospect for a Ca^+ optical frequency/wavelength standard. *IEEE Transactions on Instrumentation and Measurement*, 42(2):462–466, April 1993.

[14] D. J. Wineland and W. M. Itano. Laser cooling of atoms. *Phys. Rev. A*, 20: 1521-1540, 1979.

[15] J.D. Prestage, G.J. Dick, and L. Maleki. New ion trap for frequency standard applications. *J. Appl. Phys.*, 66(3):1013–1017, August 1989.

[16] J.D. Prestage, R.L. Tjoelker, G.J. Dick, and L. Maleki. Ultra-stable Hg^+ trapped ion frequency standard. *J. Mod. Opt.*, 39(2):221–232, February 1992.

[17] J.D. Prestage, R.L. Tjoelker, G.J. Dick, and L. Maleki. Long term stability of Hg^+ trapped ion frequency standards. In *Proc. 48th Annual International Frequency Control Symposium*, pages 755–760. IEEE Publication 94CH3446-2, 1994.

[18] N.F. Ramsey. A new molecular beam resonance method. *Phys. Rev.*, 76:966, 1949.

[19] F. Riehle, I. Schnatz, A. Celekov,T. Kisters,K. Zeiske,G. Zinner, and J. Helmcke. Calcium optical frequency standard. In *1994 Conference on Precision Electromagnetic Measurements Digest (Cat. No.94CH3449-6)*, 1994.

[20] D. Schnier, A. Bauch, R. Schröder, and Chr. Tamm. Ion trap research at PTB. In *Proc. of the 6th European Frequency and Time Forum*, March 1992.

[21] D.J. Seidel and L. Maleki. Progress toward the development of a a ytterbium ion standard. In *Proc. 48th Annual IEEE International Frequency Control Symposium*, pages 761–768. IEEE Publication 94CH3446-2, 1994.

[22] D.J. Seidel and L. Maleki. Efficient quenching of population trapping in excited Yb^+. *Phys. Rev. A*, 51(4), April 1995.

[23] Kazuhiko Sugiyama and Jun Yoda. Study of Yb^+ trapped in a rf trap with light buffer gas by irradiation with resonant light. *IEEE*

Trans. on Instr. and Meas., 42(2), April 1993.

[24] Chr. Tamm. A tunable light source in the 370 nm range based on an optically stabilized frequency-doubled semiconductor laser. *Appl. Phys. B*, B56(5):295, May 1993.

[25] R. L. Tjoelker, J. D. Prestage, G. J. Dick, and L. Maleki. Long term stability of Hg^+ trapped ion frequency standards. In *Proc. 47th Annual International Frequency Control Symposium*, pages 132–138. IEEE Publication 93CH3244-1, 1993.

[26] R. L. Tjoelker, J. D. Prestage, G. J. Dick, and L. Maleki. Recent stability comparisons with the JPL linear trapped ion frequency standards. In *Proc. 1994 IEEE International Frequency Control Symposium*, pages 739–743. IEEE Publication 94CH3446-2, 1994.

[27] A. Williams, D.J. Seidel, and L. Maleki. Generation of 369.4 nm radiation by efficient doubling of a diode laser. In *OSA Proceedings on Advanced solid-State Lasers*, volume 15, pages 250–252, 1993.

[28] D.J. Wineland, J.C. Berquist, J.J. Bollinger, W.M. Itano, D.J. Heinzen, S.L. Gilbert, C.H. Manney, and M.G. Raizen. Progress at NIST toward absolute frequency standards using stored ions. *IEEE Trans. Ultrason., Ferroelect., Freq. Contr.*, 37:515–523, 1990.

[29] M. E. Poitzsch, J. C. Berquist, W. M. Itano, and D. J. Wineland. Progress on a cryogenic linear trap for $^{199}Hg^+$ ions. In *Proc. 1994 IEEE International Frequency Control Symposium*, pages 744–746. IEEE Publication 94CH3446-2, 1994.

[30] M. Knoop, M. Vedel, and F. Vedel. Investigations of a r.f. stored Ca^+ ion cloud and observation of S-D forbidden transitions. *J. Phys. II France*, 4:1639–1650, 1994.

SESSION 1

RESONATORS I

Chairman : M.J. UNDERHILL

ACCELEROMETRIC AND MAGNETIC SENSITIVITIES OF BVA RESONATORS

J.J. Boy**, F. Deyzac*, A. Berthaut**

** LCEP / ENSMM - Route de Gray - La Bouloie - 25030 Besançon Cedex - FRANCE
* Office National d'Etudes et de Recherches Aérospatiales - BP 72 - 93322 Châtillon Cedex - FRANCE

ABSTRACT

The accelerometric and magnetic sensitivities of a quartz resonator are commonly attributed to the design of the resonator, for the main part. Since a few years, it is well-known that G-sensitivity can reach zero if the geometry of the structure (including blank geometry, orientation of the mounting, number of supports, ...) leads to the perfect symmetry around the three axes.

In this paper, we show that the BVA resonator has been well designed to exhibit a very low G-sensitivity (largely lower than $1*10^{-10}$/G). Furthermore, by finite element analysis, we have quantified the influence of some symmetry defects due to manufacturing process (as orientation of the mounting, shift between the vibration area center and the blank center, ...) on the accelerometric sensitivity. At least, we compare these theoretical results to experimental values obtained by two methods :

- measurements under random vibrations in the range [20 - 100 Hz],
- measurement of the frequency shift due to a 2G-tipover test around a given axis.

The results concerning the magnetic sensitivity have been extracted from experiments performed under low magnetic field (< 15 Gauss) on BVA and various kinds of resonators. For the BVA one, we show that the magnetostriction effect (which is a change in the dimensions of a ferromagnetic material as the resonator supports) is as negligeable as the other magneto-mechanic effects. So, this resonator exhibits a magnetic sensitivity lower than $1*10^{-12}$ a Gauss on each axis.

Keywords : Quartz crystal, BVA resonators, short term stability, accelerometric sensitivity, magnetic sensitivity.

1. INTRODUCTION

Among the improvements implemented these last years in the industrialization process of "tactical" mini-BVA (SC-cut 10 MHz third Overtone) in the LCEP laboratory, ones of them have leaded to tentatively reduce the accelerometric sensitivity. That is why some keypoints have been systematically discussed and tested, particularly in a previous paper [1], in order to define how the performances can be reproduced according to different specifications.

As indicated in [2] and [3], the principal keypoints of the design work towards perfecting the symmetries around Z-axis and the median plane. For that, the resonator is composed of two "condensers" electrode-holders on both sides of the active part, these three quartz pieces being holdered by four rigid clips associated to a symmetrical suspension consisting in 8 springs (i.e. 4 pairs). Furthermore, the vibrating part of the resonator is linked to its external support by four rigid and short quartz bridges.

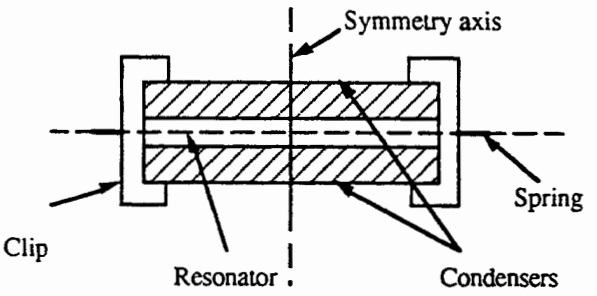


Fig. 1 : diagram of the mini-BVA (called BVA_2)

Among the tens of SC-cut BVA resonators realized a year, a few of them have been used, this time, to measure their accelerometric sensitivities as well as their frequency shift due to the application of a small magnetic field. At least, we have pointed out all the main parameters of the manufacturing process to better understand the influence of selected geometrical defects and so to improve the realiability of the product as well as these performances.

2. BRIEF REVIEW ON FINITE ELEMENTS METHOD MODELLING

With the support of DRET, ONERA, engaged in studies dedicated to reduce sensitivities of quartz resonators, has developped an appropiate simulation tool to better understand the sensitivity mechanisms of resonators and evaluate the influence of manufacturing defects. This tool involves a finite elements method for modelling the resonator and a theoritical stress/frequency model [4].

The stress state, computed by a FEM modelling, is related to resonator frequency variation. The model takes into account the crystallographic cut (φ, θ), the crystal orientation (ψ) in the section plate and calculates the wave polarization directions and the rotated coefficients of elasticity.

Here, we recall just that the main hypothesis for calculation is that the frequency shift depends only on stress variations (T_i) in the central zone of the vibrating area (i.e. center of plate without defect) because the resonator is designed to confine energy at its center (with, for example, the realisation of a convexity in one face of the plate). So :

$$\frac{\Delta f}{f} = \sum_{i=1}^{6} \tilde{k}_i T_i$$

where T_i and the coefficients $\tilde{k}_i$ are expressed in the FEM reference frame (Oxyz) whose the axes coincide with those obtained after the triple rotation of the cristallographic reference frame (Ox"y"z").

So, for the theoritical fractional frequency deviations presented below, the used axes X, Y, Z are such as indicated in the Fig. 2.

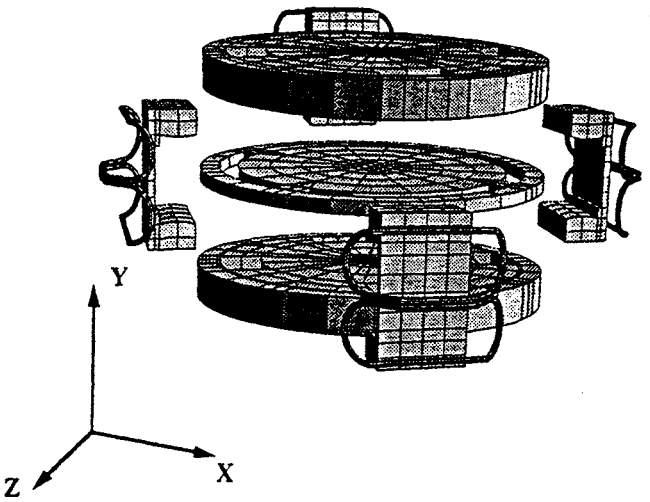


Fig. 2 : Mesh of BVA resonator

If in the case of QAS resonator [5], it was not necessary to consider an anistropic finite element model, for the BVA one, we have to take into account all sources of dissymetry such as curvature and anisotropy of the blank.

The quartz crystal frequency change due to acceleration $\vec{a}$ is usually expressed as a normalised frequency change :

$$\frac{\delta f}{f} = \vec{a} \left\| \vec{\Gamma} \right\|$$

where $\vec{\Gamma}$ is the acceleration sensitivity vector of resonator. Measurements along 3 mutually perpendicular axes X, Y, and Z are required to characterize $\vec{\Gamma}$ by its components γ_x, γ_y, γ_z . Its magnetude is related to its components by :

$$\left\| \vec{\Gamma} \right\| = \sqrt{\gamma_x^2 + \gamma_y^2 + \gamma_z^2}$$

3. INFLUENCE OF MECHANICAL DEFECTS ON THE ACCELEROMETRIC SENSITIVITY : THEORITICAL APPROACH

The simulation performed on the previous model (Fig. 2) which presents a perfect symmetry around the three axes exhibits a quasi-insentivity to the acceleration. So, we verified that this BVA structure can be ideally "super-symmetric" and, consequently, non sensitive to the acceleration.

Without speaking about the mechanical defects, this model is identical to the mini-BVA manufactured in the laboratory, except that we have not introduce the convexity on the active part of the resonator. Indeed, we have just considered the resonator as a plate with faces perfectly parallel.

So, to tentatively correlate the simulation to the experimental results, it appears necessary to introduce in the FEM modelling :

- in one hand, the convexity (which insures the energy trapping ideally in the center of the resonator) with a given and realistic value of the curvature radius,
- and, in the other one, various mechanical (or geometrical) defects to quantify their own influences on the accelerometric sensitivity.

3.1. Radius of curvature :

BVA blank is plano-convex for energy trapping. Curvature is an intentional dissymetry of blank which increases acceleration sensitivity of the perfect BVA resonator (Fig. 3).

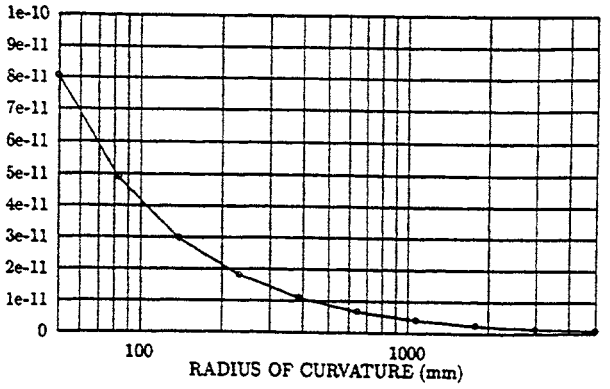


Fig. 3 : Acceleration sens. vs blank curvature

For the choosen radius curvature in the mini-BVA, the gamma vector magnitude is :

γ_x (/g)	γ_y (/g)	γ_z (/g)	$\vec{\Gamma}$ (/g)
$1.46*10^{-11}$	$-0.54*10^{-11}$	$0.96*10^{-11}$	$1.82*10^{-11}$

It appears as evident that the manufacturing of a convex-convex resonator can be so reduce drastically the frequency shift due to acceleration. However, such a realisation is not easy to perform and induces, as we can see below, other dissymetries which increase this sensitivity.

3.2. Orientation of the mounting :

We calculate here the influence of an angular error (called azimuth angle ψ) the direction of bridges and the X- (or Z-) axis which are the projections of the crystallographic axes in the plane of the blank.

As the law $\vec{\Gamma} = f(\psi)$ is $\pi/2$-periodic, Figure 4 depicts $\vec{\Gamma}$ variations of BVA resonator only when azimuth angle ψ ranges from -45° to +45°.

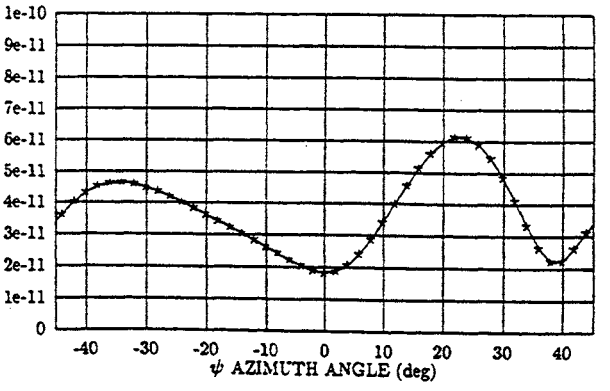


Fig. 4 : Acceleration sens. vs azimuth angle ψ.

We estimate that the ultrasonic cutting out (used to realize the four "bridges") can introduce an angular error on ψ equal to $\pm 2°$ at the maximum.

3.3. Offcentering of the curvature :

As we said below, convexity is necessary for a good trapping of acoustic energy in the *middle of the resonator*, but if the axis of curvature is not matched with this of resonator, we will assume that vibrating spot follow the curvature (Fig. 5).

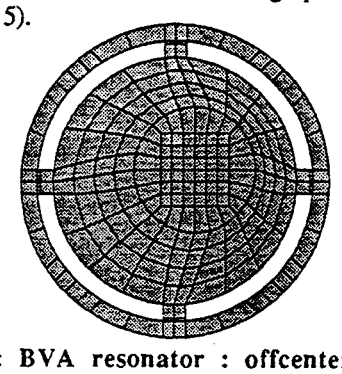


Fig. 5 : BVA resonator : offcentering of curvature (central part).

Fig. 6 depicts variations in $\vec{\Gamma}$ versus curvature offcentering ranging from 0 to 1 mm in Z direction.

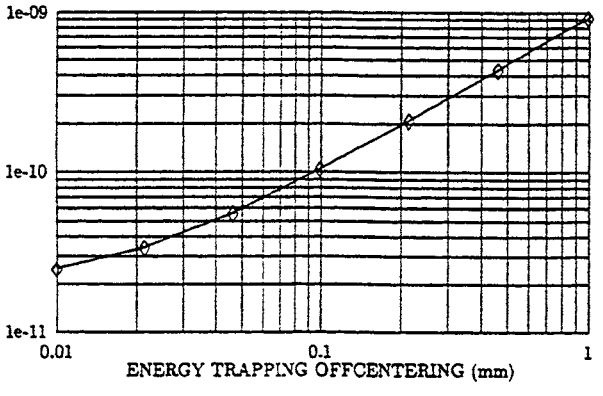


Fig. 6 : $\vec{\Gamma}$ (/g) versus curvature offcentering.

Modelization shows that it is not geometry defect (curvature offcentering) but integration defect (trap offcentering) which induces the most important effect.

3.4. Offcentering of the grooves machining :

The ultrasonic machining requires a good matching between axis of the tool and axis of the blank. Fig. 8 presents a BVA resonator with 0.3 mm machining offcentering in Z direction.

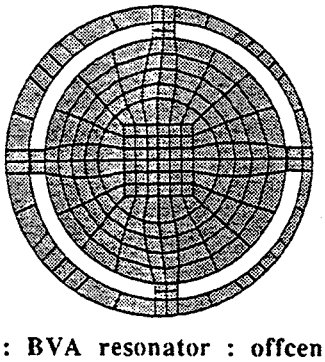


Fig. 7 : BVA resonator : offcentering of grooves machining (central part).

At least, Fig. 8 discribes variations in $\vec{\Gamma}$ when offcentering of grooves machining ranges between 0 and 0.3 mm.

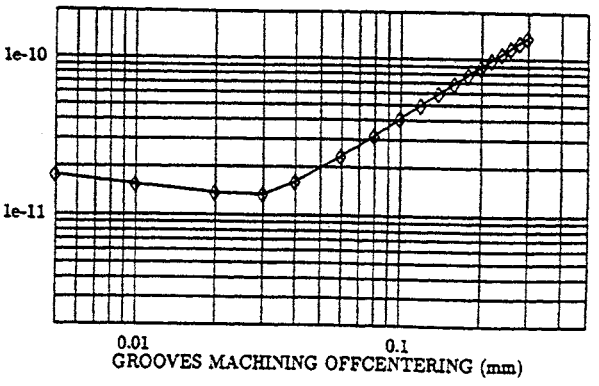


Fig. 8 : $\vec{\Gamma}$ (/g) versus grooves machining offcentering.

3.5. : Conclusion of modelizations :

Modelization of perfect (i.e. no machining defect) BVA resonator leads to a value of $1.8*10^{-11}$ /g for gamma vector magnitude. As we will see below, this value is 3 to 4 times lower than the best experimental results but the simulation of some defects (table 1) leads to nearest values.

Table 1 : $\vec{\Gamma}$ (/g) versus some defects

BVA res.	γ_X (/g)	γ_Y (/g)	γ_Z (/g)	$\vec{\Gamma}$ (/g)
"Perfect"	1.5	-0.5	1.9	$1.8*10^{-11}$
$\Delta\psi = 10°$	1.8	2.4	1.5	$3.4*10^{-11}$
δ curv.= 100 µm	2.9	-8.8	4.9	$10.5*10^{-11}$
δmach.= 220 µm	-0.5	-2.8	-9.4	$9.8*10^{-11}$

4. MAGNETIC SENSITIVITY

For new applications as, for example, precision orbit determination, some perturbation sources have not to be considered as negligible. A magnetic sensitivity of about 10^{-11}/gauss, is unacceptable in high precision measurements [6], [7]. With the BVA resonator, we gain an order of magnitude essentially due to the "super-symmetry" of the structure. Indeed, the magnetostriction effect (which is a change in the dimensions of a ferromagnetic material as the resonator mounts) does not introduce sufficiently high stresses in the active part of the resonator. The reasons exposed above about the very low accelerometric sensitivity of this resonator explain also its lower magnetic sensitivity. If a offcentering of the curvature does not seem to influence this kind of sensitivity (quartz is a non-magnetic material), the magnetostrictif effect on the feedthrough (which is in Kovar), as for it, can be drive the ring supporting the springs and create differential mechanical stresses in the resonator.

We propose some experimental results in the next paragraph.

5. EXPERIMENTAL RESULTS

The experimental results presented below concern 10 BVA resonators selected among 4 different series manufactured in the LCEP laboratory in 1993. According to our general presentation, we have pointed out the mechanical defects as offcentering of the active part or of the vibrating area, orientation of the mounting, ... The table 2 indicates these geometric defects :

Table 2 : Geometrical defects

Serial Number	Geometrical defects
30	Parallelism defect due to cylindrical rectification.
31	Small broken quartz induced by Ultrasonic machining
33	- No -
34	offcentering of the grooves machining of about 300 µm

Furthermore, during an industrialization process, it is not possible to check the parallelism of the two faces of each quartz disk at better than 2 µm (including the envelope of the curvature). Moreover, the tightening of the clips is not perfectly identical from one clip to other one. Generally speaking, in this kind of process, for time reasons, it is not always possible to use the best mechanical checking.

Most of resonators have been measured in short term stability in the oscillator used for accelerometric sensitivity under random vibrations. The mean value for $\sigma_y(\tau)$ for 3 s is about $3*10^{-13}$.

5.1. Accelerometric sensitivity :

If the reference frame used for the theoretical calculations is defined from the crystallographic axes, this one used for the measurements can not be the same because the resonator is enclosed and so, we have not access to the axes inside the cap. The following diagram presents these "new" axes slightly different for measurements under vibrations and 2G-tipover.

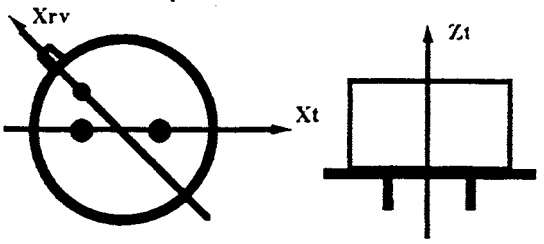


Fig. 8 : Different frames
Xrv, Yrv, Zt : under random vibrations
Xt, Yt, Zt : for 2G-tipover

We have performed two kinds of measurements :
- measurements under random vibrations in the range [20 - 100 Hz],
- measurement of the frequency shift due to a 2G-tipover test around a given axis.

The following table summarizes the results obtained for these 10 resonators.

Table 3 : accelerometric sens. (in 10^{-10}/g)

	Under Random vib.				2G-tipover test			
	Xrv	Yrv	Zt	Γ	Xt	Yt	Zt	Γ
30.01	1.0	0.56	1.19	**1.65**	0.7	0.15	0.9	**1.15**
30.03	0.56	1.25	0.28	**1.40**	0.7 0.89	0.6 0.45	1.0 0.61	**1.36** **1.17**
30.12	0.44	0.63	1.25	**1.47**	0.6	0.25	0.4	**0.76**
31.03	0.9	0.56	1	**1.45**	0.85	0.5	0.7	**1.21**
31.06	O.9	0.8	0.35	**1.25**	0.5	0.4	0.4	**0.75**
31.08	0.8	0.7	0.56	**1.20**	0.5 0.5	0.03 0.17	0.6 0.42	**0.78** **0.68**
33.05	0.67	0.56	0.56	**1.04**	0.4 0.48	*0.02* *0.03*	0.9 0.26	**0.98** **0.55**
33.06	0.8	0.9	0.56	**1.32**	0.5 0.2	0.55 0.55	0.4 0.3	**0.84** **0.66**
34.03	1.0	0.7	1	**1.58**	*0.1*	*0.9*	0.9	**1.28**
34.06	0.63	0.71	1.27	**1.59**	*0.1*	*1.0*	0.7	**1.22**
	Mean : $1.4*10^{-10}$/g				**Mean : $1.0*10^{-10}$/g**			

The 30.03 resonator has been also tested under sinus vibrations at 2 Hz (vibration level : 0.32 g peak). The results are the following :

Xrv	Yrv	Zrv	Γ
2 (!)	1.25	0.7	**2.46**

The previous table shows that the 2G-tipover test is more "optimistic" that 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, next figure (Fig. 9) shows that 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).

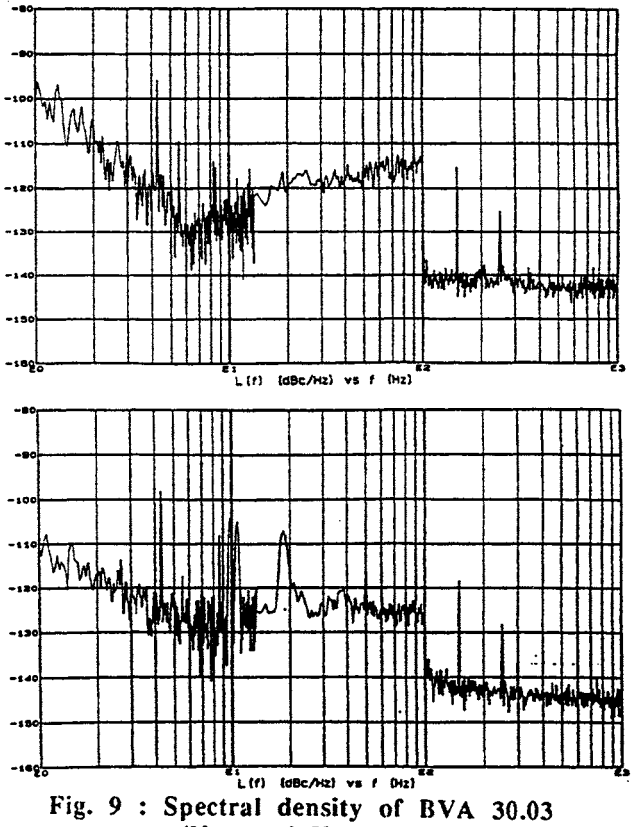


Fig. 9 : Spectral density of BVA 30.03 (Yrt- and Xrt-axes)

The following figure (Fig. 10) illustrates in what conditions we have performed the 2G-tipover test : five successive tipovers of the resonator in a mechanical structure allowing rotation around an horizontal axis. The frequency difference between the oscillator under test and a reference is recorded in a 3 s gate time. We record 10 points by step and wait 10 s for a slow tipover. The results given in a previous table correspond to the mean value of the 5 tipovers for each axis.

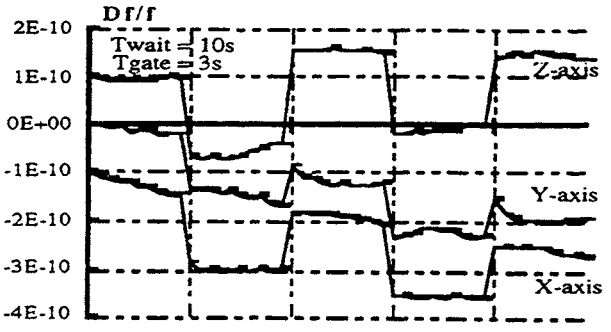


Fig. 10 : 2G-tipover on 30.03 BVA resonator

Finally, we have performed some measurements on convex-convex resonators. Unfortunately, the greatest difficulties of realisation of such resonator introduce again other symmetry defects. So, as we can see in the table 4, the results are just slightly better.

Table 4: accelerometric sens. (in 10^{-10}/g) for convex-convex resonators. (2G-tipovers tests)

N°	Xt	Yt	Zt	Γ (10^{-10})
35.02	0.16	0.14	0.40	0.45
36.05	0.11	0.08	0.42	0.44
36.01	0.52	0.12	0.77	0.94

So, in conclusion of these accelerometric sensitivity measurements, we can note that the "best" realized BVA resonators exhibit an accelrometric sensitivity 3 times higher than the results of modelization performed on "perfect" structure. In fact, we assume that it is up to now quasi impossible to realize "perfect" resonators. Particularly, the grinding machining which is a handmade operation introduces some defects in the parallelism between the two faces and so a small offcentering of the curvature (of about 40 to 50µm).

5.2. Magnetic sensitivity :

The next measurements are performed with an oscillator whichfor the resonator is deported of about 20 cm. So, only the resonator and its oven are in the center of three pairs of Helmoltz 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 drived by a computer which can inverse its polarity in the coils. A 2 A current leads to a measured magnetic field in the center of the Helmoltz coils of :

8.8 Gauss along Xt
8 Gauss along Yt
10 Gauss along Zt.

As for the accelerometric sensitivity measurements, 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. 12. The mean magnetic sensitivity is also calculated as the modulus of a vector in the reference frame previously called Xt, Yt, Zt.

Our main results are presented in the table 6 whichfor we indicate a maximum of magnetic sensitivity in accordance of the previous measures performed by R. Brendel in [7] and [8].

Table 5 : Magnetic sensitivites ($\Delta f/f$ expressed in 10^{-12} by Gauss)

	Xt	Yt	Zt	Δf/f
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 : $3.1*10^{-12}$/Gauss			

The results presented here include the effect of a residual magnetic field on the electronics of about 1 Gauss (i.e. 10% of a magnetic field applied on the resonator). We have verified that this electronics induces a non negligeable frequency shift when the field is created along X-axis. So, the sensitivity of the resonator is slightly better than indicated.

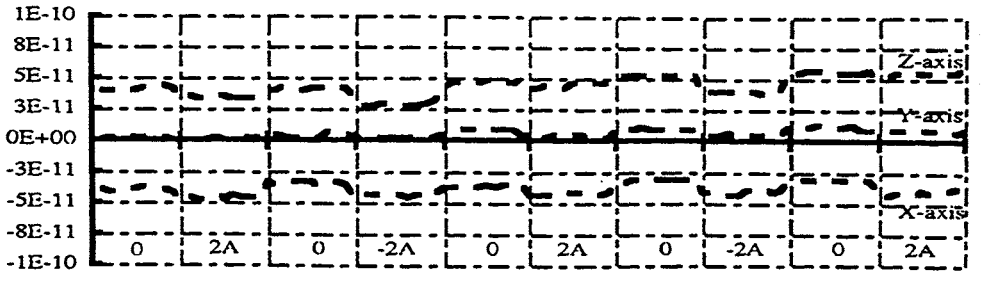


Fig. 11 : Magnetic sensitivity of the 30.03 BVA resonator (Cycle : 0/2A/0/-2A/ ...)

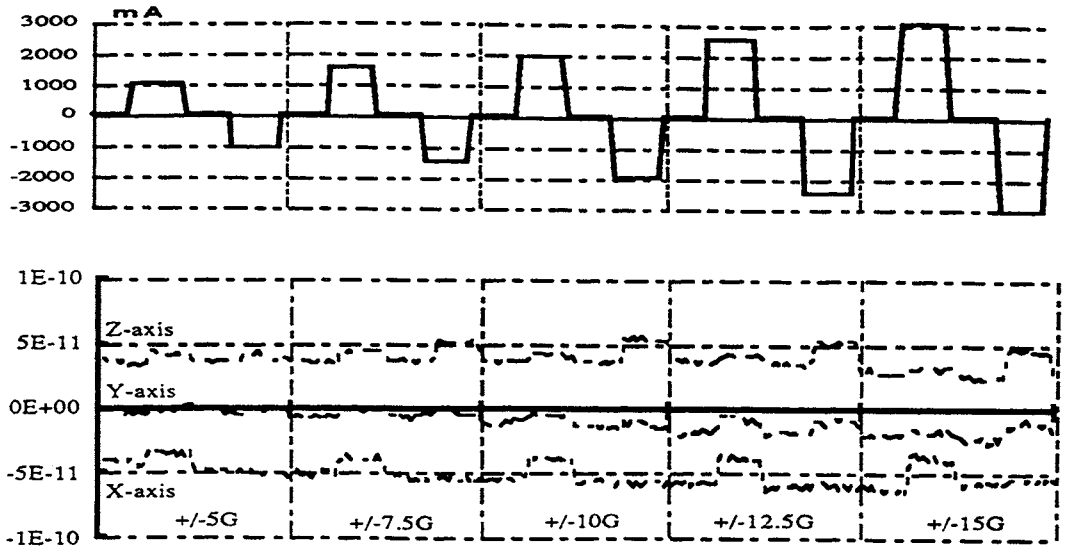


Fig. 12 : Magnetic sensitivity of the 30.14 BVA resonator

The figure 12 illustrates the frequency variations obtained on the BVA 30.14 after different magnetic field values. For each current level, we have inversed its polarity to qualify the eventual dissymetry observed by several authors for "classical" resonators. The frequency shifts are summarized in the following table (table 7).

Table 6 : Magnetic sensitivity of the 30.14

	Xt	Yt	Zt	Δf/f
+5 Gauss	1.4	0.5	1.25	2
- 5	-0.4	0.3	0.8	1
+7.5	1.8	0.9	1.2	2.4
-7.5	0.1	0.3	1.9	1.9
+10	1.8	0.5	0.6	2
-10	-0.1	0	1.9	1.9
+12.5	1.7	1	0.3	2
-12.5	0.1	0.9	1.4	1.7
+15	1.8	0.2	0.2	1.8
-15	0.3	0.9	1.4	1.7

5.3. Comparaison with other resonators :

For accelerometric and magnetic sensitivities, we have measured BVA_4 (with electrodes deposited on the active part) AT-cut resonators (10 MHz - 3rd overtone) mounted in a "classical" mount. In this case, Zt is also perpendicular to the plate as indicated below.

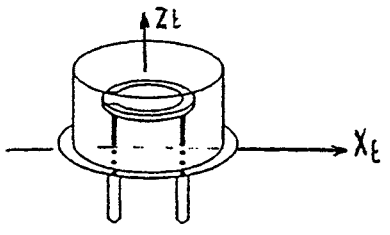


Fig. 13 : Scheme of a "classical" BVA_4 resonator

Typical results are presented in the two following tables and partially illustrated by Fig. 14 for which we note that the resonator is quasi insensitive to the magnetic field when the polarity of the current is inverted. The greatest frequency shift is obtained when the field lies in the plane of the support (Compass effect ?)

Table 7 : Accelerometric sens. for 2 BVA_4 (to 1 g)

Xt	Yt	Zt	Global ($*10^{-10}$)
1.6	5.7	2.5	6.5
2.9	2.5	2.7	4.7

Table 8 : Magnetic sens. for BVA_4 (to 1 Gauss)

Xt	Yt	Zt	Global ($*10^{-12}$)
22.5	125	50	136

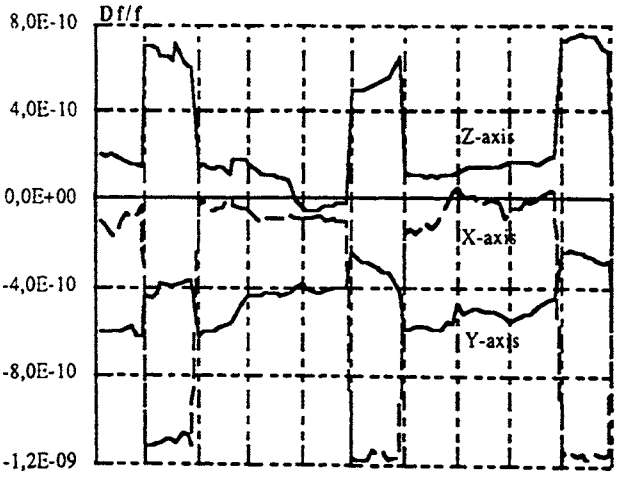


Fig. 14 : Magnetic sensitivity for "classical" AT-cut.

5.4. Calculation and measurements of the mechanical resonances :

We have noted a good correlation between measurements and FEM dynamic analysis as indicated in the following table :

Table 9 : Mechanical resonances

	FEM analysis	Measurements under sinus vibrations
Along X_{TV}	1.16 KHz	1 and 1.2 KHz
Along Y_{TV}	- idem -	- idem -
Along Z_{TV}	2.1 KHz	1.8 KHz

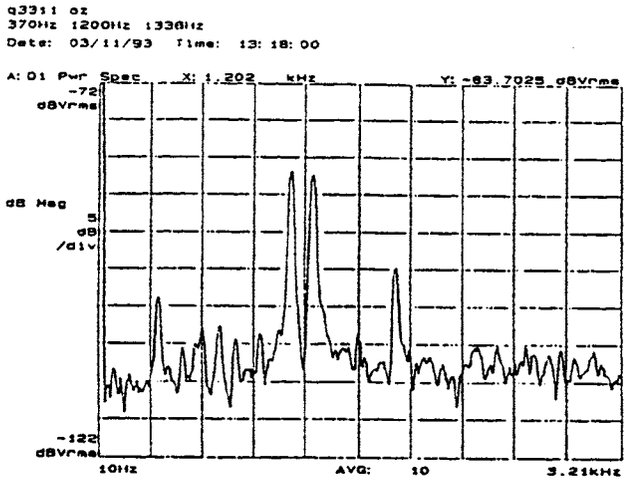


Fig. 15 : Mechanical resonances measured under vibrations

CONCLUSION

A systematic study of the design and the manufacturing of the "tactical" mini-BVA has permitted us to achieve yet some improvements in its realization.

The determination of the keypoints of the design is completed. They concern, for almost, the problems of symmetry of the structure, around Z-axis and the median plane of the resonator.

We show that FEM is a very helpful tool to understand mechanical behaviours of resonators. It is used to quantify the influence of defects associated to manufacturing process and it is very convenient to verify improvements efficiency.

We have so verified that the mini-BVA is nominally "perfect" concerning the accelerometric sensitivity. At least, we show that the main defect inducing a largest frequency shift under acceleration is the offcentering of the vibriting area due to offcentering of the curvature, but also offcentering of the electrodes, dissymetry of this vibrating area and, at least, some defects (as dislocations, for example) in the quartz raw material.

ACKNOWLEDGEMENTS

The authors would like to thank the Centre National d'Etudes Spatiales (CNES) which supports a part of this study and G. Robichon from **bva Industrie** for the accelerometric measurements under random vibrations.

REFERENCES

1. J.J. BOY, A. BERTHAUT, G. RENARD, R. ROY, D. THIEBAUD, G. ROBICHON, C. LONGET : "Characterisation of BVA in industrialisation process" - 7th Europ. Freq. and Time Forum (1993), pp. 57, 62.
2. R.J. BESSON : "A new electrodeless resonator design" - 31st Annual Freq. Control Symp. (1977), pp. 147, 152
3. R.J. BESSON and al., French patent 8509097 and US patent 4.705 982 (11-10-1987)
4. F. DEYZAC : "Finite elements method modelling of accelerometric and barometric sensitivity of quartz resonators" - 7th Europ. Freq. and Time Forum (1993), pp. 291, 295.
5. J.P. AUBRY, J.C. CRAVEUR, F. DEYZAC, M. FRAGNEAU, D. JANIAUD : "Analyse des mécanismes de sensibilité accélérométrique et barométrique des résonateurs à quartz de type QAS" - 1st Europ. Freq. and Time Forum (1987), pp. 195, 203.
6. F. DEYZAC : "Magnetic sensitivity of quartz oscillators" - 4th Europ. Freq. and Time Forum (1990), pp. 255, 258.
7. R. BRENDEL, C. EL HASSANI, M. BRUNET, E. ROBERT : "Influence of magnetic field on quartz crystal oscillator" - Proc. 43rd Ann. Symp. on Freq. Control (1989), pp. 268, 274.
8. R. BRENDEL, B. CRETIN, A. EL HABTI : "Origin and measurement of quartz resonator magnetic sensitivity" - 8th Europ. Freq. and Time Forum (1994), pp. 215, 230.
9 R. BRENDEL : Private communication "Rapport de mesures" (April 1993).

A Novel Technique for Measuring Drive Level Dependence of Quartz Units

Mitsuaki Koyama[+], Yasuaki Watanabe[++], Hitoshi Sekimoto[++] and Yoshimasa Oomura[++]

[+] Nihon Dempa Kogyo Co., Ltd., Sayama city, Saitama 350-18, Japan
[++] Tokyo Metropolitan University, Hachioji city, Tokyo 192-03, Japan

Abstract - **A new method for precisely measuring the resonance frequency - drive level dependence (DLD) of quartz crystal resonators has been developed. This technique is based on the oscillator method using the transistor Colpitts circuit. The crystal current that is controlled with the base-bias voltage, is detected with a non-contact current probe. The advantage of this method is that the resonant frequency can be measured using a frequency counter directly; thus, the reproducibility and the accuracy of the measured data are better than those obtained with passive methods. The measurement resolution of the experimental system attained to $1x10^{-11}$ for the frequency, and to 1μ A for the crystal current, respectively. The detailed explanations for the proposed technique and some experimental results will be shown in this paper. In particular, experimental results for the SC-cut crystal resonators for high-performance OCXOs clearly indicate that this technique is extremely effective for detection of the DLD minute frequency jumps that reduce their oscillation frequency stability.**

I. INTRODUCTION

The demand for mobile communications equipment such as mobile telephones and pagers has been increasing because such equipment allows users to communicate easily with anyone at any time. Mobile communications equipment is largely classified into mobile telephones & calling devices, electronic devices for a base station, communications equipment for artificial satellites and measuring instruments for inspecting these kinds of com-munications equipment. For electronic devices for base stations and communications equipment for artificial satellites, a quartz-crystal - based oscillator is used in an oven for frequency standard. The specified frequency stability required for this oscillation circuit is 1 x 10^{-9}/day to 10 x 10^{-9}/day. But with the increasing demand for communication, the frequency stability of 5 x 10^{-10}/day has recently become necessary.

In general, an OCXO (Oven-controlled crystal oscillator) using an SC-cut crystal unit has been used to meet this requirement, but in some cases, oscillating the OCXO causes a frequency jump on the order of 10^{-9} or less.[1] The frequency jump is considered to be caused by the oven drift, the influence of the spurious response providing abrupt temperature characteristics, the fluctuation of parameters on the active oscillation circuit side, and the influence of the current characteristics of a crystal unit.[1] The authors have conducted various experiments, assuming that this frequency jump results from the fluctuation of current in the crystal unit. This phenomenon is a kind of activity dips[1]. However, measurement methods [2] and [3] for the current in a crystal unit have not allowed accurate measurement of the fluctuation of 1 x 10^{-10} with the required reproducibility because of the limited performance of the measuring instruments, preventing the determination of the true cause of frequency jump. A measurement method for the frequency vs. current characteristics of a crystal unit with higher accuracy and sufficient reproducibility is required for revealing the cause of frequency jump and designing a highly reliable crystal oscillator.

We propose measurement method in which the current in an oscillating crystal unit is detected by means of a current probe and the current is controlled by changing the oscillation-circuit bias (base) voltage. A merit of this measurement method is superior reproducibility compared with the conventional resonance measurement method because oscillation frequency can be measured directly by a frequency counter. Another merit is the sufficient accuracy of frequency. This paper first gives an outline of the measurement method and the measurement system and shows an actual measurement example using an SC-cut crystal unit. Based on these results, we will further study the OCXO with the problem of frequency jump to show that proper selection of crystal unit driving current permits complete elimination of frequency jump even if the same crystal unit is used.

II. MEASUREMENT SYSTEM DESCRIPTION

This section explains the principle for measuring the current characteristics of a crystal unit using the crystal oscillation circuit as well as the measurement system.

A. *Measurement of the Frequency vs. Current Characteristics of a Crystal Unit Using the Oscillation Circuit*

Fig. 1 illustrates the concept of the measurement of current characteristics using the crystal oscillation circuit In this experiment, it is assumed that the transistor Colpitts oscillator is used. It is known that the circuit gain changes in accordance with the base bias potential in the Colpitts oscillator if the circuit constants are kept constant on the emitter and collector sides and the collector supply voltage is constant. Providing that the nonlinearity of the active element exerts little influence, chang in the base bias potential allows the high-frequency current in a crystal unit to be varied continuously. Holding other values such as collector supply voltage constant and varying the base bias potential only slightly, the current vs. oscillation-frequency characteristics of a crystal unit can be measured. As direct determination of the current in a crystal unit from this bias voltage is impossible, a high frequency-current probe (non-contact type)[1] is placed in contact with the lead wire of a crystal unit in an oscillation circuit to read the current in the crystal unit.

Fig. 2 shows an example of current vs. bias voltage characteristics measurement using the circuit shown in Fig. 1. This figure demonstrates that the relationship between the current in a crystal unit and the bias voltage is linear (0.8 mA/ V). (Our measurement results are the same as calculation results based on the simulation by *SPICE*. The details are not included in this paper.) In other words, our experiments have revealed that it is possible to vary the current in a crystal unit minutely and linearly by changing the bias voltage change and that this technique can be utilized for the frequency vs. current characteristics of a crystal unit.

B. *Measurement System*

Fig. 3 is a block diagram of the measurement system based on the above measurement principle. The output from the crystal oscillator is multiplied 1,000 times, using a multiplier, and the small fluctuation of the frequency of the crystal oscillator can be detected using the heterodyne detection method. The bias potential mentioned above is controlled by the DC voltage sweeper, and the current in the crystal unit and the oscillation frequency are printed out for each voltage value. The measurement is done automatically, using the PC shown in the figure.

III. MEASUREMENT RESULTS

This section shows the measurement results for the three SC-cut crystal units (No.1 to No.3) used in the measurement system in the previous section.

As the crystal units to be used are already incorporated

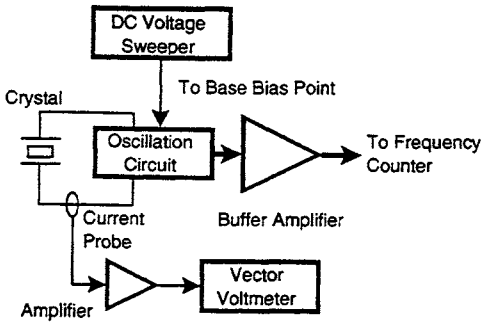


Fig. 1 Block diagram for measuring the resonance frequency vs. current characteristics of a crystal unit based on the oscillation circuit.

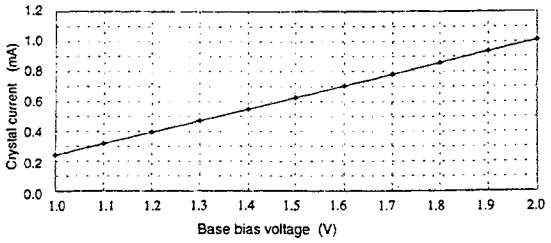


Fig. 2 Example of measuring the bias voltage and the current in a crystal according to the method shown in Fig. 1.
(Test sample: 3rd-overtone 5MHz crystal units)

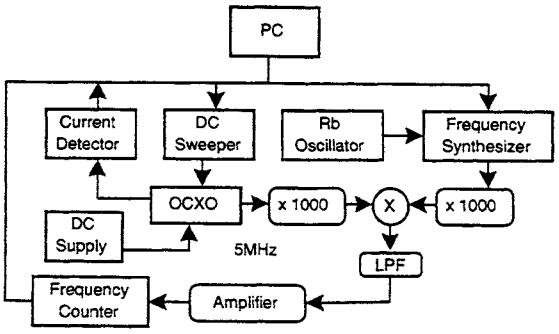


Fig. 3 Block diagram of a high-precision measurement system for measuring the frequency vs. current characteristics of a crystal unit based on an oscillation circuit.

TABLE I
SPECIFICATION OF THE RESONATORS UNDER TEST.

Item	Type/Value
Nominal frequency	5MHz
Holder type	HC-36/U
Overtone	3rd overtone
Quartz crystal	Synthetic quartz Unswept crystal
Infrared absorption	0.034
Number of etching channels	$12/m^{-6}$
Cutting angle	SC-cut
Equivalent-circuit Q	2.16×10^6
L1	5.5 H
C1	0.186 pF
C0	3.2 pF
C0/C1	17200
M	126

voltage of each oscillation circuit without using the common measurement circuit. Table 1 shows the specifications of the crystal units used.

Fig. 4 shows the change of oscillation frequency given when the current in a crystal unit is changed from 0.400mA to 0.600 mA. Our measurement results have revealed that there is no frequency jump due to the change in current in crystal units (Nos. 1 and 2). For crystal unit No.3, a frequency jump of approximately. 1.7×10^{-9} occurs when the current is approximately 0.5 mA, demonstrating that our measurement method is effective for the detection of a small frequency jump caused by the change in driving current. Compared with crystal units Nos.1 and 2, the oscillation circuit using crystal unit No.3 has larger frequency fluctuation for the change in temperature and at normal temperature[1]. As this current value is the preset value of the OCXO, the cause is considered to be the frequency jump shown in Fig. 4.

IV. Improvement of the Characteristics of OCXO Based on the Frequency vs. Current Characteristics of Crystal Units

This section explains the measurement results for the oscillation circuit in the case of adjusting the current in a crystal unit based on the result obtained in the previous section, as to the frequency fluctuation of the OCXO using an SC-cut crystal unit [1].

Fig. 5 shows the block diagram of the OCXO used for the measurement. The OCXO is composed of a transistor Colpitts oscillation circuit with an AGC, an oven, a temperature-control section, and a voltage regulator. The temperature in the oven is preset to be identical with the peak temperature of the frequency vs. temperature characteristic curve of the crystal unit to be used as the sample. In the previous report, measurement was made on the assumption that the current in a crystal unit is 0.5 mA. As a frequency jump has occurred near this current value in the crystal unit No.3, the current value has been lowered to approximately 0.45 mA for measuring the oscillation frequency in comparison with the ambient temperature and then making a comparison with the measurement results already obtained. The measurement temperature range is -10°C to +60°C, with 5°C as an increment or a decrement. In this experiment a rubidium atomic oscillator having the frequency stability of 2×10^{-11}/month is used as the frequency standard.

Fig. 6 shows the frequency vs. temperature characteristic measurement results for three OCXOs when the current in the crystal unit is 0.5 mA ± 0.01 mA. SC-cut crystal units Nos. 1 to 3 are used in each OCXO.

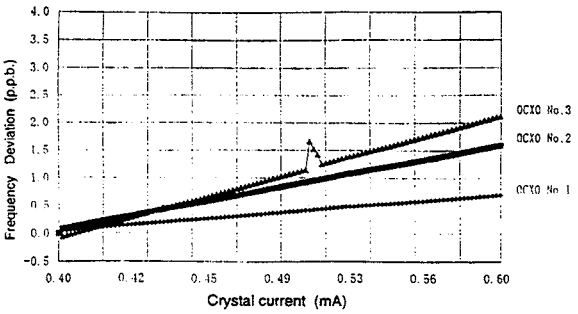


Fig. 4 Frequency vs. current characteristics of an SC-cut crystal unit measured by the system shown in Fig. 3.

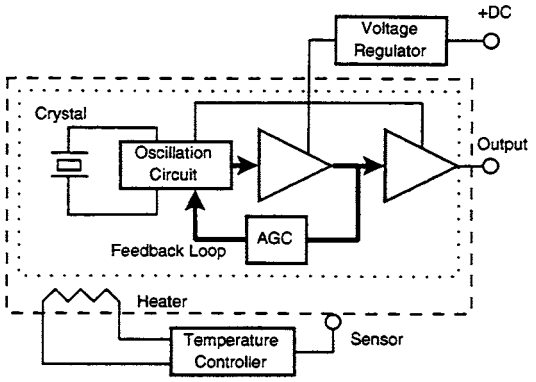


Fig. 5 Block diagram of a sample OCXO.

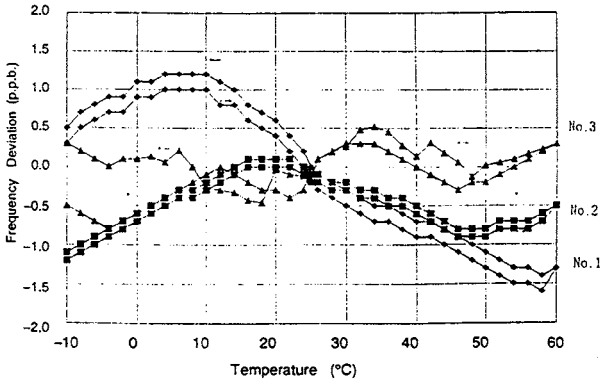


Fig. 6 OCXO frequency vs. temperature characteristic measurement result obtained when temperature is changed in the -> direction [1] (Current in the crystal unit: 0.5 mA). The frequency reproducibility of the No.3 curve is bad.

The figure shows that no frequency jump occurs for OCXO No.1 and OCXO No.2 and that the frequency vs. temperature characteristic curve is smooth. The reproducibility at the same temperature is within 0.2×10^{-9}. On the other hand, no remarkable frequency jump occurs for the OCXO No.3, but the oscillation-frequency reproducibility at the same temperature is 1×10^{-9} and thus inferior to the other two OCXOs.

Fig. 7 shows the frequency vs. temperature characteristic measurement results for the three OCXOs when the current in the crystal unit is 0.45 mA ± 0.01 mA. This current value has been preset so that no frequency jump occurs, judging from the measurement results in the previous section and the fact that the oscillation conditions differ little from those in the case of Fig. 6. This figure shows that the reproducibility of the oscillation frequency of the OCXO No.3 has become identical with that of the other two OCXOs. Properly setting the crystal-unit driving current and avoiding the frequency jump zone allows the design of an OCXO with outstanding reproducibility of the oscillation frequency.

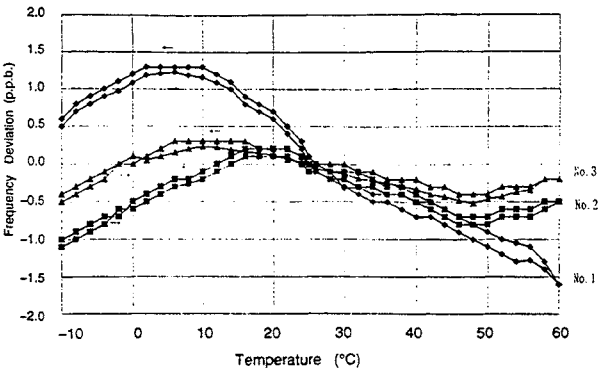


Fig. 7 OCXO frequency vs. temperature characteristic measurement result obtained when temperature is changed in the -> direction (Current in the crystal unit: 0.45 mA). The frequency reproducibility of the No.3 curve is identical to that of the other two curves.

V. CONCLUSION

We have proposed a measurement method using an oscillation circuit for the purpose of accurately measuring the resonance-frequency vs. current characteristics of an SC-cut crystal unit. This measurement method detects a small frequency jump causing the current drive level of an oscillating crystal unit to be changed by the bias voltage of the oscillation circuit.

This method has much better frequency resolution and measurement reproducibility than the conventional resonance measurement method. This measurement method has been applied to the SC-cut crystal unit used in the OCXO, which is poor in frequency stability, as reported in our last paper. It has been determined that the inferior frequency stability is attributable to the frequency jump caused by the change in current in a crystal unit. The frequency jump is on the order of 1×10^{-9} and was not detected by the conventional measurement method. Also, the measurement method proposed in this paper demonstrates that the design of an OCXO with excellent frequency reproducibility is possible if an oscillating crystal unit is driven at the proper current level. This measurement method is considered to be effective for the analysis of frequency jump, the cause of which has not been determined, during long-term frequency stability measurement.

REFERENCES

[1] Koyama M. et al. :Report of the commission on high-precision frequency generating circuit miniaturization, *IEE of Japan*, No.17-4 (July 1994, In Japanese)

[2] Watanabe Y. and Oomura Y.:*Journal of IEICE of Japan*,'86/11 Vol. J69-A No.11 (November 1986, In Japanese)

[3] Oomura, Tamaki and Watanabe:*Nation Convention Record, IEICE of Japan*, No.A-335 (March 1990, In Japanese)

[4] IEC Document:"Measurement method of drive level dependence (DLD) of crystal units", 49(Sec.)248 , (1993)

[5] Lindberg L. et al :Proc. 1994 IEEE Intern. Freq. Contr. Symp. , pp273-280 (1994)

CHARACTERIZATION OF LOSSES IN QUARTZ RESONATORS USING ELECTROMAGNETIC WAVES REFLECTION COEFFICIENT MEASUREMENTS

G.D Mansfeld*, A.D.Freik*, M.Mourey**
*Institute of Radioengineering and Electronics of Russian Academy of Sciences, 11, Mokhovaya str., 103907, Moscow, Russia.
**LCEP, Ecolè National Supèrieure de Mècanique et des Microtechniques, 25030, Besançon cedex,France

Abstract

Some results of the application of recently developed High Overtone Bulk Acoustic wave Resonator (HBAR) method of acoustic characterisation of thin layers and films to the study of losses in standard quartz resonators of various cuts in a wide frequency band are presented.A short discussion of the applicability of the method to this particular area of investigation together with specially calculated correction coefficients are given. Some important details of the experimental procedure connected with the use for measurements of the network analyzers are discussed. It is shown that in most cases the experimentally found resonator losses are mainly governed by Akhiezer mechanism and hence the quality factor of modern quartz resonators approaches to material limit.

Introduction

Recently a new method of measurement of acoustic losses and sound velocity in thin films and layers based on the use of HBAR composite resonator containing the investigated layer has been developed and demonstrated [1-4]. The main idea of the method is that in case of reflection an electromagnetic wave inciding on the composite resonator the phase of the reflection coefficient contains easily obtainable information about losses in structure. Under some conditions to obtain total absorption in structure it is only necessary to measure a distance between a resonant peculiarities on phase vs frequency characteristics. The fact that HBAR displays resonance frequencies in a very broad frequency band opens the unique possibility of measurements of frequency dependencies of acoustic losses which contain useful information about the absorption mechanisms. The purpose of this work is to apply this method to at first glance more simple objects containing only one layer - quartz resonators and to connect directly measured by this method losses with propagation characteristics of the material. The use of the HBAR method let us to investigate the resonators on very high overtones when real resonances and antiresonances already do not exist [5].

Applicability of HBAR method to quartz resonators

Acoustic resonators are multyfrequency quasiharmonic oscilating electro-mechanical systems. Their properties may conveniently be described using equivalent electrical circuit with series branch (L_n,C_n,R_n) shunted by parallel capacitance C_p [4]. Here n is an odd - number of resonance .

The information about the absorption in HBAR acoustic spectroscopy is taken only from frequency measurements of the positions of peculiarities (minimums and maximums) on reflection coefficient phase vs frequency dependencies in the vicinity of the n-th resonance as it is shown at fig.1.

The difference between frequencies corresponding to the extrema $\Delta f_n = f_n^{max} - f_n^{min}$ may be

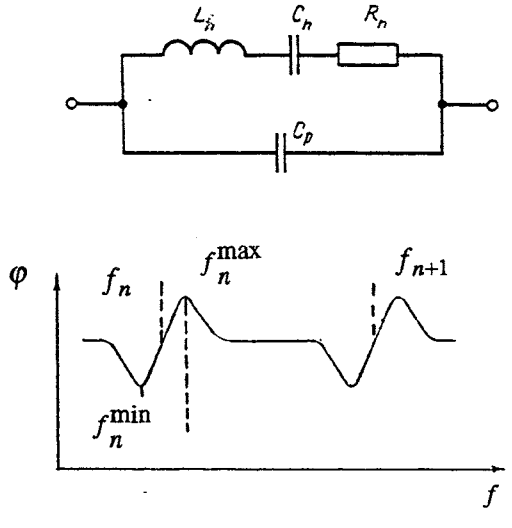


Fig.1. Equivalent circuit and an example of the reflection coefficient phase vs frequency characteristics.

connected with the real bandwidth of the series branch Δf_n^c if to take into account the characteristic impedance of the network analyzer or reflectometer used in experiments R_0 and the equivalent resistivity of series

branch R_n. In some cases it is also important to take into consideration the ratio of C_p/C_n. Then

$$\Delta f_n^c = \Delta f_n * \left[1 - \left(\frac{R_0}{R_n}\right)^2 + 16\left(\frac{C_p}{C_n}\right)^2\left(\frac{\Delta f_n}{f_n}\right)^2\right]^{-1/2} \quad (1)$$

Here f_n is the frequency of th n-th harmonic.

The apriori unknown ratio of capacitances may be expressed via electromechanical constant of the piezoelectric material using well known expression [4] :

$$C_p/C_n = n^2\pi^2/8K^2 \quad (2)$$

For quartz the third term in brackets is negligibly small even for very high add harmonic numbers. But for materials with high values of K^2 and high losses this term may become noticible.

As it can be shown using [1-3] for the case of usual quartz resonators with only one layer of material the expression for total energy losses per second α_n may be written in a very simple form :

$$\alpha_n = \pi\Delta f_n^c \quad (3)$$

It is easily seen from (1) the applicability of the method is restricted by the condition

$$R_O < R_n \quad (4)$$

So if this condition is not fulfilled the method is failed. But with the increase of the number of harmonics R_n rapidly increases and at higher frequencies (4) may be satisfied. Moreover with the increase of n real resonances and antiresonances may disappear. In this case the method is valid again and

$$\Delta f_n^c \equiv \Delta f_n \quad (5)$$

Notice that quality factor Q of resonator is connected with α_n by well known relation [6] :

$$Q = \omega_n/2\alpha_n \quad (6)$$

As a result, using the HBAR method, one can estimate the bandwidth and the quality factor of the equivalent circuit series branch of the resonator on high harmonics. An advantage of the method for material losses characterization is the absence of necessity to make the measurements of plate or resonator thickness.

Experimental results

Measurements of propagation characteristics of precision high quality resonators required the use of the network analizer or reflectometer with very high frequency stability and resolution. We used in our experiments two chanel network analizer HP 4396A. One chanel was used for measurements of the frequency dependencies of the phase of electromagnetic wave reflection coefficient and observation of Smith's charts, other - for direct observation of the phase and the amplitude of the impedance of the resonators.

We have studied various high precision and standard AT, BT and SC-cut resonators. At low harmonics (usualy up to 3-5) real resonances and antiresonances were easily seen and it possible to make direct measurements of the bandwidth using $\pm 45^\circ$ points of phase of impedance vs frequency $Z(\omega)$ characteristics (we denote this bandwidth as $\Delta f_n(\pm 45°)$. The series resistance R_n at resonance frequency is measured when on Smith's chart $\mathrm{Im}\, Z(\omega) = 0$.

At high numbers of harmonics when $\mathrm{Im}\, Z(\omega) < 0$ and real resonances disappear the bandwidth may be found only from the frequency dependence of the phase of the reflection coefficient in accordance with (5).

Comparision of bandwidth data obtained for some samples by two methods is given below :

AT-cut, planar faces

n=3,

f_n= 5MHz, $\Delta f_n(\pm 45°)$=2.2 $\pm$ .1Hz,

R_n =69.7 Ohm , Δf_n^c=2.2 $\pm$ .2Hz,

$$Q = 2.2.10^6$$

n=5,

f_n=10MHz, $\Delta f_n(\pm 45°)$=8.3 $\pm$ .1Hz,

R_n =89.6 Ohm, Δf_n^c=8.4 $\pm$ .2Hz,

$$Q = 1.2.10^6$$

SC-cut, planar faces

n=3,

f_n=10MHz, $\Delta f_n(\pm 45°)$=8.2 $\pm$ .1Hz,

R_n =63.1 Ohm, Δf_n^c=8.2 $\pm$ .2Hz,

$$Q = 1.2.10^6$$

SC-cut, BVA,

n=3,

f_n=10MHz, $\Delta f_n(\pm 45°)$=7.22± .05Hz,

R_n=85.6 Ohm, Δf^c_n=7.3± .1Hz,

$$Q = 1.39.10^6$$

It is seen that the data obtained by both methods practicaly coincide.

With the increase of the number n of harmonics the only one method of measurements of bandwidth remains.

The scope of data of the results of measurements of acoustic losses in AT BT and SC - cuts are shown on figs. 2 and 3 respectively.

As it is seen from fig.2. all points are found at somewhat higher level that it is predicted in [10], but obey ω^2 law that is typical for Akhieser losses and quantitively coincide with with experimental echo-pulse data [9].Points shown by dots at frequencies higher than 1000 MHz differ from these denoted as x by substracting losses connected with roughness of the surface.The roughness was measured directly and additional losses were calculated using expressions from [7,8].

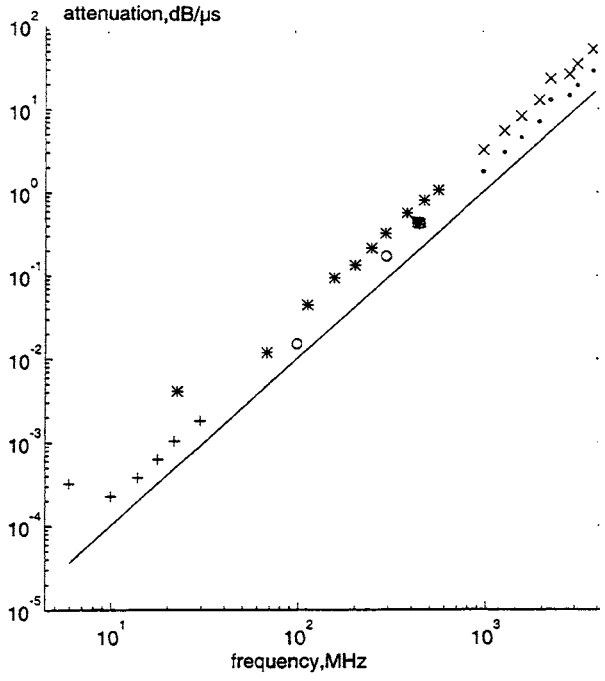


Fig.2. Acoustic losses in AT-cut quartz : +- various samples on 1-9 harmonics ;* - 1-21 (exept 19) harmonics ; o 9 and 27 harmonics ; solid line - data from [10] ; ■ - on 450MHz - result of precision measurements [9] ;x and . - our [3] data from microwave composite resonator measurements.

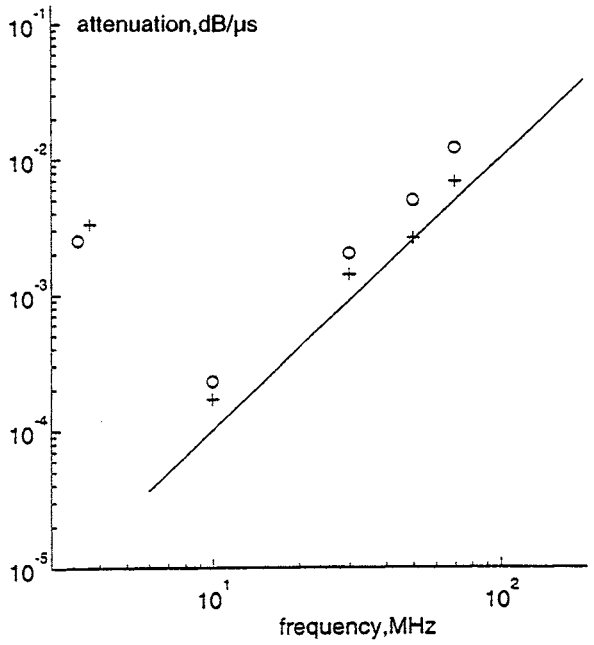


Fig.3. Acoustic losses in SC -cut quartz.o- resonator with planar faces, + - BVA "planar- convex" structure.

In the case of SC - cut total losses are somewhat smaller then in previous case.It is also seen that in the case of BVA at high harmonics losses are smaller. Experiments with BT- cut samples showed that they have the smallest losses in comparision with those in AT and BT -cut samples.

Conclusion

We presented and tested simple method for acoustic losses measurements in quartz resonators. Advanatages of the method are the simplicity, that comes from employing only frequency measurements and the opportunity to measure the same sample in wide frequency range. Unlike methods known before the presented method does not require complex calculations.

The acoustic losses on high frequencies in commonly used cuts of quartz as it is found are very close to the material Akhieser's losses.

Acknowledgements

Authors are grateful to collegues from IRE RAS I.M. Kotelyansky for manufacturing of transducers for microwave composite resonators, to academician Yu.V. Gulyaev and prof.S.N.Ivanov for support of the work.

Authors are also grateful to prof.V.B.Gruzinenko, dr. V.V. Bezdelkin (Institute "Fonon") and dr. V.A. Novoselov (Institute "Elpa") for their materials and samples used in some experiments.

Considerable part of the work was done in Laboratoire de Chronométrie Electronique et

Piézoélectricité - ENSMM, Besançon - FRANCE. Authors are greatful to C. Tellier and D. Benmessaouda for help, and Prof. R. Besson for support of this work and some samples.

References

[1] G.D.Mansfeld, A.D.Freik, B.N.Krutov. A high overtone BAW resonator as a tool for acoustic characterization at thin films and layers, in Proceeding of EFTF, 1993, pp 261-266.

[2]. G.D.Mansfeld, A.D.Freik, B.N.Krutov, High overtone HBAR resonator as a tool for acoustic characterization of thin films, Proc. of the Conference "Acousto-electronics-93 " Varna, Bulgaria, 1993, p. 20-30.

[3] G.D.Mansfeld, BAW Composite resonator spectroscopy, Proc. of IEEE UFFC Simposium, 1994, Cannes, France.

[4] G.D.Mansfeld, Measurements of acoustic wave attenuation in $La_3Ga_5SiO_{14}$ using HBAR technique, Proc.of IEEE Frequency Control Simposium, 1994, Boston, USA.

[5] J. Zelenka, Piezoelectric resonators using bulk and surface acouctic waves: Materials, technology, applications. Moscow, "Mir",1990

[6] B.A.Auld, Acoustic fields and waves in solids, Robert E. Krieger Publishing company, Melabar, Florida 1990.

[7] V.V. Kosachev, Yu. N. Lohov, V.N.Chukov, BAW seattering on statistically rough free boundaries of solids ; Sol. State Phys. (Sov.) v 31 p 105-113, 1989.

[8] S.N Ivanov, E.N. Khazanov, Scattering of BAW by rough surface. Radiotechnics and Electronics (Sov), v2 p 402-408, 1981.

[9] J. Lamb, J. Richter, Anisotropic attenuation with new measurements for quartz at room temperatures, Proc. Royal., Soc., v. 293, p 479-492, 1976.

[10] J.F. Rosenbaum, Bulk acoustic wave theory and devices, Artech House - Boston, London, 1988.

STUDY OF SC-CUT LATERAL FIELD EXCITED RESONATORS FEATURES

A.Masiukiewicz, B.Gniewińska, W.Szulc, K.Weiss

Tele and Radio Research Institute, Ratuszowa 11, 03-450, Warsaw, Poland

ABSTRACT

Several lateral field excited SC-cut resonators of different constructions were made at Tele and Radio Research Institute. Basic parameters and other features as dynamic temperature coefficient, nonlinear characteristic and 1/f noise were studied. Three oscillators employing LFE resonators were built. The high stability oven and specially developed oscillation circuits were apllied [11,12,13]. Parameters of oscillators were compared with resonators, oven and oscillation circuits parameters. Present limitations and further expectations of high stability oscillators using LFE resonators were discussed.

1. INTRODUCTION

The lateral field excited resonators (LFE) have a few well known advantages [1,2,3,4,5,6]. The main are lack of electrodes in the active part of the crystal plate (this enables achievment of Q factor close to the material limits) and very efficient suppresion of B mode of SC-cut resonators (due to different than for typical SC cut resonator field direction which don't excite B mode). Application of LFE resonators can improve both long and short term frequency instability, parameters critical for high stability oscillators.

At present the best oscillators on the market are offered by Oscilloquartz and BVA Industries The most important parameters of these oscillators are within the following ranges:

- frequency instability over operating temperature range 1E-10 to 5E-10,
- long term frequency instability 2E-11 to 5E-11/day,
- short term frequency instability 5E-13 to 1E-12/s.

To obtain such parameters the BVA resonators produced only by a few firms of difficult and expensive technology are applied. But to meet high performance requirements beside the resonator parameters several factors should be taken into account. The oven of high quality and good short term temperature stability has to be used (for example double oven) as well as low noise electronics. The subtle resonator parameters such as nonlinear behaviour and dynamic temperature coefficient have to be considered.

The lateral excited resonators can achieve very high quality factor close to the material limits with simple technology (beside the calibration procedure) but on the other hand it is followed by very small motional capacitance and very large motional resistance.

Taking into account parameters of the especially developed oscillation circuit , oven and LFE resonators parameters the short term frequency instability of 5E-13/s can be achieved.

2.BASIC PARAMETERS OF LFE SC-CUT RESONATORS

Measuremets of basic parameters which were carried out at the first stage of the work using π type four port circuit were not accurate. That method was based on motional resistance and capacitance measurements and because of very small C1 value it was difficult to achieve good repeatibility. To improve accuracy two methods were applied. The first based on HP network analyzer type 8753D measurements. Resonators were directly connected to the analyzer terminals. The quality factor and motional resistance were calculated basing on S21 measurements according to following formulas:

$$R_1 = 2Z_0(10^{\frac{-S_{21}}{Z_0}} - 1) \qquad 2.1$$

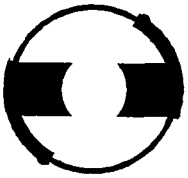
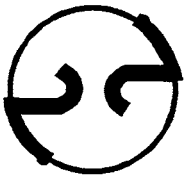
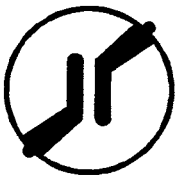
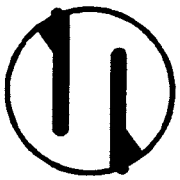
where $Z_0 = 50\Omega$

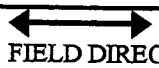
$$Q = \frac{tg\varphi_{zQ}}{2\frac{\Delta f}{f}} \qquad 2.2$$

where $tg\varphi_z = \frac{tg\varphi_{S21}}{1 - \frac{|S_{21}|}{\cos\varphi_s}}$

The second method used for results verification was the reflection method [15]. Good correlation between these two methods was achieved. The results are presented in Table 1.

TABLE 1

	ELECTRODES CONFIGURATION	BEFORE COLD-WELDED R_1 [Ω]	AFTER COLD-WELDED R_1 [Ω]	Q [E6]	NOTES
1		4000 ÷ 8000			not used because of high R1
2		2500 ÷ 3500	560÷820	1.6÷2.6	
3		5000 ÷ 9500			not used becouse of high R1
4		4500 ÷ 6000	1050÷1140	1.3÷1.4	
5		3000 ÷ 3800	770÷785	1.7÷2	
6		3100 ÷ 3200	1150	1.1	
7		3000 ÷ 3600	800÷1150	1.5÷1.9	



FIELD DIRECTION

The highest quality factor was obtained for 2 and 5 electrodes configuration. Large dispersion of quality factor and motional resistance was observed even for the same electrode types. Probable reasons were connected with technological details as for example diferences in blanks curvature.

3.NONLINEAR CHARACTERISTICS

The characteristics of frequency versus drive power were measured using HP network Analyzer type 8753D.
The results are presented in figure 1.

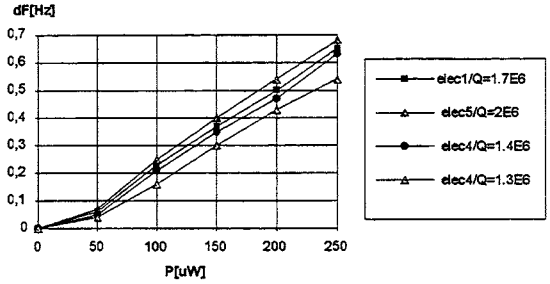


Fig.1 Frequency versus drive power characteristics of LFE resonators.

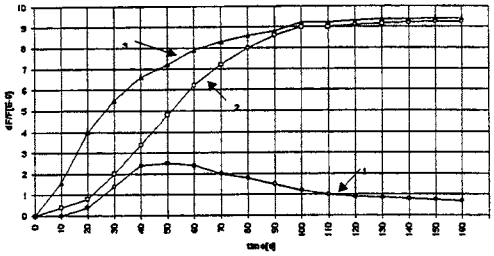
Results are close to that obtained for typical SC-cut resonators.To obtain short term instability better than 1E-12 the drive level shouldn't exceed level of 5÷10 uW.

4.DYNAMIC TEMPERATURE COEFFICIENT

Dynamic temperature coefficient was calculated basing on comparison frequency versus temperature characteristics which were measured for slow temperature drift and step temperature change.The resonator was excited by the oscillation circuit and was placed inside the oven typically used in OCXO oscillators. The temperature shock was realized as the rapid change of the oven temperature. The temperature versus time characteristic of quartz plate was measured using B mode of the typical SC-cut resonator as the temperature sensor and basing on the results the maximum velocity of temperature change was calculated. The results were as follows: Vmax=±0.24k/s for t=45s. The dynamic coefficient was measured for oven temperature lower and equal to turn over temperature of the resonator. The results were presented in figure 2 and Table 2.

TABLE 2

MEASUREMENT CONDITIONS	$\tilde{a}$ [E-7]
$T = T_e; \Delta T -$	-1.6
$T < T_e; \Delta T +$	-1.25



1 $\tilde{a}v, T = T_e, -\Delta T, V_{\max} = 0.24K/s$
2 $\alpha\Delta T + \tilde{a}v$
3 $\alpha\Delta T, T < T_e, +\Delta T, V_{\max} = 0.28K/s$

Fig.2 Dynamic temperature coefficient of LFE SC-cut resonators

5. 1/f NOISE

The 1/f noise measurement were carried out for the resonators of 4 types of electrode configuration which were not optimal from the Q factor point of view but only possible because of calibration problem. The difference of nominal frequency below 1 Hz for three resonators were achieved (shifted from 5MHz of about 500Hz). The measurements were made in bridge circuit of block diagram presented in fiure 3.

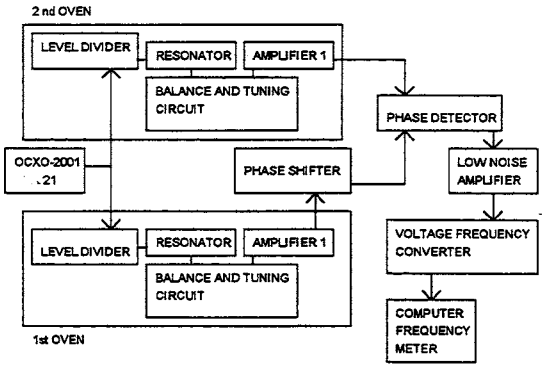


Fig.3 Block diagram of the 1/f noise measurement circuit

The 1/f noise for resonators Q≅1.3E6 was in the range σy[1s]=0.95÷1.1E-12. Assuming that 1/f noise dependence from unloaded factor quality of the resonator is according to $1/Q^4$ formula the 1/f noise for resonator of Q>2E6 could be lower than 5E-13. The present 1/f noise model of the resonator [14] doesn't discuss LFE resonators and this problem requires further studies.

6.OSCILLATORS MEASUREMENTS RESULTS

The oscillators were built on the base of three chamber oven with two symmetrically located heaters (BD139 medium power transistors) and two sensors which controlled temperature (two

10kΩ thermistors). The oscillator contains following circuit:

- Butler oscillation circuit with the resonator in the base,
- two transistors resonance output amplifier,
- voltage regulator,
- temperature control circuit (with separate voltage regulator)[11,12].

The resonator, oscillation circuit, output amplifier and voltage regulator were placed inside the oven in three chambers while the temperature control circuit was placed onto the surface of one of the oven walls. Thermal insulation was built from the poliurethane foam of 15 mm minimum thickness. The outer enclosure according to C007 IEC was used.

Three oscillators employing SC-cut LFE resonators were built and tested. Temperature characteristics of the the oscillators are shown in figure 4.

Frequency instability of 4÷8E-10 in the -10 to +60°C temperature range was achieved. In the same figure temperature characteristic of oscillator employing typical SC-cut resonator (with two chamber oven) is presented(OCXO-500P).

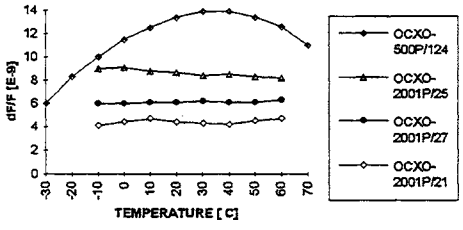


Fig.4 Temperature characteristics of OCXO-2001P oscillators

The results of long term frequency instability measurements are shown in figure 5.

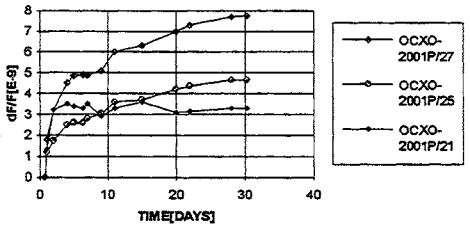


Fig.5 Long term instability of OCXO-2001P oscillators

The common feature of these characteristics is very short period of 3 to 5 days after which aging drops to the level of 1E-10/ day Short term frequency instability measurement results are presented in figure 6.

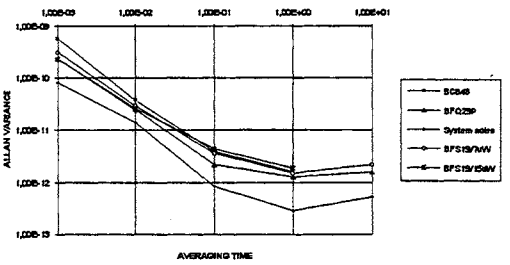


Fig.6 Short term frequency instability of OCXO-2001P oscillators

This instability was tested for two different voltage regulators LT2110 and uA723. A few types of transistors were tested in oscillation circuit. Measurements were carried out for two power loss levels in the resonators.

Taking into account the phase noise measurements results of oscillation circuit (with element of low Q=5 instead of the resonator)[11] with the assumption that the resolution was not worse than 3dB the following influence can be evaluated:

- oscillation circuit 3÷6E-13,
- resonator 0.95÷1.1E-12,
- temperature 2E-13[12].

The calculated σy[1s] was within the range 1.05÷1.2E-12 and the measurements results confirm these calculations.

7.CONCLUSION

The temperature behaviour of the oscillators is quite good and the temperature range can be extended. Long term frequency instability measurements were carried out only for one month so probably the lower aging could be achieved but it requires longer measurements .

Short term frequency instability is worse than expected but practically the level of internal resonator noise was achieved. Using resonators of higher Q this parameter can be improved.

At present the most difficult problem in the oscillator technology concerns the calibration procedure of the resonator. Resonators used in oscillators were calibrated within the range of 1Hz between themselves but they were shifted from 5MHz of about 500Hz.

The chemical treatment method used for calibration hasn't still assured proper efficiency and repeatibility and the the dispersion of calibration is with in the range 5÷500Hz.

8.REFERENCE

1.A.Ballato,E.R.Hatch: Lateral Field Excitation Of Quartz Plates. 37AFCS, 1983

2.R.C.Smythe: An Approximate Expression For The motional Capacitance Of A Lateral Field Resonator. 41st AFCS, 1987

3.A.W.Warner: Lateral Field Resonators. 39th AFCS, 1985
4.A.W.Warner: Measurement Of Plano Convex SC Quartz Blanks Using Lateral Field Excitation. 42nd AFCS, 1988
5.M.M.Driscoll: Fabrication And Methods For Evaluation And Circuit Utilization Of Prototype Lateral Field Resonator 41st AFCS, 1987
6.B.Capelle: Synchrotron Radiation X Ray Topography Study Of Lateral Field Resonators. 6th EFTF, 1992
7.B.Gniewińska: Effets Thermiques Dans Les Oscillateurs a Quartz. Seminaire "Etalons de Frequency". Besancon, 1981
8.D.B.Leeson: A SImple Model Of Feedback Oscillator Noise Spectrum. Proceedings of IEEE, 1966
9.J.P.Valentine: Analyse de Quelqes Aspects du Compartement Thermique das Resonateures a Quartz a Ondes de Volume. These a L'Universite de Franch Comte , 1983
10.Kurt O. Wessendorf: The Lever Oscillator for Use In High Resistance Resonator Applications. IEEE FCS, 1993
11.B.Gniewińska, A.Masiukiewicz: Optimization Of LFE SC-cut Resonators Oscillation Circuits. 8th EFTF, 1994
12.A.Masiukiewicz, B.Gniewińska, W.Szulc: Application Of LFE Resonators At High Stability Oscillators. IEEE FCS, 1994
13.A.Masiukiewicz, B.Gniewińska, W.Szulc: PreciseQuartz crystal Oscillators With Lateral Field Excited Resonators. PIEZO'94, Zakopane, Poland, 1994
14.F.L.Walls, P.H.Handel, R.Besson, J.J.Gagnepain: A New Model Of 1/f Noise In BAW Quartz resonators. IEEE FCS, 1992
15.M.Wójcicki, A.Lisowiec: A Direct Reflection Method And Six-port Measuring System For Accurate Determination Of Piezoelectric Resonators Parameters. 8th EFTF, 1994

SESSION 2

SPACE APPLICATIONS

Chairman : M. BRUNET

Automatic control systems for ground- and space- based hydrogen frequency standards

Yuri G. Gouzhva, Arvid G. Gevorkyan, Victor Z. Gankin, and Alexandre L. Myasnikov

Russian Institute of Radionavigation and Time

Abstract

A hydrogen frequency standard comprises a number of automatic control systems, which quantity and degree of complexity are determined by the list of device parameters to be realized under specific operational environment.

The data are given concerning a phase-noise spectral power density for two modifications of H-masers, quartz oscillator, frequency multiplier, and receiver. These data are used for building an optimal phase-loop in order to provide the minimum frequency instability.

Described is a building of autonomous resonator auto-tuning system using an amplitude technique for resonator tuning, and given is an analysis of its parameters. Estimated are limit parameters of hydrogen frequency standard depending on time constant of resonator auto-tuning system.

Introduction

While designing the present models of H-masers for space vehicles and for ground applications as well, a goal was to create devices that could operate autonomously and have a maximum unification of circuitry decisions.

A phase-lock loop (PLL) for H-maser comprises: high-precision quartz oscillator (QO) with loops for coarse and fine control; buffer amplifier; superheterodyne receiver with triple frequency conversion; frequency multiplier generating heterodyne frequencies; readjusted frequency synthesizer and frequency-phase discriminator.

PLL function is to transfer of signal frequency of atomic oscillator to a standard frequency with a minimum stability loss. For PLL optimization, spectral parameters of each functional unit were investigated.

The similar tasks were considered and solved by a number of authors, for example [1].

Resonator auto-tuning system (RATS) performs a resonator tuning of atomic oscillator to a frequency of emission line, taking into account a frequency shift compensation due to the spin-exchange atom interaction within a storage bulb, and maintains this state under an action of disturbing forces providing desired long-term frequency stability and reproducibility from turn-on to turn-on.

The goal of this work is PLL/RATS analysis for H-masers of various applications and contribution of these systems into frequency instability.

I. PLL optimization

Block-diagram of PLL is well-known (Fig.1). Using a known mathematical model, let us present a phase-noise power spectral density (PNPSD) with power functions [2]:

$$S(f) = h_{(i+2)}\, f^i, \qquad i=0,-1,-2,-3,-4. \tag{1}$$

In general, the following kinds of noise are typical for various units: for H-maser - white frequency and white phase noise - $S_H(f)=h_{0H}f^{-2}+h_{2H}$; for receiver - white phase noise - $S_r(f)=h_{2r}$; for QO - white frequency and white phase noise and frequency-phase flicker noise - $S_{QO}(f)=h_{-1QO}f^{-3}+h_{0QO}f^{-2}+h_{1QO}f^{-1}+h_{2QO}$; for frequency multiplier - white phase and phase flicker noise - $S_{fm}(f)=h_{1fm}f^{-1}+h_{2fm}$; h^{fs}_{0H} - coefficient for full-sized H-maser PNPSD reduced to the QO frequency is determined from:

$$h^{fs}_{0H} = \frac{k\, T_r\, f^2_o}{P_o\, Q^2_l\, N^2}, \tag{2}$$

where k is Boltzmann constant, T_r - resonator temperature, f_0 - oscillator operational frequency, P_o - atomic emission power within a resonator, Q_l - line quality, $N=f_o/f_{QO}$, f_{QO} - QO frequency. For $P_o=5\cdot10^{-14}$ W, $Q_l=2\cdot10^9$, $T_r=325$ K, $h^{fs}_{0H}=5.6\cdot10^{-13}$ rad^{2}Hz, h^c_{0H} and h^c_{2H} - coefficients for PNPSD of compact H-maser with resonator Q-regeneration reduced to the QO frequency can be determined through relations [3]:

$$h^c_{0H} = \frac{k\, T_{rn}\, f^2_o}{P_o\, Q^2_l\, N^2}, \tag{3}$$

$$h^c_{2H} = \frac{F\, k\, T\, (1+\beta_1+\beta_2)+\beta_1\, k\, T_{rn}}{P\,\beta_1\, N^2}, \tag{4}$$

where T_{rn} - resonator noise temperature, F and T - regenerator noise coefficient and amplifier temperature, β_1 and β_2 - coefficients of input and output resonator coupling with a regenerator.

A noise temperature of regenerated resonator is determined through [4]:

$$T_{rn}=T_r\left[1+\sqrt{T(F-1)/T_r(1+\beta_1+\beta_2)/\beta_1}\;(1-1/J)\right]^2, \tag{5}$$

where $J=Q_r/Q_{ld}$ - regeneration coefficient, Q_r - regenerated resonator quality , Q_{ld} - loaded resonator quality.

For compact hydrogen frequency standard, when $P_o=7\cdot10^{-14}$W, $Q_l=8\cdot10^8$, $Q_{ld}=4500$, $Q_r=70\,000$, F=3 dB, $\beta_1=0.5$, $\beta_2=0.1$, T=317 K, T_r=320 K. For this device, T_{rn}=2275 K and $h^c_{0H}=1.75\cdot10^{-11}$rad^{2}Hz, $h^c_{2H}=1.05\cdot10^{-11}$rad^{2}Hz.

h_{or} - coefficient for the receiver PNPSD is determined as:

$$h_{or} = \frac{k\,T_r\,(F_r - 1)}{L\beta_1\,P_o\,N^2}\,, \qquad (6)$$

where T_r is receiver temperature, F_r - receiver noise coefficient, L - coefficient of power transfer to the receiver input (L=1 for full-sized H-maser which frequency is put directly to a receiver and L=0.5 for compact H-maser), K - amplifier transfer ratio in a regeneration circuit.

For full-sized H-maser $h^{fs}_{0r}=1.7\cdot10^{-11}$rad^{2}Hz, for compact H-maser - $h^c_{0r}=3.1\cdot10^{-12}$rad^{2}Hz.

PNPSD for frequency multiplier (FM) and quartz oscillator are experimentally determined with use of low-frequency spectrum analyzers. The data were got for FM three modifications with various elements and components distinguished in circuitry decisions. The measurements have shown that the phase flicker noise was observed for all the FM modifications over the investigated frequency range (up to 10 kHz at least). Using measurement results, the following values of h_{1fm} were realized: 1 - $6.3\cdot10^{-10}$rad^2, 2 - $1.6\cdot10^{-12}$rad^2, 3 - $2\cdot10^{-13}$rad^2.

The results of QO PNPSD measurements were extrapolated into the range of lower frequencies. As a result of averaging results, the coefficients being typical for QO are found: $h_{-1QO}=7.9\cdot10^{-10}$rad^{2}Hz, $h_{0QO}=0$, $h_{1QO}=1.25\cdot10^{-12}$rad^2, $h_{2QO}=2.1\cdot10^{-14}$rad^{2}Hz.

The data are plotted in Fig.2,3 using the double logarithmic scale.

The optimum value of PLL maintaining band, proceeding from the criterium of reaching minimum frequency root-mean-square variation, is determined by that Fourier frequency value, at which QO PNPSD and total PNPSD of other units forming a synchronization circuit coincides, that is

$$h_{-1QO}\cdot f^{-3} = h_{2r} + h_{1fm}\cdot f^{-1} + h_{0H}\cdot f^{-2}, \qquad (7)$$

it means that a desired Fourier frequency value can be found from a cubic equation solution:

$$h_{2r}\cdot f^3 + h_{1fm}\cdot f^2 + h_{0H}\cdot f - h_{-1QO} = 0\,, \qquad (8)$$

or from graphic representation.

II. Autonomous RATS for H-maser

RATS with modulation of resonator tuning frequency and amplitude detection was considered in works of a number of authors [5]. Now we shall consider its application.

An error signal in this system is generated due to a difference in amplitude of H-maser generation signal at the deflection of resonator tuning mean frequency from the emmission line frequency.

F_m, modulation signal frequency, should satisfy following conditions:

1. $\Delta f_l << F_m \le \Delta f_r$; (9)
2. $F_m > \Delta F_{mb}$;
3. $0.2 < (f_o/F_m - E\,[f_o/F_m]) < 0.8$,

where Δf_r, Δf_l - transmission band of resonator and emission line, respectively, ΔF_{mb} - PLL maintaining band in H-maser, f_o - emission line frequency, E - an integer part of a number.

Realization of condition 1 enables to modulate of resonator tuning frequency, without disturbing a line emission, condition 2 - to decrease significantly an influence of spurious phase modu lation on PLL and to maintain H-maser output signal spectral purity, and condition 3 - to eliminate an influence of n-th harmonic of modulation signal on emission line.

A square type of modulation signal can be considered as optimal, as far as resonator frequency takes two values that can be easily fixed. Further, while using a digital signal synthesis, the circuitry is simple to the greatest extent, and, finally, the compencation of spin-exchange frequency shift is fortunately realized through a change of pulse period-to-pulse duration ratio of pulse train.

The main parameter of automatic control system is its transfer ratio with open feedback loop (K) or time constant ($\tau_a=1/K$);
τ_a is restricted from upper side (due to necessity to detect an error signal in noise) and from low side as well (due to damping of environmental temperature variations), that is:

$$\frac{(1+r^2)^3\,k\,T_{rn}}{16\,\pi\,P_o\,Q^2_l\,r^2\,(\Delta f/f)^2_{all}} < \tau_a < \frac{(\Delta f/f)_{all}\,f_o}{(Q_r/Q_l)\,\alpha_r\,\Delta T_d/K_{ov}}\,, \qquad (10)$$

where $(\Delta f/f)$ is an allowable H-maser fractional frequency drift, α_r - resonator frequency temperature coefficient, K - resonator temperature stabilization coefficient caused by an ovenizing system, ΔT_d - maximum daily change of environmental temperature in a mode of limited frequency stability.

If $(\Delta f/f)_{all}=1\cdot10^{-14}$, $\alpha_r=20000$ kHz/K, $\Delta T_d=2$ K/day, $K_{OV}=3000$, with optimal r=0.7 and other parameters for H-maser with resonator Q-regeneration given above, we have 940 sec$<\tau_a<$1050 sec. Thus, an optimal RATS time constant for compact H-maser is within a narrow band of allowable values.

A generalized block diagram of autonomous RATS is given in Fig.4.

Realization of conditions limiting τ_a is provided through a choice of a digit number in a digital integrator and digital-to-analog converter (DAC). In such a case, it should be noted that the DAC volume should meet an additional requrement, i.e. quantization noise minimization, that is the discrete step of resonator retuning should not worsen frequency stability for time measurement intervals $\tau<\tau_a$.

Let us note two main destabilizing factors in RATS operation which elimination is a principal matter for realizing H-maser high accuracy parameters.

a). While modulating resonator tuning frequency, a spurious phase modulation occurs, along with useful amplitude modulation, with modulation index $m = (r \cdot f_o)/(F_m \cdot Q_l)$, which appears as a synchronous stray signal.

The phase-modulated signal, transmitting selective circuits of receiving channel, results in spurious amplitude modulation, which compensation needs an equivalent resonator detuning, thus causing a frequency shift of H-maser output signal:

$$(\Delta f/f)_{PM} = \frac{0.46 \cdot f_o}{\Delta F_r \cdot Q_l^2} . \qquad (11)$$

Using the last relation, it can be shown that for compact H-maser F_r (receiver channel bandpass up to the amplitude detector) should not be less than 1.3 kHz.

b). While processing an error signal, synchronous detector and generator can produce H-maser uncontrolled frequency shifts due to dependence of its parameters on environmental conditions.

A synchronous detector transforms an error signal into constant voltage being equal to zero for fine-tuned resonator. Electronic keys performing such an operation, as well as voltage comparator generating the logic levels dependent on detuning sign, produce a certain constant voltage at the comparator output due to the thermal currents of its elements, when there is no useful signal. This voltage is compensated through equivalent detuning when RATS loop is closed. Since a spurious voltage level depends on environmental temperature, its change results inevitably in H-maser output frequency instability through a resonator frequency variation.

To eliminate this non-desired effect without selecting elements is possible by means of interchanging with F_m/k-frequency (k - an integer number) channels of transmitting error signals to the comparator input and synchronous changing count direction within a digital integrator.

III. Experimental testing electronics parameters of hydrogen frequency standard (HFS)

A series of HFS stability measurements was performed. The measurement circuits used are given in Fig.5. A variety of circuits is caused by a desire to have the valid data concerning a contribution of HFS functional units into a frequency instability of the generated signal.

H-maser frequency stability measurement using 5-MHz output is given in Fig.5a.

Measurements performed according to the circuit in 5b enable to eliminate the H-maser physical package (HMPP) contribution into the instability of HFS output signal, i.e. permit to estimate HFS electronics quality.

When a frequency of one of H-maser changes by a value near to 1 Hz, it is possible, by using an operational mode with a large magnetic field, to perform measurements with an increased resolution at the FM output (Fig.5c).

Fig.5d shows a way to reach this effect through a checking of HFS electronics parameters, where the detuning of m Hz at a high frequency is provided by a shift of synthesized frequency within one of HFS by a value $(F_{FS1}-F_{FS2}) = m \cdot N/N_{fm}$. Besides, while using such measurements, a contribution of frequency multiplier into frequency instability of checked signal is decreased in $|N/(N_{FM}-N)|$ time, because it is compensated at this output by means of feedback loop over time measurement intervals $\tau > 1/2\pi\Delta f_{mb}$, thus enabling to estimate FM quality on the basis of comparison of measurement results according to Fig.5b and 5d.

The results of measurements performed for HFS electronics with FM-2 are presented in Table.

Measurement circuit	Fig.5a		Fig.5b	Fig.5c		Fig.5d
RATS status	turn-on	turn-off	turn-off	turn-on	turn-off	turn-off
Measurement time interval, sec:	Allan variance, $n \cdot 10^{-13}$					
1	12	13	11	9	9	2.4
10	2	1.7	15	1.5	1.4	.31
100	.21	.07	.07	.15	.17	.04
1000	.06	.06	.07	.03	.05	.005

While analyzing these data, one can conclude as follows:
1). PLL system provides a realization of high performance with respect to short-term and long-term frequency stability.
2). Frequency stability of electronic circuits is determined over time measurement intervals interesting for us, beginning with 1 sec, by frequency multiplier. This is in a full correspondence with measurements performed in frequency domain for frequency multiplier (Fig.2), thus confirming an allowance for its extrapolation into low Fourier frequencies.
3). Insertion of RATS does not worsen HFS frequency stability over small time measurement intervals.

Conclusion

Automatic control systems were investigated used in hydrogen frequency standards for ground and space applications. This enabled to realize a ground-based standard with a full-seized spherical resonator designed by VNIIFTRI (Russia) with frequency instability $\sigma(\tau)=9 \cdot 10^{-13} \tau^{-1}$ for $\tau<\tau_a$ and $3.2 \cdot 10^{-15}$ over 1 day and HFS for space application with frequency instability $\sigma(\tau)=2 \cdot 10^{-12} \tau^{-1/2}$ for τ from 1 to 10^5 sec (Fig.6). HFS overview is given in Fig.7,8.

References

1. L. Prost, J. Buska, F.I. Gardiol. Limitation of Hydrogen Maser Frequency Standards Due to Receiver Noice. IEEE, IM, vol.IM-33, No.2, June 1984.

2. J.A.Barnes, A.R. Chi, L.S. Cutler, D.J. Healey, D.B. Leeson, T.E. McGunigal, J.A. Mullen, W.L. Smith, R.L. Sindor, R.F.C. Vessot, G.M.R. Winkler. Characterization of Frequency Stability. IEEE, IM, vol.IM-20, No.2, May 1971.

3. P. Lessage, C. Audoin. Frequency Stability of an Oscillating Maser: Analysis of The Effect of an External Feedback Loop. IEEE, IM, Vol.IM-30, No.3, March 1984.

4. A.L. Myasnikov, G.M. Kryukov, O.N. Kornyshov, I.V. Sergachev, S.D. Teplova. Construction and parameters of compact H-maser. Radionavigation and Time. 1993, No.1,2 (Russian).

5. H.E. Peters. Design and Performance of New Hydrogen Masers Using Cavity Frequency Sweetching Servos. Proceedings, 38th Symposium on Frequency Control, 1984.

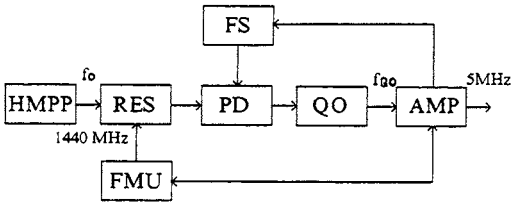


Fig.1. Block diagram of PLL.
HMPP - H-maser physical package; RES - receiver; FS - frequency synthesizer; FMU - frequency multiplier; PD - phase detector; QO - quartz oscillator; AMP - amplifier.

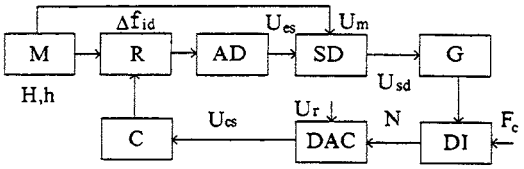


Fig.4. Block diagram of an autonomous RATS.
M - modulator; R - resonator; AD - amplifier-detector; SD - synchronous detector; G - generator; DAC - digital analogous convertor; DI - digital integrator; C - controller; U_{cs} - error signal; Δf_{id} - initial descrepansy of frequency; U_r - reference signal; U_{sd} - output of SD; U_m - modulation signal; F_c - clock frequency; N - control signal in digital code.

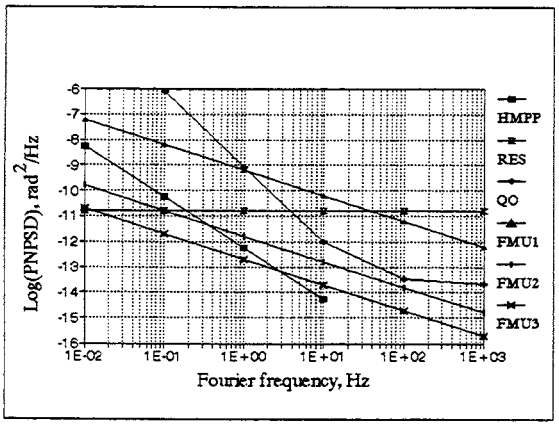


Fig. 2. PNPSD for full-size H-maser units and three versions of FMU.

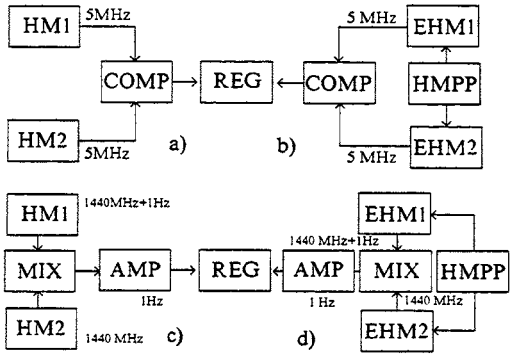


Fig.5. Measurement schemes:
a). 5 MHz output signals of two H-masers;
b). 5 MHz output signals of two electronic units;
c). 1440 MHz output signals of two H-masers;
d). 1440 MHz output signals of two electronic units.

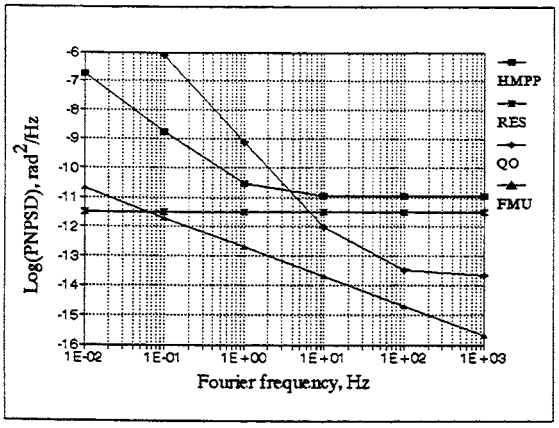


Fig. 3. PNPSD for compact H-maser units.

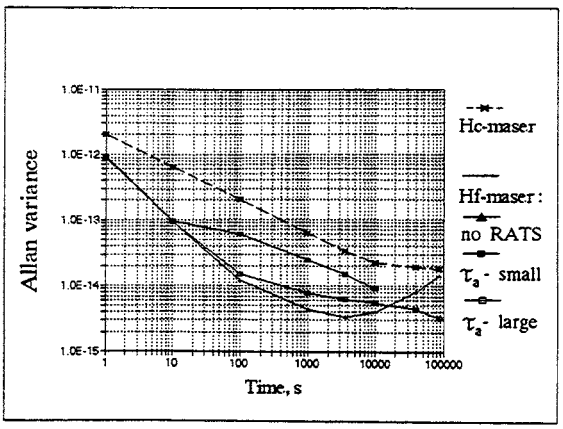


Fig. 6. H-maser of various modifications frequency stability.

Fig. 7. View of full-sized H-maser.

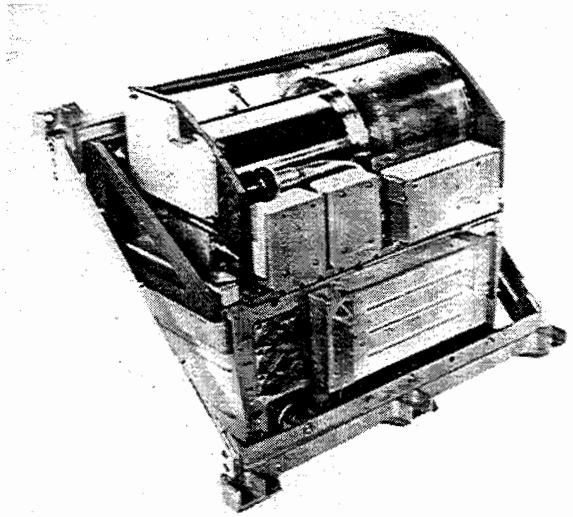


Fig.8. View of H-maser for space applications.

MONITORING OF THE PRARE CLOCK ON-BOARD METEOR-3: OPERATIONS AND PERFORMANCE ANALYSIS

S. Bedrich[1)]

DLR, Institute of Radio Frequency Technology, P.O.Box 1116, D-82230 Oberpfaffenhofen

A. Härting[2)]

GeoForschungsZentrum Potsdam (GFZ), Dept. Kinematics and Dynamics of the Earth, DLR/D-PAF, P.O.Box 1116, D-82230 Oberpfaffenhofen

Abstract

In the space segment of the satellite tracking system PRARE all frequencies are derived from a single ultra stable crystal oscillator. To calibrate the measurements, time marks generated by the oscillator are observed with a ground based reference clock and used to model frequency offset and drift continuously. We report first on the routine operations and their results in terms of long-term oscillator characteristics. Then, special analyses are presented using the oscillator for highly precise time transfer over single contact periods of a few minutes. The results show very good characteristics which are not degraded by the overall system design, indicating that an extremely stable oscillator, such as an H-maser, could extend PRARE to a next-generation system for global time transfer and clock synchronization.

1 PRARE System Overview

The Precise RAnge and Range-rate Equipment (PRARE) is a two-way microwave satellite tracking system [1,2]. The measurements are performed on-board via active ground station transponders. In addition to the round-trip X-band carrier a down-link is established in S-band to obtain direct data for the correction of ionospheric delay. A master station and commanding facilities receive the measurement data stored on-board and upload ephemeris and schedule data which are distributed to all ground stations by the PRARE space segment. It is thus possible to serve a global network of ground stations from a centralised control segment.

In the present PRARE system all frequencies are derived from a single source, the ultra stable on-board VCXO [3]. To correct the measurements for deviations from the nominal oscillator frequency and to transform the PRARE time tagging to the UTC scale it is necessary to

---

[1)] now at: GeoForschungsZentrum Potsdam (GFZ), Dept. Kinematics and Dynamics of the Earth, D-PAF, P.O.Box 1116, D-82230 Oberpfaffenhofen

[2)] now at: Fachbereich Seefahrt, Fachhochschule Oldenburg, P.O.Box 260, D-26926 Elsfleth

monitor the oscillator carefully. This is done at the master station Oberpfaffenhofen during satellite contact by receiving 1-pps signals from the satellite and comparing those to the local reference clock.

Presently, only the master station is capable of receiving time signals, but the modifications necessary for a standard ground station are minute. Therefore, PRARE can easily be extended to a high performance time transfer system. A study has shown that by replacing the on-board oscillator by an H-maser an unprecedented time transfer performance can be expected [4]. With the oscillator signals received at the master station analyses have been carried out that corroborate the assumptions made on the time transfer capabilities of the system.

Since February 1994, PRARE is being tested on the Russian METEOR-3 satellite and will soon be launched on the European remote sensing satellite ERS-2 as the prime navigation payload. Future applications with highest quality oscillators are considered as time transfer systems [5].

2 Time Keeping Operations

To monitor the PRARE oscillator a master clock is operated at the Oberpfaffenhofen site. This master clock consists of a GPS-disciplined Rb-oscillator in phase locked duplicate for redundancy purposes. During contact with the satellite, the PRARE station processor synchronizes itself to the X-Band downlink and produces a 1-pps signal that stops a time interval counter which has been started by the master clock. Data from this counter are written on disk storage by the same software that steers the entire measurement process.

For the geodetic purposes of orbit determination and point positioning, a timing accuracy of about 1 μs, during which the satellite moves by approximately 7 mm, is fully sufficient. To meet this requirement, the above described equipment is more than adequate.

Time tagging of measurements is achieved by a clock model which provides a transformation from PRARE-time T to UTC t. Each record of the clock model file consists of the quantities

$$t_{val},\ T_0,\ a_0,\ a_1,\ a_2.$$

The "valid-to"-time t_{val} is given in Julian days from January 1, 2000 and is only used to identify the applicable record in the file. Usually, one record is valid for three days. The epoch T_0 is the PRARE-time at the beginning of the validity period of that record, a_0, a_1, and a_2 are polynomial coefficients. This way, PRARE-time T, after selecting the correct record, is transformed into UTC t as

$$\Delta T = T - T_0$$

$$\delta t = a_0 + a_1 \Delta T + a_2 \Delta T^2$$

$$t = T_{reset} + T + \delta t$$

where T_{reset} is determined by the instant PRARE was switched on.

To generate the clock model the observed differences between the 1-pps signals from the reference clock and those received from the space segment are accumulated over three days. A least squares fit to these data by a second degree polynomial can only be done after correcting the PRARE signals for travel time. This would optimally be done using the distances measured at the same time by the PRARE system itself. This is, however, a tedious procedure, which has only been done for a few sample cases that will be reported on below. Operationally, we are compelled to calculate distances from the predicted orbit, which may be off by up to a few 100 m.

3 Discussion of Results

A typical result of a 3 days' fit (from August 1994) is shown in figure 1. Per day there are 6 to 7 contact periods longer than 10 minutes which appear as clusters of data. The curve represents the clock model with constant time offset, frequency offset (average slope) and frequency drift (curvature).

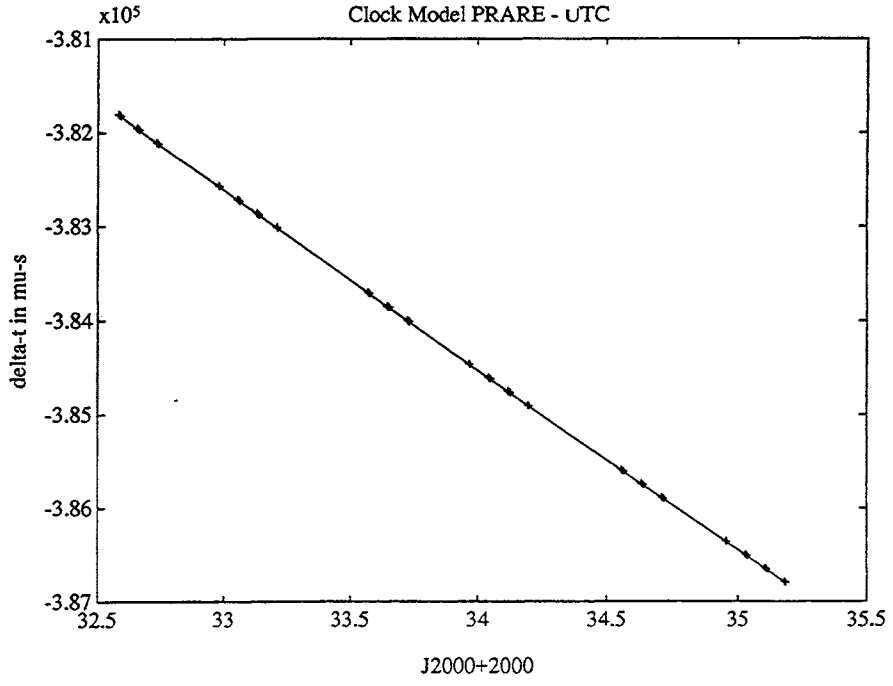


Fig. 1 Least squares fit of a second degree clock model polynomial to the PRARE-time - UTC difference measurements.

A detailed look at the residuals of the polynomial fit reveals that most of the uncertainty in the clock model is due to the limited accuracy of the predicted orbit. The residuals exhibit a pronounced structure in every satellite pass, leaving a total σ of typically 1 μs, whereas pure measurement noise is less than 5 ns. Some of the residuals may also be due to short-term behaviour of the oscillator which, however, cannot be distinguished in this routine data processing.

From the clock model we study the long-term characteristics of the on-board oscillator. For 1994, the frequency offset (3 days' average) with respect to the nominal 8.489 GHz carrier down-link is shown in figure 2. After switching on we observe a stabilisation of the drift. A period of uncertain data around day 125 has been skipped, around day 170 the oscillator was switched off for about 6 days. The slope between days 100 and 290 corresponds to a drift of 0.34 Hz/d or, in relative terms, $4.0 \cdot 10^{-11}$ per day. Around this value, the individual three-day averages are scattered with a standard deviation of $3.5 \cdot 10^{-11}$ per day.

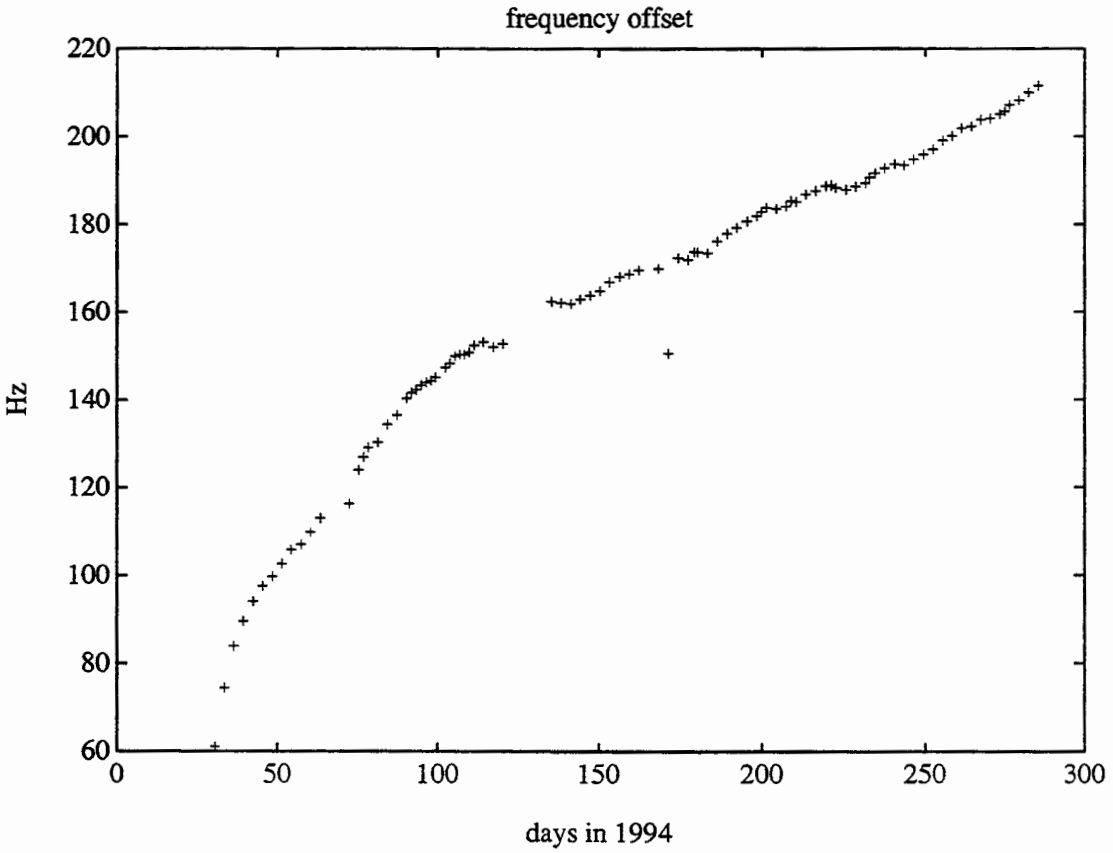


Fig. 2 Frequency offset of the PRARE oscillator on-board METEOR-3 with respect to the nominal carrier frequency of 8.489 GHz, averaged over 3 days.

4 Time Transfer Properties

To overcome the uncertainties in signal travel time correction, which are introduced by the predicted orbit, and to verify the time transfer and clock synchronization potential of the PRARE system, a set of high elevation passes over the master station has been analysed in detail. In the present work, the on-board two-way carrier Doppler measurements, which are recorded each second for range-rate determination, have been evaluated and compared to the 1-pps measurements on ground.

After correcting for the different down- and uplink carrier frequencies and for the change of the Total Electron Content during signal transmission which is, as mentioned, measured by

the ground stations, the one-way Doppler has been integrated continuously and transformed to an individual one-way signal travel time for each second. Not corrected is the travel time inside the hardware, which is also measured regularly by the system, but remains nearly unchanged during one pass and results only in a constant bias.

When differencing now the Doppler from the 1-pps measurements, the on-board oscillator frequency offset becomes visible. Its slope corresponds very precisely to the routine evaluation results shown in figure 1. After correction for the time drift with a linear fit and smoothing by a 60-seconds sliding average, the PRARE space segment clock and the master station Rb-clock are synchronized as presented in figure 3.

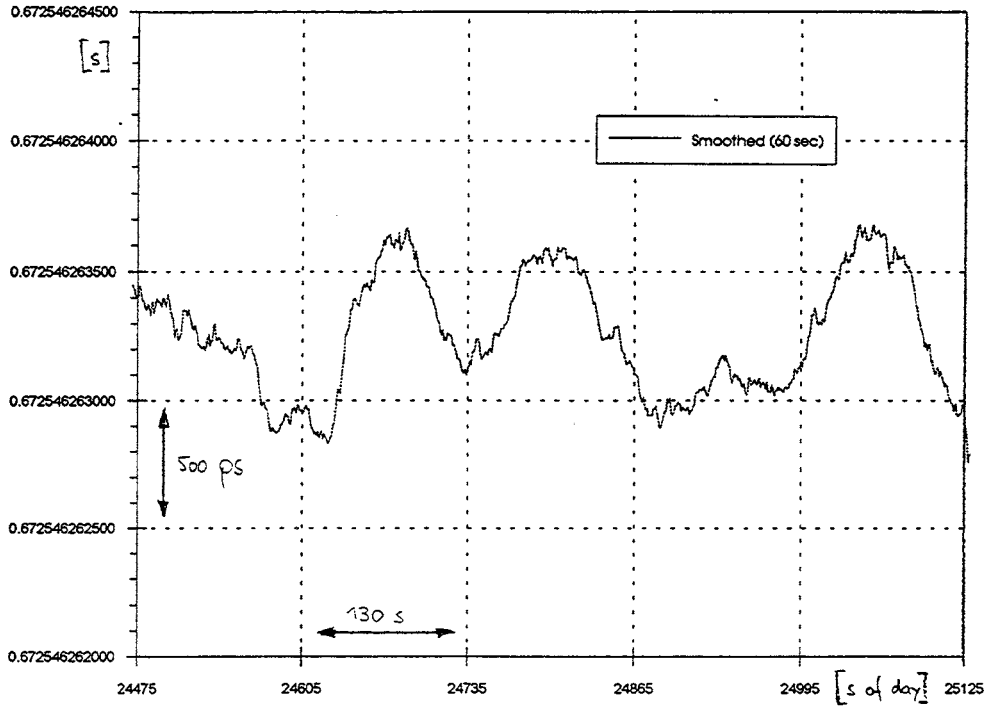


Fig. 3 Correlation of the PRARE on-board and the master station ground clock around culmination time of a master station pass after system measurement analysis.

The figure shows a 650-seconds detail around culmination point where tropospheric and multipath influence on signal propagation is a minimum. Residual noise of the Doppler measurement is nearly zero by nature and further reduced by the smoothing process. The resulting signature of the curve is assumed to be due mainly to the noisy 1-pps measurement on ground, to secondary oscillator effects, and to residual propagation effects.

The figure proves at the same time that the system is well applicable for high performance time transfer. It can be read out that the space oscillator lags around 0.6725 s behind the ground clock, having an accuracy well below 1 ns. These values can further be improved by a longer smoothing procedure and by including all time biases which are present, but neglected in this preliminary analysis.

5 Conclusion

Time comparison measurements between the VCXO on-board the PRARE space segment and the ground based master station clock are regularly taken. The data show that the oscillator performs as specified as far as the long-term characteristics are concerned. The routine operations, however, are uncapable of fully exploiting the extremely low measurement noise, which is indicative of very good time transfer properties of the system.

Detailed analyses of measurements confirm the excellent clock synchronization potential of the system showing a residual accuracy below 1 ns. The highly precise ranging and range-rate determination capabilities fully maintained, the PRARE system in combination with an atomic oscillator seems to have the potential to serve as a core element in an extended next-generation satellite time transfer and navigation system.

References

[1] Reigber, Ch. et al., "Das PRARE-System auf dem Satelliten METEOR 3/7", Die Geowissenschaften 12 (1994) 296-299.

[2] Hartl, Ph., "The Precise Range And Range-rate Equipment PRARE", Proc. Workshop on ERS-1 Radar Altimeter Data Products (ESA SP-221), Frascati, 1984.

[3] Schäfer, W., H. Wilmes, "The Precise Range And Range-rate Equipment PRARE - System Design and Status of Development", Proc. Workshop on Advances in Satellite Radio Tracking, Austin, 1986.

[4] 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.

[5] Reigber, Ch., W. Schäfer, E. W. Grafarend, H. Nau, PRARETIME - Precise Time Transfer and Tracking for Earth Parameter and Orbit Determination, Proposal to ESA, Potsdam/Stuttgart/Oberpfaffenhofen, 1994.

TIME ASPECTS OF THE EUROPEAN COMPLEMENT TO GPS: SUMMARY OF THE EXPERIMENTAL RESULTS.

P. Uhrich, B. Juompan, R. Tourde.
Laboratoire primaire du temps et des fréquences [LPTF],
Observatoire de Paris, Bureau national de métrologie, Paris, France.

M. Brunet, J-F. Dutrey.
Centre national d'études spatiales [CNES], Toulouse, France.

Abstract.

The CNES project of a European Complement to GPS [CE-GPS] is conceived to fullfill the needs of Civil Aviation for a non-precise approach phase with GPS as sole navigation means. This generates two missions: a monitoring mission - alarm of failure -, and a navigation mission - generating a GPS-like signal on board the geostationary satellites. The host satellites will be the Inmarsat III constellation. The CE-GPS missions lead to some time requirements, mainly the accuracy of the GPS time restitution and of the monitoring clock synchronization.

An experiment has been scheduled over the last three years, using the Inmarsat II F-2 payload and specially designed ground stations based on 10-channel GPS receivers. This paper presents the results obtained in the time domain during the continental phase (two stations) and the transatlantic phase (three stations) of the CE-GPS experiment. The time aspects are the accuracies of the GPS time restitution and of the monitoring clock synchronization either by Two-way time transfer through Inmarsat satellite, GPS Common-view, or using the restituted GPS time.

Introduction.

The CNES (French Space Agency) has conceived the project of a European Complement to GPS (CE-GPS in the following), dedicated to the needs of the Civil Aviation to achieve the requirements for a non-precision approach phase with GPS used as sole navigation means. This generates two missions: a monitoring mission - alarm of failure -, and a navigation mission - generating a GPS-like signal on board the geostationary satellites. Many functions have to be fullfilled by the system, for which the time requirements are reaching the state of the art of the techniques used by the time metrology. The CE-GPS project started more than four years ago, and the experimental part three years ago.

The concept of the CE-GPS and the experimental system were presented at EFTF 93 [1]. The space segment of the CE-GPS experiment, beside the GPS NAVSTAR constellation, is the Inmarsat II F2 AOR-E geostationary satellite. The ground segment is made up of three specially designed ground stations based on 10-channel navigation GPS receivers. The CE-GPS experiment was divided into three stages: a first stage in May-June 1993, the "calibration phase", with two stations connected to the same clock [1]; a second stage from September to November 1993, the "continental phase", with one station at the LPTF (Paris, France) and another one at the CNES space center of Toulouse (France) [2]; a third stage from May to June 1994, the "transatlantic phase", with three stations located at the CNES space centers of Toulouse (TLS, France), Kourou (KRU, French Guiana) and Hartebeeshoek (HBK, South Africa).

The results presented in this paper concern only the time aspects of the experiment during the continental and transatlantic phases: some calibrations of the ground stations; the restitution of the GPS time; the synchronization of the monitoring clocks. For the last purpose, GPS Common-view, Two-way time transfer through the geostationary satellite, and the restitution of the GPS time in each station have been used. The results obtained are compared to the requirements of the CE-GPS project.

1. Time aspects of the CE-GPS.

The description of the system architecture, and the discussion about time and frequency servo-control techniques, or orbitography aspects, have been made elsewhere [1,3]. Concerning the navigation mission, the time requirements are about the synchronization with respect to the time scale broadcasted by the GPS satellites. The standard deviation of the time difference between an event issued from the geostationary payload and an equivalent event issued from any GPS satellite should be less or equal to 120 ns with Selective Availability (S.A.) on [4]. With S.A. off, this requirement drops to 20 ns. Concerning the monitoring mission, the time requirements deal with the relative synchronization of the monitoring clocks, within 10 ns (1 σ) in accuracy if they are in view of the same geostationary satellite, or within 15 ns (1 σ) if not.

The method for restituting the GPS time is based on the statistical behaviour of the S.A. noise [2, 5, 6]. Some experimental results were obtained during the continental phase with a four-channel time dedicated GPS receiver: over an averaging period of 144 min, the GPS time could be restituted on the ground with an accuracy of 14 ns (1 σ), assuming that the uncertainty on the ionospheric delays is negligible. The results obtained during the transatlantic phase are described in this paper.

The synchronization of the monitoring station clocks is achieved either by GPS Common-view, by using the GPS time separately calculated in remote stations, or by Two-way satellite time transfer (TWSTT) through the geostationary payload. The well known method for the synchronization of remote atomic clocks is the GPS Common-View technique [7]. The BIPM schedule was implemented inside the GPS time dedicated receivers included in the CE-GPS ground stations. When using MITREX Modems and a Eutelsat satellite, an accuracy of 1.7 ns (1 σ) was estimated for TWSTT over a short distance ($\simeq$ 800 km), and directly compared to the results of the relevant GPS Common-views [8]. LPTF proposed to use the Inmarsat C- to L-Band transponder, and the spare C/A GPS Gold PRN-codes [1, 2, 4, 9]. Inmarsat II F2 AOR-E is located -15.5 ° E. The Sagnac effect for all links is easy to determine within 0.01 ns (1 σ). At the start of the experiment, no data were available concerning TWSTT performances when using the CE-GPS equipment. Four 15 min sessions were scheduled every day, with a basic sampling of 0.6 s.

2. Experimental set-up.

Inside each ground station (figure 1) are implemented two SERCEL NR106, which are ten-channel GPS navigation receivers. Thanks to the switcher 1, they are related either to a common GPS antenna (L1 carrier), or to the receiving antenna of the Inmarsat signal (L-Band: 1533.475 MHz). The GPS antenna is also connected to a SERCEL NRT2, which is a 4-channel time dedicated GPS receiver. The NRT2 is directly supplied with the 1pps (1 pulse per second) output of the Cesium clock, as in a classical set-up for GPS Common-view. Because the navigation receivers NR106 have no input for dating external events, like the 1 pps from the Cesium clock, it was necessary to build up a so-called 'GPS signals generator'. It generates a sequence of C/A code synchronized with the 1 pps output of the clock, which modulates a L1 carrier in order to be dated by the internal counter of both NR106. This signal is denoted

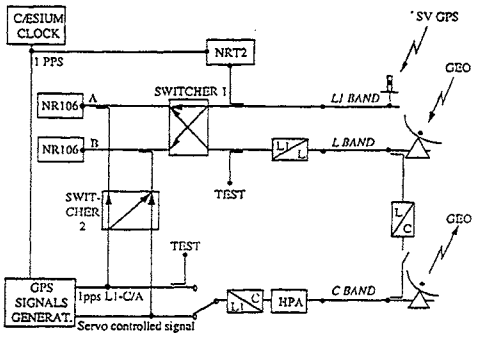


Figure 1. Set-up of a CE-GPS ground station.

'1 pps L1-C/A'. For all CE-GPS stations, the C/A code chosen for this internal link is the spare code 33 of the operational GPS. For refering any external signal to the Cesium clock of the station, that external signal and the 1 pps L1-C/A signal have to be dated simultaneously. An other function of the 'generator' is to output a servo-controlled signal towards the transmitting antenna (C-Band: 6428.475 MHz), similarly to the transmitting part of a Modem in conventional TWSTT, the equivalent to the receiving part of conventional TWSTT being the NR106 receivers.

The Cesium clocks of the CE-GPS stations were a HP 5071 A option 1 at the LPTF, a HP 5061 A option 4 at TLS and KRU, and an Oscilloquartz at HBK. All of them have proven to remain stable enough to enable a reliable evaluation of the CE-GPS equipments.

3. Atmospheric delays.

Both GPS and TWSTT need an estimation of the atmospheric delays. Models for tropospheric delays (only for GPS) are working reasonnably well at the nanosecond level. For ionospheric delays, calibrators are needed. Codeless receivers were used at all CE-GPS stations for measuring the ionospheric delays along the line of sight of the GPS satellites. It was anticipated that a polynomial mapping method should be used to determine the delays in the direction of the geostationary satellite. But many problems occured during the running of the ionospheric calibrators. Given the short distance between the stations involved in the continental phase ($\simeq$ 800 km), it was decided to use the STANAG [4] model for the GPS satellites, and a Bent model for the C- and L-Band ionospheric delays in the direction of the Inmarsat satellite.

During the transatlantic phase, the ionospheric calibrators have worked only part of the time. It was decided again to use the STANAG results as the ionospheric delays in the direction of the GPS satellites. Because a lot of data were missing for the TWSTT measurements, the LPTF proposed to compute the ionospheric delays as following: for each 15 s sampled measurements in the direction of the GPS satellites, a VTEC (vertical total electronic content) is calculated; a mean value of these VTECs is estimated, and projected in the direction of the geostationary satellite; for each TWSTT 15 min session where VTECs are available, a mean value of the ionospheric delays is computed; during the whole periods where TWSTT measurements have been made continuously, an average value of the ionospheric delays in the direction of the geostationary satellite is computed for each of the 4 daily sessions.

Three periods of 5 to 7 days duration of continuous TWSTT were identified during the transatlantic phase of the experiment. For the uncertainty of the ionospheric delays, it was proposed to adopt the highest standard deviation of the 15 min mean values, which was 5.1 ns (1 σ). An attempt to use IGS (International GPS Service for geodynamics) observations for post-processing is underway, but the results are not available yet.

4. Calibrations.

All the internal delays of the ground stations were either measured (cables) or estimated (electronic components). One of the critical issues is the calibration of the 'GPS signal generators', whose group delay appear in the synchronization equations. Frequent calibration sessions were scheduled in each station during the whole experiment. A calibration consists roughly in measuring the time delay between the arrivals into the NRT2 of two homologous 1 pps signals: a direct 1 pps signal from the Cesium clock and the 1 pps L1-C/A signal supplied through the generator output.

The results of the successive calibrations during the continental phase are presented in figure 2. The calibrations at TLS, which was switched off for a while on MJD 49266 (October 6), seem to be more stable than the LPTF station measurements, where the data are scattered within 25 ns. It has been shown by the manufacturer of the stations that, owing to the components used for this equipment, the signal generator could exhibit stepwise varying delays, each step being a multiple of 2.44 ns. By processing in deferred time the dating of the 1 pps L1-C/A performed continuously by both NR106 of each station, those variations could be used to improve the calibrations. But due to missing data, a lot of discrepencies actually remain: these discontinuities in the delays are disregarded. The average values of the calibration sessions are adopted along with the relevant standard deviations as uncertainties, giving for the signal generators delays during the continental phase:
- at TLS: mean value = 1007.5 ns; standard deviation = 4.8 ns.
- at LPTF: mean value = 998.5 ns; standard deviation = 7.1 ns.

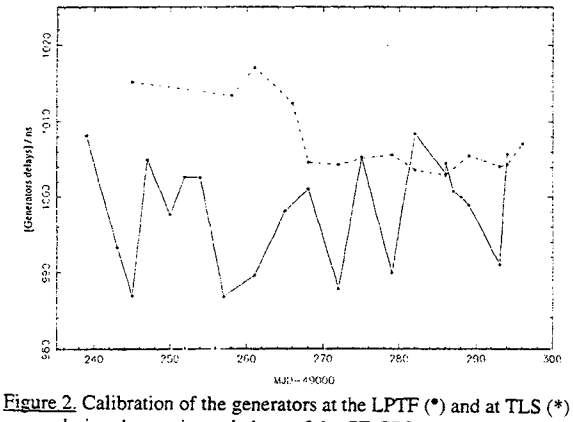


Figure 2. Calibration of the generators at the LPTF (•) and at TLS (*) during the continental phase of the CE-GPS experiment.

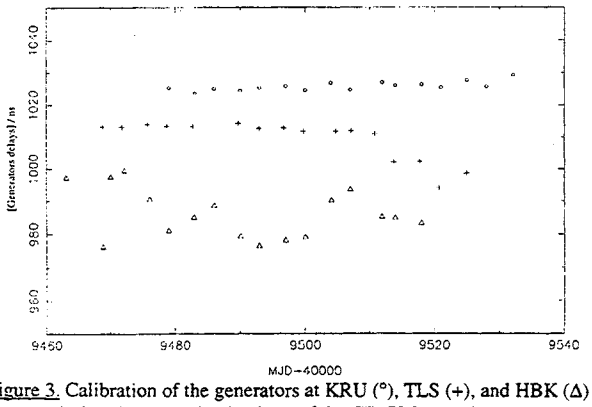


Figure 3. Calibration of the generators at KRU (°), TLS (+), and HBK (Δ) during the transatlantic phase of the CE-GPS experiment.

A similar behaviour of the signal generators was observed during the transatlantic phase of the experiment (figure 3). The manufacturer of the equipment believes that the generator of the KRU station has reached the best achievable stability for the components used. As for the continental phase, the average values of the calibration sessions were adopted along with the relevant standard deviations as uncertainties:
- at TLS: mean value = 1009.4 ns; standard deviation = 6.3 ns.
- at HBK: mean value = 986.1 ns; standard deviation = 7.6 ns.
- at KRU: mean value = 1025.8 ns; standard deviation = 1.4 ns.

A calibration by transportation of two GPS receivers was

organized only during the continental phase of the experiment, giving the usual uncertainty of 1.5 ns (1 σ) on the differences between the two time scales. No temperature effect was noticeable on all the sets of data during the continental phase of the CE-GPS experiment.

5. Continental experimental phase.

An operational time transfer link between LPTF and TLS by GPS Common-view is used for the computation of the 'Temps atomique français' (TAF). It was then interesting to set up the CE-GPS ground stations at both places for the purpose of testing the equipment in terms of time metrology.

The software able to build [REF - GPS] data from the NR106 pseudo-range measurements was not completed at the date of the continental phase of the CE-GPS experiment: only GPS data from the NRT2 receivers were used.

5.1. Two-way time transfer through Inmarsat.

The Modified Allan deviation for the 15 min sessions of TWSTT shows that white phase noise is preponderant on each 15 min session over a period long enough to allow the computation of an average value. The precision of one 0.6 s sampled CE-GPS observation is about 9 to 10 ns. It is ten times worse than MITREX precision [8], but it remains consistent with the magnitude expected when using C/A code along with standard discriminators on signals affected by a C/No of about 40 dB(Hz).

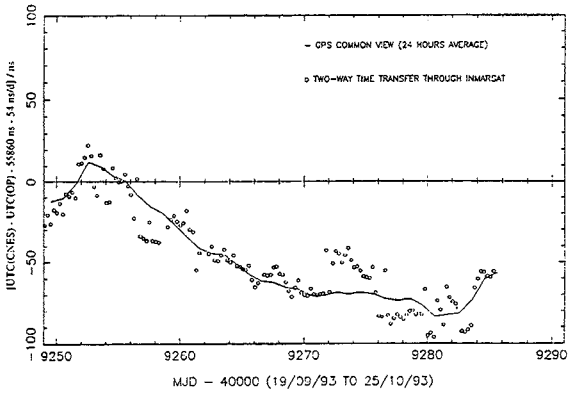


Figure 4. UTC(CNES) - UTC(OP) from GPS Common-views daily means (-) and TWSTT 15 min sessions (o) [continental phase of the CE-GPS experiment].

Figure 4 presents the average values of the TWSTT 15 min sessions, compared to the daily averages of GPS Common-views between the time dedicated receivers connected to the CE-GPS stations. The TWSTT points are well distributed around the GPS curve, even though scattered: many sessions have suffered from the above mentionned stepwise varying delays of the generators, whose uncertainties are obviously the most degrading part for the accuracy of TWSTT. We propose for this technique the following uncertainty budget in nanoseconds (1 σ):

Ionospheric delays [Bent model]	2.0
Equipment delays (LPTF - TLS)	6.2
UTC(LPTF)-1 pps REF [short line]	0.5
UTC(TLS)-1 pps REF [long line]	1.0
Two-way accuracy	6.6

where 'Equipment delays' accounts for the signal generators and various internal electronic delays.

This accuracy is given for any session average result, and remain well below the upper limit of the CE-GPS requirement (10 ns). It surely would drop to a more interesting value compared to the MITREX results [8] if the behaviour of the generators could be withdrawn from the measurements. The uncertainty of the comparison between GPS Common-view [7] and TWSTT can be estimated in nanoseconds (1 σ) by:

TWSTT (CE-GPS continental)	6.6
GPS Common-view [800 km baseline]	3.4
Uncertainty of the comparison	7.5

This uncertainty appears consistent with the data plotted in figure 4.

5.2. GPS time restitution.

The 4 channel NRT2 receivers were used for the restitution of the GPS time, by use of a schedule of 89 tracks of 13 min duration, per day and per channel. When averaging over the 4 channels, the accuracy of the GPS time restitution on the ground was estimated to 14 ns (1 σ), over a period of 144 min to limit the noise (S.A. + Observation noises) to 10 ns [2].

Figure 5 shows the results of time transfer either by GPS Common-view, or by using the GPS time restitution in both stations. The plots resulting from the computation of the GPS time are all inside the scatter of those from the regular common-views. Due to the short baseline between LPTF and TLS stations, the use of the GPS time restitution for the synchronization of the remote clocks has the same meaning as averaging Common-views. Moreover the statistical behaviour of these data has proven to be similar between the two techniques. The small shift that can be noticed between the two sets of data shows a remaining calibration error: the common-views were measured with the receivers used for the TAF link, while the GPS time restitutions were computed from measurements made by the receivers of the CE-GPS link. Over a baseline of less than 1000 km, the usual estimated accuracy for daily averaged GPS

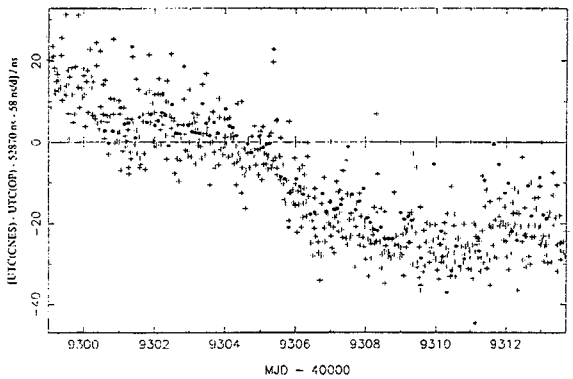


Figure 5. UTC(CNES) - UTC(OP) from GPS Common-views (+) or from the GPS time (•) as restituted at LPTF and TLS [continental phase of the CE-GPS experiment].

common-views is then the accuracy of a time transfer by using the GPS time restitution, that is 3.4 ns (1 σ).

6. Transatlantic experimental phase.

The transatlantic phase of the CE-GPS experiment is a good test for the ground segment of a future operational Global Navigation Satellite System. Due to local problems in the equipment, a lot of data were unfortunately lost. Only three short periods of continuous TWSTT measurements (5 to 7 days each) were identified [9]. Moreover, only one week of GPS data were available, for computing either Common-views or the GPS time. This prevents us to perform a comprehensive comparison of the three methods involved over long baselines.

6.1. Two-way time transfer through Inmarsat.

A statistical analysis of the TWSTT data from the transatlantic phase shows a predominant white phase noise, as for the continental phase. In a very similar way as for the continental phase, we propose the following uncertainty budgets in nanoseconds (1 σ):

Ionospheric delays		5.1	
Equipment delays	TLS - HBK	HBK - KRU	KRU - TLS
	7.0	5.5	4.6
UTC(k)-1 pps REF [two stations]		1.0	
TWSTT accuracies	8.7	7.6	6.9

These accuracies are near the limit of the CE-GPS requirement (10 ns). Assuming that the 1 pps L1-C/A generator at KRU exhibits the regular behaviour of an operational CE-GPS ground station, and that the ionospheric delays can be measured with a better uncertainty, one could adopt the following uncertainty budget for the TWSTT method in nanoseconds (1 σ):

Ionospheric delays	2.0
Equipment delays	1.5
UTC(k)-1 pps REF [two stations]	1.0
TWSTT accuracy	2.7

This would be the best achievable accuracy of the TWSTT method, with similar stations more than 10,000 km apart, provided a global coverage of the satellite's antenna.

A test of consistency, which is also a check of the computation options, is to calculate the deviation from the closure between the three stations. For this purpose, daily independant TWSTT results have been built up at common epochs: for two of the three

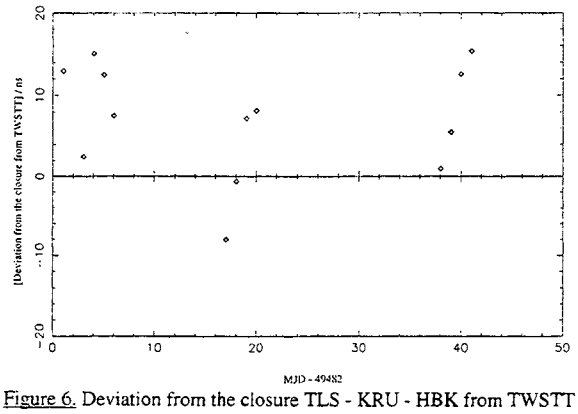


Figure 6. Deviation from the closure TLS - KRU - HBK from TWSTT through Inmarsat [transatlantic phase of the CE-GPS experiment].

TWSTT links, the results were obtained by interpolating between two independant measurements sessions on both sides of the date of the third link result. The deviations from the closure are plotted in figure 6. Again, the behaviour of the signal generators and the ionospheric delays are suspected to account for the most important part of the bias, either the bias between the three different sets of data or the deviation from the closure. The results are scattered between - 8 ns and + 15 ns. These values are consistent with the estimated accuracies given above. But the weak number of sessions do not allow us to perform a comprehensive analysis.

6.2. GPS time restitution.

For the restitution of the GPS time, one week of 15 s sampled [REF(k) - GPS(i)] data were available from the NR106 receivers, for any visible satellite i at each station k. Only few satellites could be tracked simultaneously compared to the number of channels available on the NR106 receivers. A similar behaviour is observed for every station, every satellite, and from day to day: the signature of a white phase noise starting after a period of 240 s, as previoulsy noticed [5,6]. In figure 7 are plotted the modified Allan standard deviations of [REF(TLS) - GPS(i)] computed from the TLS clock compared to the GPS time as restituted by the 24 satellites tracked on MJD 49486. Three of the satellites are either a Block I survivor (PRN 12), or Block II satellites with S.A. parameters set to 0 (PRN 15 and 20). These satellites exhibit a global white phase noise behaviour starting from 15 s, along with some 'observation

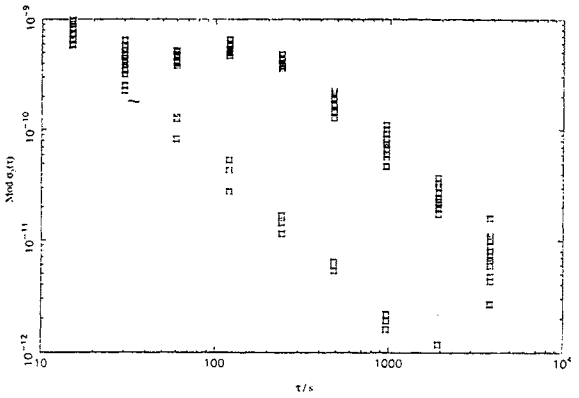


Figure 7. Frequency stability of the TLS clock compared to the GPS time as restituted by the 24 tracked satellites during MJD 49486 [transatlantic phase of the CE-GPS experiment].

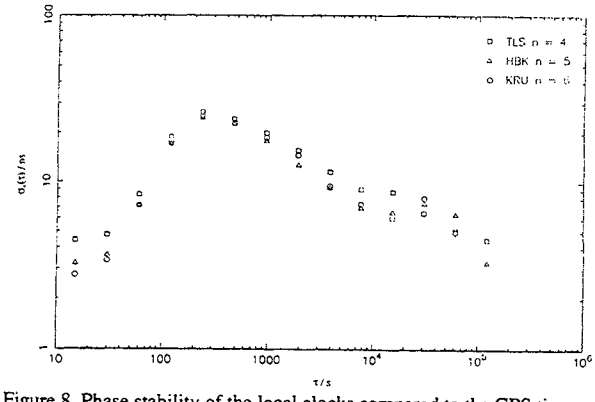


Figure 8. Phase stability of the local clocks compared to the GPS time as restituted by the mean of n simultaneously tracked satellites [transatlantic phase of the CE-GPS experiment].

noise' that can also be noticed on all other Block II satellites.

This behaviour allows the computation of a mean value over a suitable period so as to limit the contribution of the S.A. and observation noises. The measured noise at 240 s for any satellite is about 55 ns. Compared to 84 ns previously observed [6], it shows that the magnitude of the S.A. noise was not the same during the period of the CE-GPS measurement. In figure 8 are plotted the Allan phase standard deviations of the [REF(k) - GPS](15 s) averages over n simultaneously tracked satellites for the three stations involved. The number of satellites simultaneously tracked was only 4 at TLS, 5 at HBK and 6 at KRU. But the magnitude of the noise over the average was roughly the same for all the stations. The rising observed around a period of 20,000 s could come from a periodic behaviour of either the S.A., or the signal generators of the CE-GPS stations, or even from the GPS time itself. A longer observation period would allow a better confidence interval (not plotted in figure 8) around each plot, for a more comprehensive study of this phenomenon.

Setting at 10 ns the limit for the contribution of the noise to the accuracy of the GPS time restitution, in the worst case where n = 4 (TLS), the observations must be averaged over 90 min. The global uncertainty on the STANAG ionospheric delays, is estimated to be 20 ns. It is assumed that half of this value is included in the noise, and that only half of it remain as a bias. The accuracy of the GPS time restitution every 90 min can be estimated in nanoseconds (1 σ) by:

S.A. + observation noise	10
Observation bias [ionospheric delays]	10
Accuracy of the GPS time restitution	14

The CE-GPS requirement is about a signal issued by the geostationary payload. Assuming that the simulation of a space

clock can be achieved within 10 ns [3], the time difference between an event issued by the geostationary payload and an equivalent event issued by any GPS satellite is estimated within 17 ns (1 σ). This value is given considering the S.A. behaviour as observed during the transatlantic phase of the experiment, which seems to be slightly different in magnitude from what was previously observed [5, 6]. If the ionospheric delays could be mesured with an uncertainty negligible compared to 10 ns, the accuracy of the GPS time restitution could even be better.

These results have been used to compare the three monitoring clocks of the CE-GPS experiment. TWSTT data and data issued from the GPS time restitution are plotted over a common period in figures 9, 10 and 11. There is an intermission of roughly one day at the KRU CE-GPS station, but only for the GPS data. Moreover, there is obviously a daily periodic term, of about 50 ns in magnitude, in the data computed from the KRU station (figure 9 and 11). A shift between the two sets of data can also be noticed in figure 10, probably due to the use of different ionospheric models for the two methods. From the accuracy of the GPS time restitution given above, a time transfer using these data would be achieved within 20 ns (1 σ). This value is above the limit of the CE-GPS requirements. But the comparison of these two methods makes this accuracy appear rather conservative: when used for time transfer, some of the satellites included in the computation of the GPS time are in common-view from both stations, and the contribution of the S.A. noise is then 0 for these particular links.

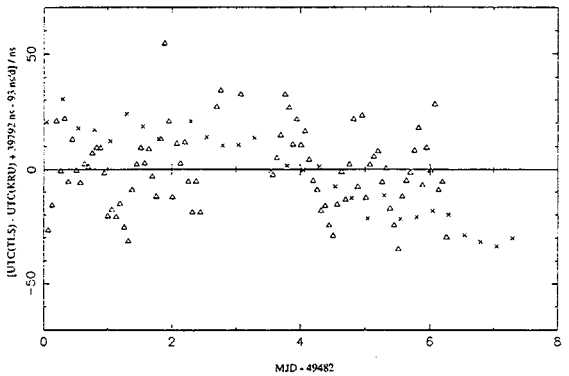


Figure 9. UTC(TLS) - UTC(KRU) from TWSTT through Inmarsat (x) and from the GPS time restituted in each station (Δ) [transatlantic phase of the CE-GPS experiment].

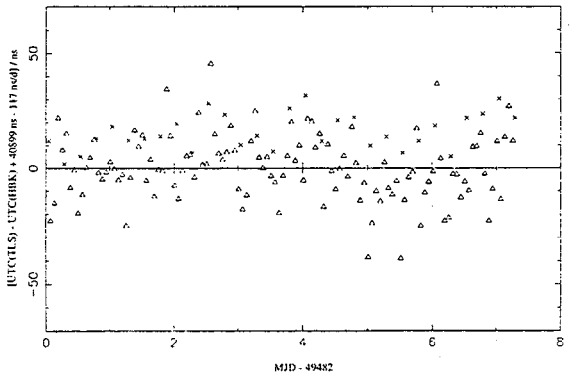


Figure 10. UTC(TLS) - UTC(HBK) from TWSTT through Inmarsat (x) and from the GPS time restituted in each station (Δ) [transatlantic phase of the CE-GPS experiment].

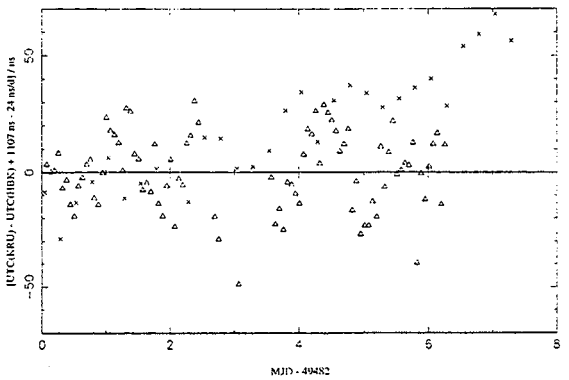


Figure 11. UTC(KRU) - UTC(HBK) from TWSTT through Inmarsat (x) and from the GPS time restituted in each station (Δ) [transatlantic phase of the CE-GPS experiment].

The deviations from the closure, plotted in figure 12, were computed for either method over the same period. The Allan's standard deviation of the closure from the GPS time restitutions shows a white noise signature over a period of 12 h. It allows an averaging of these data over that period, to be more consistent with the daily mean values of the closure computed from the TWSTT. The TWSTT data are scattered between + 2 and + 15 ns, while the GPS time data are between - 3 and - 28 ns. These deviations are probably mainly due to the ionospheric delays. The fact that all GPS data are below the 0 theoretical closure, and that all TWSTT data are above 0 in figure 12, is probably also related to the models used for the computation of the ionospheric delays. It appears yet that these results are consistent with the estimated accuracies.

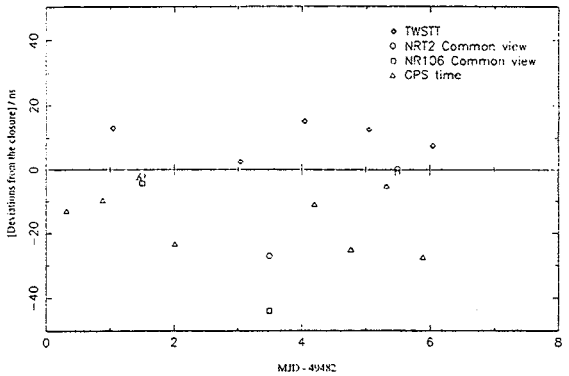


Figure 12. Deviation from the closure TLS - KRU - HBK from TWSTT through Inmarsat (+), from the GPS time restituted in each station (Δ), or from GPS Common-views by NRT2 (o) or NR106 (□) receivers [transatlantic phase of the CE-GPS experiment].

6.3. GPS Common-view.

There were only 10 daily Common-views between TLS and HBK, 14 between TLS and KRU, and only 6 between KRU and HBK, either for the NRT2 or the NR106. The statistical analysis of the Common-views showed a white or a flicker phase noise behaviour over a period of about 2 days, leading to 3 mean values for one week of data. In figure 12 are also plotted the deviations from the closure computed for GPS Common-views, with data from the NRT2 and the NR106 receivers. Two of the data issued from both receivers exhibit a very small similar deviation, unlike the third one. The difference between the two sets of data could again be related to the behaviour of the signal generators, which is not entirely smoothed in the computation of the closure from the Common-views involving the NR106. Our estimation of the accuracy of each link is about 16 ns (1 σ), given an uncertainty on the differential ionospheric delays of about 14 ns (1 σ), which should not be reduced in the computation of the closure.

With so few data, no further analysis can be achieved. It is obvious that, without any ionospheric measurements, the accuracy of the GPS Common-view method for long baselines is largely over the requirements of the CE-GPS.

Conclusion

The time requirements of the CE-GPS could be fullfilled with ground stations comparable to those built up for the purpose of the experiment presented in this paper. Despite the instability of the signal generators, the synchronization of the monitoring clocks was

performed using the TWSTT technique through an Inmarsat geostationary satellite. It was possible to achieve an accuracy of 6.6 ns (1 σ) for a 800 km baseline, or an accuracy between 7 and 9 ns (1 σ) for stations roughly 10,000 km apart. The best achievable accuracy is estimated to 2.7 ns (1 σ), but reliable ionospheric measurement systems are needed. It is anticipated that the Inmarsat III constellation will provide a two frequency downlink, which would allow to reduce this part of the uncertainty budgets.

The ionospheric delays are also the most degrading part of an uncertainty budget for GPS Common-view over long baselines, or for the restitution of the GPS time, provided the S.A. noise can be averaged out. This was the case with the behaviour of the S.A. during the period of the CE-GPS experiment. The averaging method provides an accuracy estimated to 14 ns (1 σ), for the GPS time restituted on the ground every 90 min by tracking simultaneously at least 4 satellites. But again, reliable ionospheric measurement systems are needed, for all the methods of clock synchronization using GPS, to achieve the CE-GPS requirements.

Beside TWSTT and GPS, other time transfer techniques could be used with signals transmitted by the Inmarsat transponder. Among others, the servo-control could be the most interesting due to its near real-time time transfer capability. In this case however the synchronization is less accurate than the best achievable with post-processed data.

The real-time aspects were not the purpose of that experiment. It is nevertheless obvious that, at the publication date of this paper, only the TWSTT technique would achieve the CE-GPS requirements about the monitoring clock synchronization in near real-time.

Acknowledgements.

The authors gratefully acknowledge Dr C. Thomas (BIPM) for many fruitfull discussions. Co-workers on the CE-GPS experiment, G. Fréon and P. Blondé (LPTF), J. Barbier, M. Deleuze, J-L. Issler, D. Berges, J-P. Lefèvre, N. Suard, T. Trémas, R. Tribes, and C. Yven (CNES), F. Varieras and F. Saffre (IN-SNEC), are here acknowledged. The authors thank J. Berthier (Observatoire de Paris) for his help in the processing of the raw data.

References.

[1] J. Barbier, T. Tremas (CNES). European Complement to GPS: presentation of the concept and experimental system. Proceedings of the 7th EFTF, Neufchâtel, March 93.

[2] P. Uhrich, B. Juompan, G. Fréon, R. Tourde, P. Blondé (LPTF), M. Brunet, J-F. Dutrey, J. Barbier, J-L. Issler, N. Suard, T. Trémas (CNES). Time aspects of the European Complement to GPS: some experimental results. Procee dings of the 8th EFTF, Munich, March 94.

[3] J. Barbier, M. Deleuze, J-L. Issler (CNES), P. Uhrich, B. Juompan (LPTF), *et al.* European Complement to GPS: main experimental results. Proceedings of the ION-GPS 94, Salt Lake City, September 94.

[4] STANAG. Technical characteristics of the NAVSTAR GPS, June 91.

[5] D.W. Allan (Allan's TIME) and A. Lepek (Tech Projects). Trends in International timing. Proceedings of the 7th EFTF, Neufchâtel, March 1993.

[6] C. Thomas (BIPM). Real-time restitution of the GPS time. Proceedings of the 7th EFTF, Neufchâtel, March 1993.

[7] W. Lewandowski, G. Petit and C. Thomas (BIPM). Preci sion and Accuracy of GPS Time Transfer. IEEE Trans. on Instr. and Meas. Vol. 42, No. 2, 1993, 474-479.

[8] D. Kirchner (TUG), H. Ressler (IWF), P. Grudler, F. Bau mont, C. Veillet (OCA), W. Lewandowski (BIPM), W. Han son (NIST), W. Klepczynski (USNO), P. Uhrich (LPTF). Comparison of GPS Common-view and Two-way Satellite Time Transfer Over a Baseline of 800 km. Metrologia 1993, 30, 183-192.

[9] P. Uhrich, B. Juompan, R. Tourde (LPTF), M. Brunet, J-F. Dutrey (CNES). Time aspects of the European Complement to GPS: Continental and transatlantic experimental phases. Proceedings of the 26th Annual PTTI Meeting, Reston, De cember 94.

Frequency and Time Transfer with Geodetic GPS Receivers: First Results

P. Baeriswyl, Th. Schildknecht, J. Utzinger, G. Beutler

Astronomical Institute
University of Berne / Switzerland

Abstract

The standard method for time transfer with the Global Positioning System (GPS) is the so-called Common-View technique: using the CA-code, ionospheric delay corrections, precise orbits and accurate station coordinates a precision of a few nanoseconds can be achieved over intercontinental baselines when integrating over a period of one day.

In geodetic applications the GPS receivers acquire all observables of all satellites in view. The advantages are obvious: in addition to the CA-code from one satellite and two receivers (Common-View method) the phase and P-code on both carriers from all available satellites and all stations of a network may be processed in an interferometric mode (geodetic method). This results in an accuracy of a few millimeters over long baselines provided precise orbits are used. Therefore the phase between two receiver clocks may be monitored to within a few 0.01 nanoseconds, consequently time transfer on the sub-nanosecond level seems feasible.

Two GPS time transfer terminals were built to transfer time between two stations. Each one is equipped with a modified Ashtech Z12 receiver which meets some specific requirements.

In addition to these modifications the temperature of the receivers and the electronics is stabilized in both terminals. Keeping the temperature within 1 degree, the internal delays do not change more than 100 picoseconds.

Several tests on short baselines (including zero baselines) were performed. Their results are very promising and proof the concept.

Introduction

Since more than ten years the Global Positioning System (GPS) is used for transferring time between remote clocks. The accuracy in the 'Common View Mode' is a few nanoseconds for time transfer and one part in 10^{14} for frequency intercomparision after several days of integration even over very long distances (Lewandowski, 1991).

By exploring the full potential of the GPS system (use of all GPS observables of all available satellites) we aim to improve the time and frequency transfer by one order of magnitude even on long baselines. We use geodetic receivers to acquire phase and P-code on both frequencies and the C/A-code. The use of a geodetic software allows to process the data in an interferometric mode. The idea was first discussed in (Schildknecht et al., 1990) where the authors showed that the phase data is essentially determining the drift and that the code observations are determining the initial clock offset.

In particular we make use of

* Ashtech Z12 receivers modified for time transfer purposes,
* precise orbits stemming from the International GPS Service for Geodynamics (IGS) (Beutler et al. 1994),
* the Bernese GPS Software (Rothacher et al., 1993) modified for time transfer purposes.

This results in the following advantages:

* Using a geodetic GPS receiver the antenna coordinates can be obtained within one centimeter with respect to a well established reference frame like the International Terrestrial Reference Frame (ITRF) (Boucher et al., 1991) after a maximum of one day of observation.
* Since the receivers use all available observables the ionosphere can be estimated by forming a linear combination of the two frequencies.
* The use of IGS orbits allows to reduce synchronization errors stemming from orbit errors to 200 ps on baselines of up to 5000 km since the orbit accuracy is of the order of 0.2 m (Bauersima, 1983).

Data Processing Technique and Experimental Setup

The Bernese GPS Software makes use of so-called single difference observations. These are the differences of the observations of one satellite measured from two receivers.

The corresponding observation equations for code (c'_{lk}) and phase (p'_{lk}) are as follows:

Code observation equation:

$$\Delta d_k + (c + \dot{d}_k) \cdot \Delta t_k + (\Delta\delta_{lk})_{ion} + (\Delta\delta_k)_{trop} - c'_{lk} = v_{clk} \ ,$$

Phase observation equation:

$$\Delta d_k + (c + \dot{d}_k) \cdot \Delta t_k + (\Delta\delta_{lk})_{ion} + (\Delta\delta_k)_{trop} + N_l\lambda_l - p'_{lk} = v_{plk} \ ,$$

where

l	carrier index (1,2);
k	epoch index;
Δd_k	slant range difference between satellite S and the two stations R1 and R2 involved;
Δt_k	synchronization error between the two receiver clocks;
c	speed of light;

$\dot{d}_k$ radial velocity of satellite S with respect to R1 at time t_k;
$(\Delta\delta_{lk})_{ion}$ differential ionospheric range correction;
$(\Delta\delta_k)_{trop}$ differential tropospheric range correction;
N_l initial ambiguity parameter;
λ_l wavelength of carrier l;
c'_{lk}, p'_{lk} code respectively phase observation;
v_{clk}, v_{plk} residuals.

The parameters of interest in these equations are the series of receiver clock synchronization terms Δt_k. Using the precise orbits from IGS and appropriate troposphere models we may assume that the geometry and the tropospheric refraction correction (i.e. Δd_k, $\dot{d}_k$, and $(\Delta\delta_{lk})_{ion}$) are perfectly 'known'. The ionospheric range correction $(\Delta\delta_{lk})_{ion}$ is frequency dependent and may thus be determined by using both carrier frequencies.

For the phase observations the only critical unknowns are the parameters N_l. Apart from these ambiguities we expect the precision of the determined clock synchronization errors to be of the order of the measurement noise which is a few millimeters! This means that the syntonization error may be estimated with the corresponding accuracy even over short time intervals.

On the other hand we may use all code measurements to estimate the two unknowns N_l yielding finally the 'absolute' synchronization errors Δt_k. This means that we 'smooth' the noisy code measurements (rms of a few nanoseconds) by means of the accurate phase measurements.

For a detailed discussion see (Schildknecht et al., 1990).

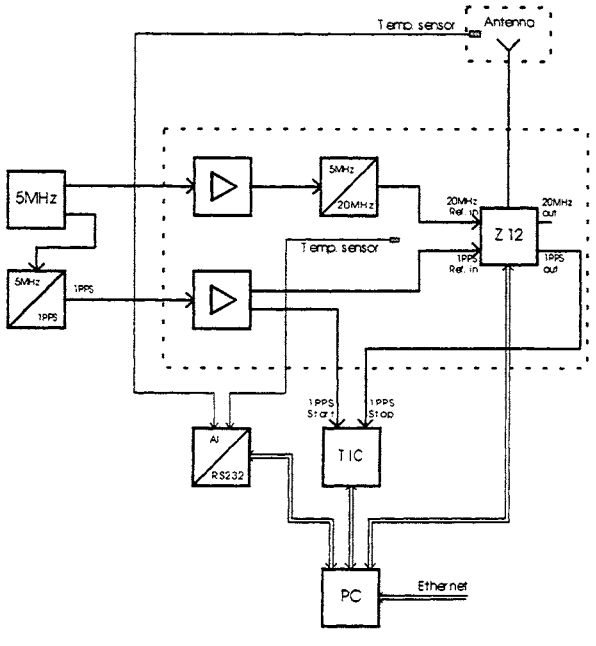


Figure 1: *Functional layout of a prototype of the Geodetic GPS Time Transfer Terminal. Components within the doted boxes are temperature stabilized (antenna: ±3 degree, receiver and electronics: ±0.2 degree).*

Since we want to compare external clocks (and not receiver clocks) we must know the relation between the receiver internal clock and the external (station) clock. We also have to make sure that the receiver clock is the same for all observables (in particular for phase and code).

In order to have 'access' to the internal clock the Z12 receivers were modified by Ashtech in the following way:

* The receivers internal reference frequency base is replaced by an external reference (20 MHz, sin).
* The receivers internal 1 PPS is replaced by an external 1 PPS. This reference second has to be synchronized to GPS time within 1 µs.
* For monitoring and calibrating purposes the receivers are equipped with a reference frequency output and a 1 PPS output.
* In order to stabilize the receiver temperature the Z12 power management was modified (unused processors are not switched off).

These modifications guarantee a perfect syntonization of both the receiver and the station frequency base. The receiver internal 1 PPS is strictly related a) to this fundamental frequency and b) to the 1 PPS of the station clock.

More information concerning these modifications and the synchronization principle of the receiver with the reference clock may be found in (Baeriswyl, 1993).

Figure 1 shows a functional layout of the prototype time transfer terminal:

* The GPS receiver is driven by the station clock.
* The receiver internal delay of the 1 PPS signal is measured with a time interval counter (Stanford SR620).
* The temperature of the GPS receiver and the signal distributing electronics is stabilized.
* The temperature of the antenna will be stabilized.
* The system is controlled with a local PC. The 1 PPS delay and the temperatures are measured continuously (e.g. once per minute) and the GPS data are downloaded from the receiver once per day.
* The PC is linked to a remote computer from where the time transfer terminal may be controlled and where all the data are downloaded.
* The GPS data from two or more stations are processed on a central computer with a modified version of the Bernese GPS Software.

Some of the above listed points may change for the operational system in order to keep the hardware setup simpler. For instance if tests show that the temperature dependencies of the antenna preamplifier and / or the receiver can be modeled with a sufficient accuracy it would not be necessary to put these parts in a temperature stabilized box (doted box).

Test Measurements and Results

The following questions have been investigated:

* Precision of clock synchronization determined from phase and P-code measurements on a zero baseline.
* Temperature dependence of phase and P-code synchronization error on a zero baseline.
* Precision of clock synchronization determined from phase and P-code measurements on a short baseline (0.5 m).

* Relation between synchronization error between different observations.
* Repeatability from session to session.
* Repeatability after switching off one receiver.

All tests were made with a common external frequency base (rubidium standard) for both receivers. GPS data were acquired every 30 sec. The time interval between the 1 PPS outputs of the two receivers as well as receiver and room temperature were recorded for controlling purposes. The minimal session length was 12 h. Figure 2 illustrates a simplified experimental setup for the zero baseline tests. Short baseline tests were performed with almost the same setup. The only difference was that each receiver had its own antenna.

The antenna temperature was not stabilized since variations in the delay cancel out on the short baseline. Assuming that both antennas experience the same temperature variations and that the dependence of the delays from the temperature is the same for both antennas (which should be true according to information from Ashtech (Baeriswyl, 1994)).

All following results were obtained with the Bernese GPS Software where a few modifications had to be implemented to derive the clock synchronization. Clock synchronization errors (Δt_k) were computed in the following way:

$$\Delta t_k = \frac{1}{n}\sum_{i=1}^{n} \Delta t_k^i \ ,$$

where

k	epoch number;
n	number of satellites contributing a synchronization error at epoch k;
Δt_k^i	synchronization error at epoch k determined from observation of satellite i.

On average 5 satellites were used to compute this mean value. Figure 3 to Figure 10 show plots of the synchronization errors (Δt_k) for the different tests and the different observables.

In general we find that:
* The rms of a single synchronization error determined from the phase observations is smaller than 20 ps on the zero baseline and smaller than 35 ps on the short baseline.
* The rms of a single synchronization error determined from the code observations is smaller than 6 ns on the zero baseline and smaller than 10 ns on the short baseline respectively.
* The rms of a single synchronization error determined from P1 shows sometimes values up to 40 ns. These outliers (approximately 20 per day) may easily be rejected.
* All computed synchronization errors have a value of about 60 ns (this common offset was eliminated on all figures) although both receivers were synchronized to the same station clock to within 0.5 ns. This offset is due to the two following effects:
  - A contribution of 50 ns may be explained by the fact that the receiver internal 1 PPS is generated at a leading edge of the 20 MHz reference but delayed by 3 - 4 20 MHz cycles with respect to the incoming 1 PPS (Baeriswyl, 1994). External measurements with a time interval counter showed that the delay between the 1 PPS receiver input and the 1 PPS receiver output had values of about 207 ns for Z12A and 261 ns for Z12B respectively.
  - The remaining 10 ns may be ascribed to receiver internal electronic and filter delays.

We may repeat that all synchronization errors computed with the phase observations contain an arbitrary offset with respect to the synchronization error determined from the code data.

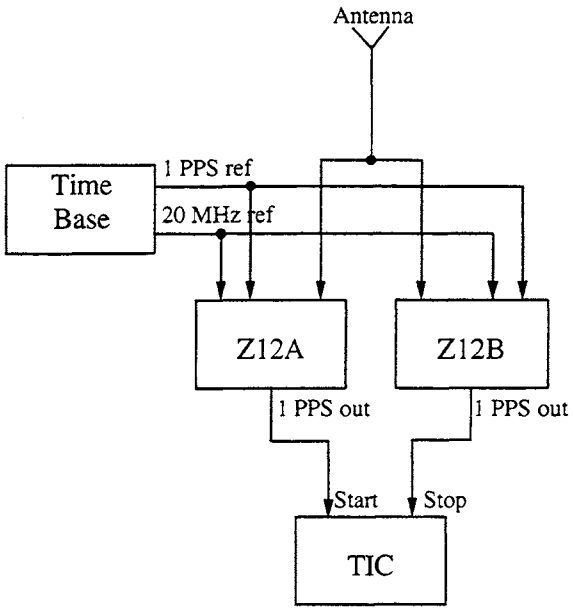


Figure 2: *Simplified scheme of the experimental setup for the zero baseline tests.*

Let us discuss the different tests:

0-baseline and constant Z12-temperature: Figure 2 shows the experimental setup for this test: two receivers are connected to the same antenna and have a common time base. With a time interval counter (TIC) we measure the interval between the receivers 1 PPS outputs.

Figure 5 indicates a peak-to-peak variation of the synchronization error (Δt_k) over a session of one day of about 15 ps for L1 and 25 ps for L2, respectively. The corresponding value for the P1-code observations over the same period is 3 ns (outliers rejected) and 2 ns for P2-code observations, respectively (see Figure 3 and Figure 4).

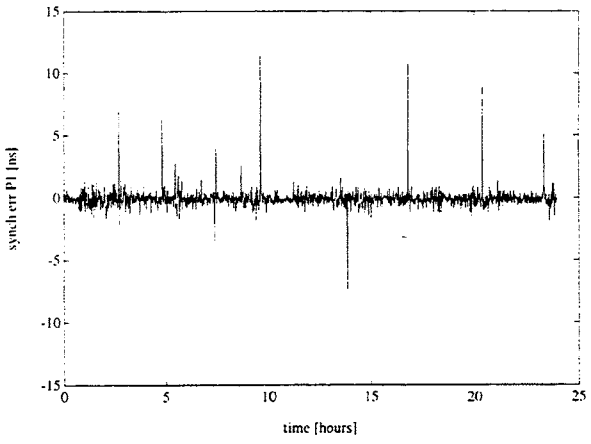


Figure 3: *Synchronization error computed with P1 on a zero baseline with constant temperature. An offset of about 60 ns has been removed.*

The rms for the phase measurements is about 2 ps and 3 ps, respectively. The corresponding values for the code observations are more than two magnitudes larger: 0.8 ns for P1 and 0.6 ns for P2, respectively. The error of the estimated clock synchronization error (this is simpliy the mean value) is less than 100 ps after 10 minutes of observation.

The results from P1 and P2 show an offset of about 3.5 ns. This difference originates in the different filter delays for L1 and L2 (Baeriswyl, 1995)

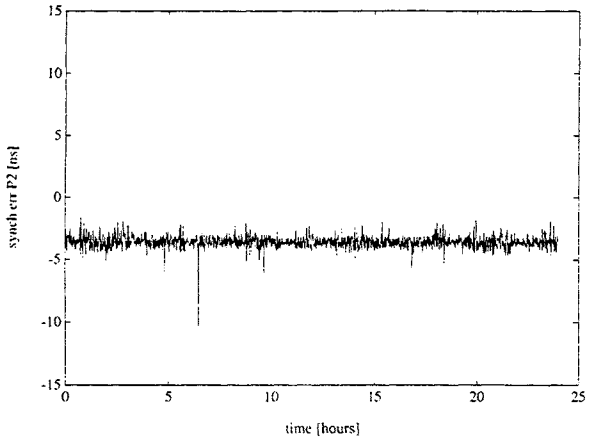


Figure 4: *Synchronization error computed with P2 on a zero baseline with constant temperature. An offset af about 60 ns has been removed. Compared to the synchronization error computed with P1 (see Figure 3) a difference of 3.5 ns can be observed.*

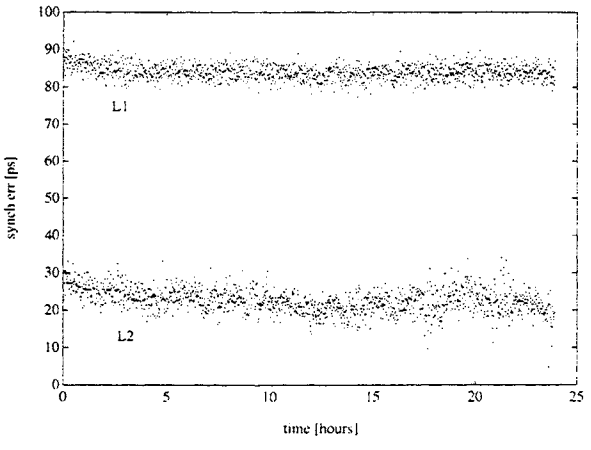


Figure 5: *Synchronization error estimated from L1 and L2 on a zero baseline with constant temperature. Both synchronization errors are affected by an arbitrary offset.*

0-baseline and variable Z12-temperature: The experimental setup is the same as for the previous test (Figure 2) with the exception that the temperature of Z12B was increased by 12 ^{0}C (see Figure 6).

Neither the peak-to-peak behaviour nor the rms do change significantly when the temperature of one receiver is changed.

Nevertheless the results shown in Figure 7 and Figure 8 are surprising at first glance since the synchronization error has a positiv temperature coefficient for the phase (ca. +20 ps/^{0}C) but a negativ for the code observations (ca. -100 ps/^{0}C). However, the delays for phase and code have a completely different origin: the phase delay is a function of delays in the local oscillator divider chain, whereas the code delay is a function of group delays through filters and delays in the digital section (Baeriswyl, 1995).

In addition we notice a sudden change of the phase delays (or rather the synchronization errors determined with the phases) at approximately 8:00 (Figure 8). This is probably due to a mechanical shock experienced by one of the receivers. The effect may be neglected given its smallness compared with the rms of the synchronization errors determined by the code.

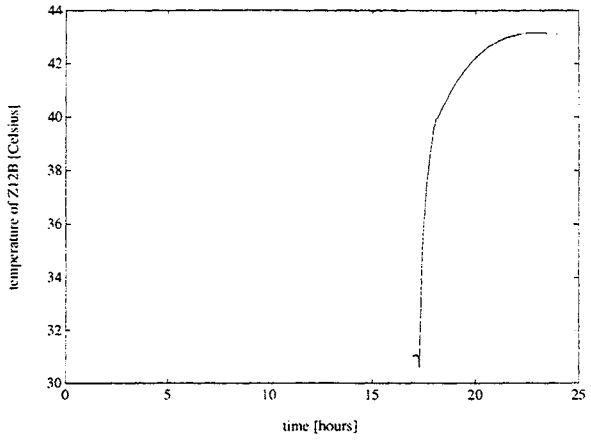


Figure 6: *Temperature evolution on the surface of the receiver Z12B. (Before 17:00 the temperature was not measured but held constant at ±1 on both receivers).*

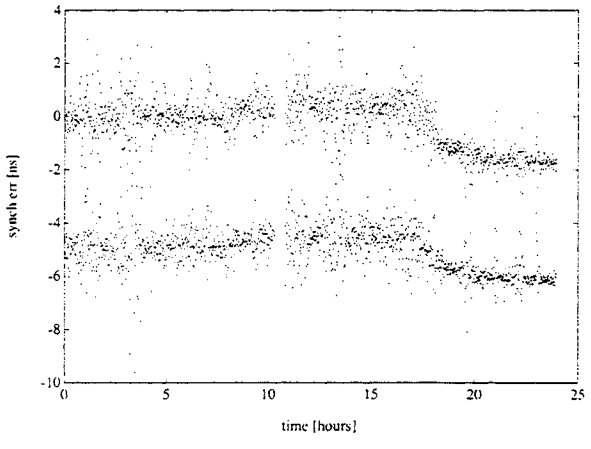


Figure 7: *Synchronization error computed with P1 and P2 on a zero baseline with variable temperature (see Figure 6). An offset of about 60 ns has been removed.*

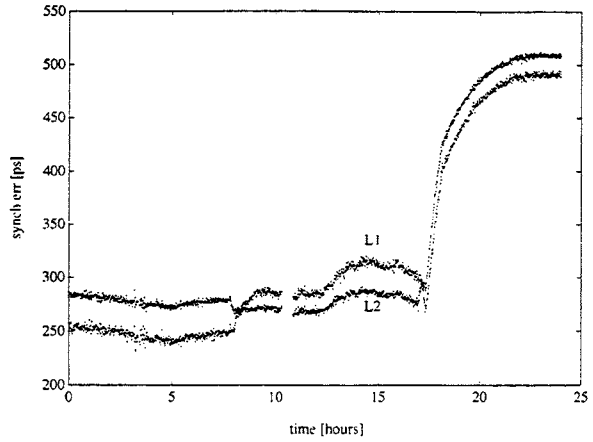


Figure 8: *Synchronization error computed with L1 and L2 on a zero baseline with variable temperature (see Figure 6). Both synchronization errors are affected with an arbitrary offset.*

Short baseline: As an the previous test a common time base was used but this time each receiver had its own antenna. The baseline between the two phase centers was 50 cm.

On a short baseline we expect additional noise stemming from the antenna preamplifier and the antenna cables since both cancel on a zero baseline. We also want to examine the behaviour of the 'absolute' synchronization error computed with the code: are these values the same for the zero baseline and for the short baseline setup (assuming that antenna cable delays were already corrected)?

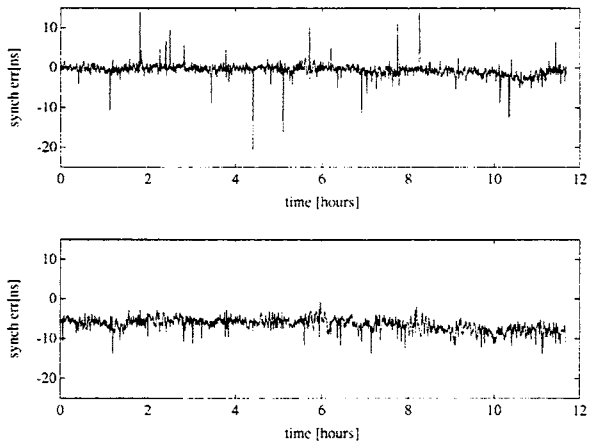


Figure 9: *Synchronization error computed with P1 and P2 on a short baseline and constant temperature. An offset of about 60 ns has been removed. A difference of 5 ns between P1 and P2 can be observed.*

The rms of the synchronization error determined by P1 is about 1.0 ns. The rms for P2 is 1.4 ns. Although somewhat worse the corresponding error of the mean after 10 minutes of observation is still less than 100 ps! The carrier phase data shows a rms of 5 ps for L1 and L2 which is only slightly larger compared with the zero baseline data.

In Figure 9 we notice small systematic trends of about 2 ns. We have no firm explanation for these changes. Possibly these are multipath effects. This could be verified by acquiring data over several days and checking for repeated signatures (the geometrical satellite constellation repeats after one sideral day for a given site).

Due to these systematics the rms on both code observations increases for longer observation intervals (1.8 ns for P1 and 1.7 ns for P2, respectively).

For the synchronization error estimated by the phases (Figure 10) there is a large systematic trend at the beginning of the session (during the first two hours). This may be explained by the fact that the second antenna was connected to the receiver immediately before the start of the session and increased the temperature of the preamplifier and therefore changed the signal delays.

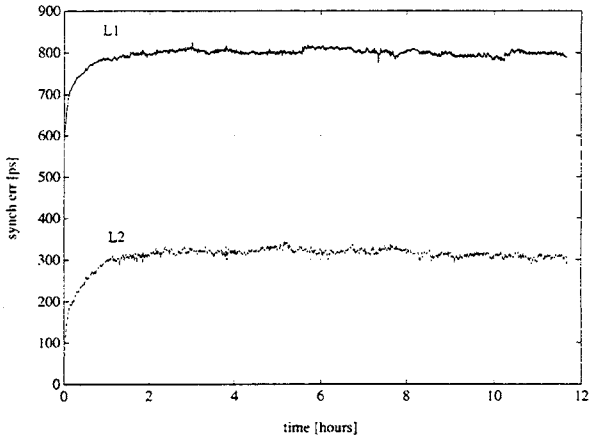


Figure 10: *Synchronization error computed with L1 and L2 on a short baseline with constant temperature. Both synchronization errors are affected with an arbitrary offset.*

Conclusions

In order to proof the feasibility of sub-nanosecond time transfer with geodetic GPS receivers a series of test was conducted. Special attention was given to potential temperature dependencies of the receiver internal and the antenna delays. Therefore the experimental setup of the time transfer terminal includes temperature control of the receiver and the antenna. Because we focused in these first tests on the critical effects stemming from the time transfer terminal itself all experiments were performed on short baselines (including zero baselines) and with a common time base.

The results show that sub-nanosecond time transfer is clearly possible under these conditions even after a few minutes of observation only.

In order to achieve this precision all receiver and antenna delays must either be stabilized by controlling the temperature or by applying corresponding corrections as a function of the temperature. Possible humidity dependent changes of the delays (which were not yet studied) should be also be taken into account. In addition we assume that the geometry and the tropospheric delay corrections are 'perfectly' given by the International GPS Service for Geodynamics (IGS).

In a next phase experiments over longer baselines will be necessary. We expect in particular the multipath effects on the P-codes to be the limiting factor.

References

Baeriswyl, P., Th. Schildknecht, W. Gurtner, G. Beutler (1993). Frequency and time transfer with geodetic GPS receivers. *Proceedings of the 7th European Frequency and Time Forum*, Neuchâtel, 1993.

Baeriswyl, P. (1994). Private communcations Ashtech.

Baeriswyl, P. (1995). Private communcations Ashtech.

Bauersima, I. (1983). *Radiointerferometrische Satellitenbeobachtungen (GPS II).* Mitteilungen der Satellitenbeobachtungsstation Zimmerwald, Nr 10, Astronomisches Institut der Universität Bern, Schweiz, 1983.

Boucher, C., Z. Altamimi (1991). ITRF 89 and other realizations of the IERS Terrestrial Reference System for 1989. *IERS Technical Note 6*, April 1991, Observatoire de Paris.

Boucher, C., Z. Altamimi (1991). The IERS GPS Terrestrial Reference Frame. *IERS Technical Note* No. 6, April 1991, Observatoire de Paris.

Lewnadowski, W., C. Thomas (1991). GPS time transfer. *Proceedings of the IEEE 79*(7).

Rothacher, M., G. Beutler, W. Gurtner, E. Brockmann, L. Mervart (1993). *Bernese GPS Software, Version 3.4.*

Schildknecht, Th., G. Beutler, W. Gurtner, M. Rothacher (1990). Towards subnanosecond GPS time transfer using geodetic processing techniques. *Proccedings of the 4th European Frequency and Time Forum*, Neuchâtel, 1990.

SESSION 3

MATERIALS I

Chairman : J.P. AUBRY

ON THE CURRENT STATE AND PROBLEMS OF DETERMINATION OF THE NONLINEAR CONSTANTS OF QUARTZ

Carl K. Hruska

Piezoelectricity Research Laboratory
Department of Mathematics and Statistics, York University
North York, Ontario M3J 1P3, Canada

ABSTRACT

Nine independent reasons are described that can negatively affect the quality of the estimated values of the third-order nonlinear material constants of α-quartz.

An active presence of one of them - the substitution - has been positively identified in the process of determination of the eight existing sets of the third-order elastic, electroelastic, electrostrictive, and dielectic constants defined at zero strain and electric field. This constitutes a sufficient reason to generally consider both the values and the standard errors of the respective nonlinear constants distorted and unreliable.

The wealth of more than 200 experimental data accumulated over the past decades by means of various experiments can be re-used and the nonlinear constants recalculated to remove the distortions using a new strategy which is described.

I INTRODUCTION

The main body of knowledge about the electromechanical nonlinearities in α-quartz is represented by the tensors of the third-order elastic, electroelastic, electrostrictive and third-order dielectric constants (third-order permittivity) defined in the natural state and referenced to zero strain and dc field. These are third-order nonlinear electromechanical constants briefly referred to as the nonlinear constants.

At this point each of the four tensors has been determined two times, by independent experiments and in a complete (or a very nearly complete) form[1,2,3]. A review of errors frequently made in the process of determination of the nonlinear constants presented here allows for a critical evaluation of these tensors, in particular with respect to their desired 'material' character. This is the main objective of this work.

A method to substantially improve the quality of the existing values of the nonlinear constants and to simultaneously employ all existing experimental data to produce their single, complete and reliable set is also described.

II DEFINITION OF THE NONLINEAR CONSTANTS

The nonlinear material constants of crystals are defined by means of the following constitutive equations

$$(1)\quad T_{IJ} = c_{IJKL}\cdot\eta_{KL} - e_{NIJ}\cdot E_N + \frac{1}{2}\, c_{IJKLMN}\cdot\eta_{KL}\cdot\eta_{MN} - \frac{1}{2}\,\varepsilon_0\cdot l^{rel}_{MNIJ}\cdot E_M\cdot E_N + f_{MIJKL}\cdot E_M\cdot\eta_{KL} - \frac{1}{2}\, f_{MIJKL}\cdot\eta_{IJ}\cdot\eta_{KL}\,,$$

$$(2)\quad D_M = \varepsilon_{MN}\, E_N + e_{MIJ}\,\eta_{IJ} + \frac{1}{2}\,\kappa_{MRS}\cdot E_R\cdot E_S + \varepsilon_0\cdot l^{rel}_{MNKL}\cdot E_N\cdot\eta_{KL}\,,$$

where T_{IJ}, η_{KL}, D_M, E_N are, respectively, the components of the stress tensor, the Cauchy-Green finite strain tensor, the electric displacement vector, and the electric field intensity vector, all defined in the reference state of the crystal. The material tensor components in the above constitutive equations - in the order of their appearance - correspond to

- c_{IJKL} - the elastic stiffness tensor,
- e_{NIJ} - the piezoelectric stress tensor,
- c_{IJKLMN} - the third-order elastic stiffness tensor,
- l^{rel}_{MNIJ} - the relative electrostrictive tensor,
- f_{MIJKL} - the electroelastic tensor,
- ε_{MN} - the dielectric permittivity tensor,
- κ_{MRS} - the third-order dielectric permittivity tensor.

The constant ε_0 is the permittivity of free space.

All uppercase indices in Eqs.(1) and (2) take on the values of 1,2,3, while the Einstein summation rule is in effect. The relative electrostrictive tensor l^{rel}_{MNIJ} appearing in the constitutive equations is usually replaced by the total electrostrictive tensor l_{MNIJ} according to Nelson[4]

$$l^{rel}_{MNIJ} = l_{MNIJ} - (\,\delta_{MN}\delta_{IJ} - \delta_{MI}\delta_{NJ} - \delta_{MJ}\delta_{NI}\,)\,,$$

where δ_{MI} is the Kronecker delta. The relative electrostrictive tensor l^{rel}_{MNIJ} differs from the total electrostrictive tensor by the Maxwell vacuum electric-stress tensor (Ref.4, Eq.(54))

$$\delta_{MN}\delta_{IJ} - \delta_{MI}\delta_{NJ} - \delta_{MJ}\delta_{NI}\,.$$

The introduction of l_{MNIJ} to replace l^{rel}_{MNIJ} generally simplifies the algebra in the applications of the constitutive equations.

Contracting the upper case tensor indices of the material tensor components f_{MIJKL}, l_{MNIJ}, κ_{MRS} and c_{IJKLMN} to their lower case matrix form and making use simultaneously of the interchange symmetry among them converts these material tensor components, respectively, to the electroelastic constants f_{ijk} (i = 1,2,3, j,k = 1,2,...,6), the electrostrictive constants l_{ij} (i,j = 1,2,...,6), the third-order permittivities κ_{ijk} (i,j,k = 1,2,3) and the third-order elastic constants c_{ijk} (i,j,k = 1,2,...,6).

The constants f_{ijk}, l_{ij}, κ_{ijk} and c_{ijk} are called the third-order nonlinear material constants of crystals. In general, the total number of independent values of these constants can be as high as 165. In

the case of α-quartz their total number is only 31, the reduction being due to its crystal symmetry.

Making the conventional choice of the independent constants, the nonlinear constants of quartz consist of fourteen third-order elastic constants

(3) c_{111} , c_{112} , c_{113} , c_{114} , c_{123} , c_{124} , c_{133} c_{134} , c_{144} , c_{155} , c_{222} , c_{333} , c_{344} , c_{444} ,

eight electroelastic constants

(4) f_{111}, f_{113}, f_{114}, f_{122}, f_{124}, f_{134}, f_{144}, f_{315} ,

eight electrostrictive constants

(5) l_{11}^{rel} , l_{12}^{rel} , l_{13}^{rel} , l_{14}^{rel} , l_{31}^{rel} , l_{33}^{rel} , l_{41}^{rel} , l_{44}^{rel} ,

and one third-order dielectric constant

(6) κ_{111} .

The 31 material constants (3)-(6) are the fundamental third-order material constants of quartz in the reference frame according to Ref.5.

III COMPUTING THE NONLINEAR CONSTANTS FROM EXPERIMENTAL DATA

The nonlinear constants of quartz are computed from experimental data using overdetermined linear systems and the linear regression commonly referred to as the least-squares method or the least-squares fit. This method produces estimates of the nonlinear constants and their standard errors.

Past experience indicates that the theoretical relationship of m independent observations of an experimental quantity E to the constants of quartz can be recorded using a system of linear equations

$$(7)\quad E_i - A_i = M_{i1}.x_1 + M_{i2}.x_2 + \ldots + M_{ij}.x_j + \ldots + M_{in}.x_n ,$$

where i = 1,2,...,m.

In Eq. (7) E_i represents the individual observations of the experimental quantity E; the coefficients M_{ij} as well as the absolute terms A_i are functions of the linear material constants of quartz, crystallographic orientation of the quartz experimental specimens, and of the geometry and other characteristic features of the particular experiment which produces observations E. Their numerical values are fully calculable and known when system (7) is formulated. The symbols x_j represent the sought nonlinear material constants (3)-(6).

Defining

$$L_i = E_i - A_i$$

as the left-hand side of the above equations and **M** as the matrix of coefficient M_{ij}, the above system can be recorded as

$$(8)\quad \mathbf{L} = \mathbf{M} \,.\, \mathbf{X} ,$$

where **L** and **X** are column vectors with components L_i and x_j, respectively.

It is generally accepted that the observations of E_i are not perfectly accurate because of experimental errors. To minimize the effect of the errors, the number m of observations E_i is made as large as possible. It is always larger than the number n of the unknown nonlinear constants to be determined from the linear system. The errors in E_i then make the linear system overdetermined.

The values of the unknown nonlinear constants x_j are sought from system (8) using a data reduction process called the least-squares fit. It provides the best estimates of the values of the nonlinear constants x_j and their standard errors $\vartheta(x_j)$

$$(9a)\quad x_j = (\mathbf{M}^T.\mathbf{M})^{-1}.\mathbf{M}\,.\mathbf{L}$$

$$(9b)\quad \vartheta(x_j) = \sqrt{[(\mathbf{M}^T.\mathbf{M})^{-1}.SS_R/(m-n)]} ,$$

where SS_R is known as the residual sum of squares[6] and j = 1, 2, ..., n.

IV WHAT CAN GO WRONG

Nine independent reasons have been identified as possible causes for the above procedure to produce incorrect results. They are:

1) errors in theory;
2) errors in linear constants;
3) errors in crystallographic orientation angles of the quartz specimens used;
4) systematic experimental errors;
5) under-representation;
6) small number of degrees of freedom;
7) arbitrary multipliers;
8) near colinearity; and
9) substitution and mechanical interpretation of the subsequent solution.

These reasons are analyzed below.

1.-3. Errors in the theory, linear constants and orientation angles

For the least-squares solution of system (8) to produce undistorted estimates of the nonlinear constants and valid standard errors, the values of matrix elements M_{ij} and absolute terms A_i should be determined with absolute accuracy. For this reason

(A) their theoretical definitions representing the implementation of the nonlinear theory should be correct;

(B) they should be calculated using correct linear (elastic, piezoelectric and dielectric) material constants of quartz; and

(C) they should be calculated using exact orientations of the experimental specimens used.

Severe violations of (A) have been quite common in the past and the possibility that some of them, probably minor ones, are still waiting to be spotted cannot be disregarded. Moreover, all implementations of the nonlinear theory are likely to be only approximate so that requirement (A) can hardly be fully satisfied.

It is no surprise that the choice and gradual improvements of theory may have a considerable effect on the estimates of the nonlinear constants. A comparison of results in Refs.7 and 8 or in 9 and 10 can serve as a suitable illustration.

Given the multitude of existing sets of linear constants of quartz, it is difficult to say how to satisfy (B) by selecting the correct set. The same applies to (C) where the likelihood of errors is further increased when the crystallographic orientations of quartz specimens needed to study the nonlinearities are unusual for their suppliers.

The uncertainties in the linear constants and orientation angles are sources of errors which are frequently overlooked. However, their effect is not totally trivial. A preliminary study of the problem has been made by Hruska[11].

4. Systematic experimental errors

According to the assumptions of validity of the least-squares process, the only uncertainty allowed in system (8) are random experimental errors in observations E_i. They are the sole reason for the nonzero values of the standard errors $\vartheta(x_j)$.

Consequently, the experimental values E_i should be free of systematic errors.

The elimination of systematic errors requires in a number of cases a very special effort; classical examples here are the creation of a well defined uniaxial stress[12] or a uniform dc electric field[13] in the cubes of quartz during the transit-time experiments; in some cases the success could not be guaranteed (e.g., sufficient uniformity of a lateral dc field in the resonator experiment[14]).

5. Under-representation

To avoid excessively large standard errors $\vartheta(x_j)$, each nonlinear constant x_j should be well represented in the linear system (8); i.e., associated with coefficient M_{ij} the value of which is as large as possible. This increases the contribution of each constant to the experimental value E_i which thereby increases the accuracy with which the constant can be determined from the experiment. Coefficients M_{ij} can be maximized by proper design of experimental samples.

Large standard errors are known to accompany the current estimates of some of the electrostrictive and third-order dielectric constants[2]. They are most likely caused by under-representation. The small size of some of the coefficients M_{ij} accompanying these constants can be seen in Ref.10.

6. Small number of degrees of freedom

For the standard errors $\vartheta(x_j)$ to be reliable and have a simple meaning it is necessary that the number of degrees of freedom, $m - n$, be made as large as possible; 10 or less (used in some studies) is definitely inadequate. The same applies to the numbers of observations made of which E_i is the mean value; this assures that the random errors in E_i can be treated as normally distributed.

7. Arbitrary multipliers

A necessary condition of validity of the least-squares solution (9a,b) is that the left-hand sides L_i, as random quantities, have the same standard deviation. This implies that the standard deviations of all experimental values E_i should be equal. If this is not the case, then, to rectify the situation, each equation in (7) is to by adjusted by a suitably chosen multiplier k_i. Linear system (7) is then

$$k_i.L_i = k_i.M_{i1}.x_1 + k_i.M_{i2}.x_2 + \ldots + k_i.M_{ij}.x_j + \ldots + k_i.M_{in}.x_n ,$$

where, generally, $k_1 \neq k_2 \neq \ldots, \neq k_m$. The solution of the new system is then found using the same expressions (9a,b) as before. The application of the multipliers is called 'weighting' and the 'ordinary' least-squares process with $k_i = 1$, $i = 1,2,\ldots,m$, is thereby replaced by the 'weighted least-squares'[6].

It appears that in some cases these multipliers are assigned to individual equations on the basis of an algebraic convenience rather then their true function. This leads to an unwarranted uncertainty as to what the formally correct solution of linear system (8) is. An excellent illustration can be found in Table I of Ref.15. The uncertainty in the calculated nonlinear constants given there reaches tens of percent. In general the arbitrary choice of multipliers can lead to serious errors.

8. Near colinearity

A phenomenon frequently present in matrix **M** of system (8) is the colinearity of its columns. It may be difficult to detect analytically because the algebraic definitions of the matrix elements M_{ij} are rather complex. [A single case of colinearity predicted (using numerical means) by Hruska and Khogali[16] has been demonstrated analytically by Kittinger and Tichy[17].]

The colinearity alone presents no problem. It only means that the least-squares process cannot be executed because matrix $\mathbf{M}^T.\mathbf{M}$ becomes singular and its inverse $(\mathbf{M}^T.\mathbf{M})^{-1}$ needed in Eqs.(9a,b) cannot be found. The corrective action required here is simple.

The danger associated with the colinearity is that it can be turned into a near colinearity when the matrix elements M_{ij} are not computed with sufficient accuracy. Then, the near colinearity will change the singular matrix $\mathbf{M}^T.\mathbf{M}$ into an ill-conditioned matrix which can completely distort the results of the least-squares process or a similar process[18] intended to estimate the values of the nonlinear constants.

The calculated values of the nonlinear constants obtained under this condition have an order of magnitude which is usually inconsistent with the expectations based on the magnitude of the experimental quantity E and they are usually accompanied by excessively large standard errors. The estimates show extremely large changes in response to minute (artificially induced) changes in the values of E_i. Most importantly, they nowhere nearly represent the true values of the sought nonlinear constants. Moreover, different pre-packaged least-squares subroutines are likely to deal with the same ill-conditioned matrix differently and produce different results.

Examples of the unfortunate consequences of this phenomenon date back to 1986 when a disagreement existed of nearly three orders of magnitude among several electroelastic constants of α-quartz found by different authors[9,19,20] with no obvious explanation at hand[21]. The cause, identified by Hruska[22,23] as the near colinearity, rendered most of the least-squares estimates of the electroelastic constants in Refs.19 and 20 invalid. Another demonstrated example[24] of the same problem is the case of an extremely large temperature dependence of the electroelastic constants, about 100 times larger than expected, which was also reported in Ref.19.

The colinearity in matrix M can turn into a near colinearity if the matrix elements M_{ij} are computed using only a single precision. For this reason it is advisable to use the double or quadruple precision.

Matrix M with nearly colinear columns can also be created if the values of the material constants used fail to reflect the quartz symmetry in an exact manner. A classical example is the relation for the elastic moduli

$$c_{66} = 0.5\,(c_{11} - c_{12}) . \qquad (10)$$

Using $c_{11} = 86.74$, $c_{12} = 6.99$, in combination with $c_{66} = 39.88$, all in 10^9 N/m^2 (Ref.25), can create nearly colinear columns; it is safer to use $c_{66} = 39.875$ or let the computer do the calculation using Eq.(10).

In general, matrix **M** should be tested for the near colinearity and suitably adjusted prior to the execution of the least-squares process.

9. Substitution and mechanical interpretation of the subsequent solution

It seems to be the accepted opinion that any linear system (8) is used most effectively if those nonlinear constants in it which have been determined earlier are substituted for by their published values; also if all nonlinear constants believed to make a negligible contribution in system (8) are eliminated from it. Since the elimination of a constant from

system (8) is equivalent to the substitution of a zero for its value, one can talk only about substitutions.

Traditionally, the substitutions are made prior to the least-squares fit which is then executed for the nonlinear constants remaining in the system to conclude the solution process. Its results are then interpreted mechanically as the best estimates and standard errors of these remaining constants.

It is not difficult to recognize that the substitutions must in some way affect the least-squares solution of the system. However, what the effect actually is, has been surprisingly little cared for.

A recent study made by Hruska[26] shows that this neglect has been completely unwarranted. The substitutions lead to a number of undesired phenomena which have so far been completely overlooked.

In system (8), each nonlinear constant is associated with one column vector of system matrix **M**. If a substitution is made for a certain constant, then the impact of it on the subsequent least-squares estimates and standard errors is defined by the properties of the column vector with which the constant is associated. The deciding factor is whether this vector can be expressed as a linear combination of the rest of the column vectors of the system matrix (a) wholly, (b) partly, or (c) not at all. A detailed analysis of this has been made in Ref.26, where the three cases are labeled Case 1, 3, and 2, respectively.

According to Ref.26, the undesired phenomena resulting from the premature substitutions are

in Case 1: a misinterpretation of the least-squares estimates and of the standard errors and a possible distortion of the estimates;

in Case 2: a loss of valuable information and a distortion (and a likely inflation) of the standard errors;

and in Case 3: a misinterpretation of the least-squares estimates and of the standard errors, loss of potentially valuable information, a possible distortion of the estimates, and a distortion (and a likely inflation) of the standard errors.

For these phenomena to be noticed, put under control or avoided, the traditional sequence of the solution steps should be reversed. The process should start with the least-squares fit performed on the original system. The substitutions should follow and be made, if still required, into the results of the least-squares process. The mechanical acceptance and interpretation of results of the least-squares process should be avoided at all cost.

The only difficulty connected with selecting a real-life illustration for these undesired phenomena is the large number of them in literature. Two of them, together with the necessary analysis and explanation, can be found in Refs.2, 26, 27 and 28.

V AVAILABLE TENSORS AND THEIR INTERDEPENDENCE

The first complete tensor of the third-order elastic constants of α-quartz (3) was published by Thurston, McSkimin and Andreatch[1] in 1966. Their values have quickly attained wide acceptance and have been the sole source of the third-order elastic constants ever since.

The experimental data for their determination were obtained by observations of the travel time of acoustic pulses propagating through bulk quartz. This method is known as the transit-time method. Changes in the travel time were measured under the effect of hydrostatic and unidirectional pressure. The experiment produced 34 observations collected by McSkimin, Andreatch and Thurston[29] and Thurston, McSkimin and Andreatch[1].

These constants are the first nonlinear constants obtained using a rigorous application of the nonlinear theory. However, as was natural at that time, the theory ignored the piezoelectric character of quartz. It disregarded any possible effect on the calculated third-order elastic constants due to the remaining nonlinear phenomena: electroelasticity, electrostriction and third-order permittivity. It was much later that this effect was recognized as very substantial. For this reason these constants must today be referred to as 'effective' rather than 'material' constants.

The current values of the remaining three tensors of nonlinear constants (4)-(6) have been determined independently using two different experiments. These experiments were originally intended to determine only the electroelastic constants[30,9]. However, it was found out that neither experiment could be properly interpreted without the general inclusion of the electrostrictive and third-order dielectric constants[31,32] and that each experiment should in fact be used to determine all three types of nonlinear constants. In order to obtain these constants, it was necessary to use the third-order elastic constants from Ref.1.

Historically older of the two methods, the resonator method, consists in observations of the series resonance frequency of piezoelectric resonators and its changes occurring when the resonators are subjected to a dc field bias[33]. A total of 184 observations obtained by Hruska and Khogali[16], Kinigadner[34], Hruska, Merigoux and Kucera[35], and Hruska and Brendel[7] have been selected as a source of the most recent set of the electroelastic, electrostrictive and third-order dielectric constants calculated on the basis of this method. Their current values were computed by Hruska: first[36] using the linear constants by Bechmann[25], then[2] - for a better consistency with Ref.1 - with the linear elastic constants according to McSkimin[37]. The results did not include isolated electrostrictive constants l_{12} and l_{44} as the experimental data used did not provide a direct access to these two quantities.

The second of the two methods is a new version of the transit-time experiment, where the observations of the changes in the travel time are caused by a dc biasing field[9]. The most recent values of the electroelastic, electrostrictive and third-order dielectric constants calculated from this experiment have been obtained by Hruska[2]. The experimental set of 24 transit-time observations used originates from Kittinger, Tichy and Friedel[10].

It has always been a source of anxiety to the author of this paper that in order to obtain the values of the electroelastic, electrostrictive and third-order dielectric constants, it was necessary to rely on the effective third-order elastic constants[1]. The importance of this fact was temporarily played down by the huge success of the transit-time and the resonator method which independently produced almost identical values of the electroelastic constants.

The transit-time experiment by Thurston et al. of 1965-66 was repeated in recent years on very high purity quartz by Wang in the course of his Ph.D. work[3]. The aim of the work was to calculate the truly 'material' third-order elastic constants. This was possible relying on the values of electroelastic, electrostrictive and third-order dielectric constants that had become available after 1966. Details of this work have not been released for publication.

VI A SURPRISING RESULT AND ITS DOMINO EFFECT

The impressive agreement[2] between the electroelastic constants obtained independently by the transit-time and resonator method boosted confidence in their absolute accuracy as material constants. Moreover, the agreement was also interpreted as an indication that the distortion of the electroelastic constants (due to the effective third-order elastic constants used in the computations) had to be

negligible.

A simple numerical experiment made by Hruska[38] refuted the above beliefs. The analysis of the transit-time and resonator experiments was repeated using an arbitrarily chosen set of third-order elastic constants. Predictably, the two experiments produced two new, substantially different sets of the electroelastic constants. Being based on an arbitrary set of third-order elastic constants, both sets of the electroelastic constants had to be wrong. Surprisingly, the differences between the two sets remained unchanged. The same was true also for all electrostrictive and third-order dielectric constants calculated simultaneously.

This surprising outcome means that the agreement between the results of the two experiments cannot be interpreted as a verification of the quality of the calculated constants; it is merely an indication of an inherent consistency of the two methods. It offers nothing to support the belief that the electroelastic constants (and other nonlinear constants obtained simultaneously) are the true material constants of quartz.

In combination with what has been said in the previous two sections the following conclusions cannot be avoided:

1) The effective third-order elastic obtained by Thurston et al[1] may be viewed as obtained from a properly formulated linear system (8) where the values of the unknown electroelastic, electrostrictive and third-order dielectric constants have been eliminated, i.e., substituted for by zeros. As noted, substitutions generally have an adverse effect on the the values and standard errors of the remaining nonlinear constants calculated from the system. Because of the order of magnitude of the eliminated constants, this effect is likely to be substantial and the third-order elastic constants in Ref.1 are not reliable as material constants.
2) The current[2] nonlinear constants of electroelasticity, electrostriction and third-order permittivity, obtained using the transit-time method and the resonator method, were computed using the third-order elastic constants taken from Ref.1. This was done by substituting a set of effective and, therefore, unreliable values into linear systems (8) formulated for the two methods. As a result, the current values of the electroelastic, electrostrictive and third-order dielectric constants, together with their standard errors, are not the desired material constants of quartz.
3) The third-order elastic constants obtained by Wang[3] had to be calculated with the help of some - current or earlier but definitely incorrect - values of the constants mentioned in 2). Therefore, they themselves are likely to be afflicted by the same problem.
4) Hence, all nonlinear constants in Refs. 1-3 as well as their standard errors are likely to be unreliable.

VII WHAT NEEDS TO BE DONE - A SUGGESTED STRATEGY

Past experience indicates that:

1) on its own, no single method (experiment) can lead to the determination of all 31 nonlinear constants of quartz (3)-(6); at least two different experiments are required; one experiment can only produce the values of some isolated constants and of some linear combinations of other constants;
2) in order to resolve the linear combinations and/or to determine all constants pertaining to one of the four tensors, it is necessary to know some other nonlinear constants beforehand; there is no source to take the values of these constants from for the first experiment; and
3) the undesired effects described in the previous section can be completely avoided if the least-squares fit is executed on system (8) with no substitutions made.

For these reasons it is suggested to utilize the accumulated experimental data as follows:

1) Solve system (8) for each method/experiment independently, making no substitutions. In this manner results will be obtained from each experiment representing solely its true and undistorted contribution to the knowledge of the nonlinear constants. Also, the values of standard errors will not be misinterpreted, distorted or inflated.
2) Having obtained such results from two (or more) different experiments, search for any disagreement among them. If found, such a disagreement represents a conflict between the two experiments with no confusing participation and distorting effects from any external factors. Explain the disagreement or remove its cause.
3) Once all significant conflicts have been dealt with, combine all individual linear systems into one large system and find its solution. This should produce a single set of nonlinear constants and their standard errors consistent with all available experimental data.
4) If the combined system fails to produce all 31 linear constants in isolation or if the standard errors of some nonlinear constants are too large, additional observations using better-designed experimental samples (or a completely new experiment) may be necessary.

A first step in the right direction has already been made. The solution of system (8) for the data obtained by the resonator method was obtained without any substitutions. The same was done for the data obtained by the transit-time experiment using dc field interactions. As a result, 15 nonlinear quantities have been obtained[2] ready for comparison with results obtained independently by the two methods, all with valid standard errors.

VIII CONCLUSION

Nine problems which can negatively affect the quality of the existing values of nonlinear constants of α-quartz are: errors in the theory used, errors in the linear constants used, errors in the crystallographic orientation angles of the experimental specimens used, systematic experimental errors, under-representation of some nonlinear constants, small number of degrees of freedom, the use of arbitrary multipliers, near colinearity, and the substitution and mechanical interpretation of their subsequent solutions.

Substitutions have been identified as a sufficient reason to conclude that none of the eight sets of the nonlinear constants discussed in this paper are reliable material constants of quartz. The same applies to their standard errors. As these constants represent a major portion of what is known about the electromechanical nonlinearities in quartz in general, improvements are definitely necessary.

To start with, the nonlinear constants of quartz can be re-computed using the experimental data accumulated over the past 30 years. To avoid the error of substitution the proposed calculation method should be used. This may also help to remove the current under-representation of some constants and thus reduce their excessively large standard errors.

If this does not produce satisfactory results, additional data will have to be collected and added to the existing observations.

While the adverse effect of substitutions seems to be sufficient to make the recalculation of the

nonlinear constants of quartz necessary, it is by no means the only reason for which the constants may be incorrect.

This paper discusses the determination of the third-order nonlinear electromechanical constants of α-quartz. However, its principles are generally applicable to the determination of any material constants of crystals where the least-squares method is used.

ACKNOWLEDGEMENTS

The author wishes to acknowledge the assistance of Pauline Hruska with the final stages of preparation of this paper. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada and by the Faculty of Arts, York University.

REFERENCES

[1]Thurston R. N., McSkimin H. J. and Andreatch P., Jr, J.Appl.Phys. 37(1966), pp.267-275.

[2]Hruska C. K., Proceedings of the 5th European Frequency and Time Forum, Communautes Scientifiques de Besancon et de Neuchatel, Besancon, March 1991, pp.137-142. Available from E.N.S.S.M., Universite de Besancon, France.

[3]Wang Q., "Nonlinear elastic properties and vibrational anharmonicity of very high purity quartz", Ph.D. Thesis, University of Bath, January 1993.

[4]Nelson D. F., J.Acoust.Soc.Am. 63(1978), pp.1738-1748.

[5]IEEE Standard on Piezoelectricity, IEEE Std 176-1978. Available from The Institute of Electrical and Electronics Engineers, Inc., 345 East 47th Street, New York, NY 10017.

[6]Draper N.R. and Smith H., "Applied regression analysis", J.Wiley, New York 1966, pp.59-63 and 77-81.

[7]Hruska C. K. and Brendel R., J.Appl.Phys. 65(1989), pp.715-717.

[8]Hruska C. K., Brendel R., J.Appl.Phys. 67(1990), pp.1676-1679.

[9]Reider G. A., Kittinger E., Tichy J., J.Appl.Phys. 53(1982), pp.8716-8721.

[10]Kittinger E., Tichy J., and Friedel W., J.Appl.Phys. 60(1986), pp.1465-1471.

[11]Hruska C. K., J.Appl.Phys. 75(1994), 77-83.

[12]Wang, private communication.

[13]Reider G. A., Kittinger E., Tichy J., J.Acoust.Soc.Am. 71(1982),pp.1283-1284.

[14]Hruska C. K., Proc. of the 2nd European Frequency and Time Forum, Communautes Scientifiques de Besancon et de Neuchatel, Neuchatel, March 1988, pp.621-627. Available from E.N.S.S.M., Universite de Besancon, France.

[15]Kittinger E. and Tichy J., J.Appl.Phys. 69(1991), pp.7910-7911.

[16]Hruska K. and Khogali A., IEEE Transactions on Sonics and Ultrasonics, 18(1971), pp.171-176.

[17]Kittinger E. and Tichy J., J.Acoust.Soc.Am. 91(1992), pp.1484-1488.

[18]Marquardt D. W., 'Numerical solution of nonlinear algebraic equations', Research Project 611830, Engineering Research Laboratory, E. I. du Pont de Nemours & Co. (Inc.), 1958.

[19]Kusters J. A., Proceedings of the 24th Annual Frequency Control Symposium (AFCS), April 1970, pp.46-54. Available from NTIS, Springfield, VA, as document AD746210.

[20]Brendel R., J.Appl.Phys. 55(1984), p.608.

[21]Kittinger E., Tichy J. and Friedel W., J.Appl.Phys. 56(1984), pp.2584-2585.

[22]Hruska C. K., Proc. of the 31st AFCS, June 1977, pp.159-170. Available from NTIS as document ADA088221.

[23]Hruska C. K., J.Appl.Phys. 61(1987), pp.1127-1129.

[24]Hruska C. K., Transactions on Ultrasonics, Ferroelectrics and Frequency Control, 35(1988), pp.637-640.

[25]Bechmann R., Phys.Rev. 110(1958), pp. 1060-1061.

[26]Hruska C. K., J.Appl.Phys. 72(1992), pp.2432-2439.

[27]Hruska C. K., Proc. of the IEEE 1990 Ultrasonics Symposium, Honolulu, Dec. 1990, vol 3., pp.1249-1254. Available from The Institute of Electrical and Electronics Engineers, Inc., New York, NY. 10017.

[28]Hruska C. K, and Ng P., J.Acoust.Soc.Am. 93(1993), pp, 1426-1430.

[29]McSkimin H. J., Andreatch P, Jr and Thurston R. N., J.Appl.Phys. 36(1965), pp.1624-1632.

[30]Hruska C. K., Czech.J.Phys. B12(1962), pp.338-353.

[31]Hruska C. K., Proc. of the 31st AFCS, June 1977, pp.159-170. Available from NTIS as document ADA088221.

[32]Kittinger E, Tichy J. and Friedel W, Jap.J.Appl.Phys. 24(1985), Supplement 24-2, pp.707-709.

[33]Hruska K., Czech.J.Phys. B12(1962), pp.338-353.

[34]Kinigadner A., Ph.D.Thesis, University of Innsbruck, Austria, 1978.

[35]Hruska C. K., Merigoux H., Kucera M., J.Can.Ceram.Soc. 57(1988), pp.53-57.

[36]Hruska C. K., J.Appl.Phys. 68(1990), pp. 41-45.

[37]McSkimin, J.Acoust.Soc.Am. 34(1962), pp.1271-1274.

[38] Hruska C. K., J.Appl.Phys. 73(1993), pp.343-347.

PROGRESS in CRYSTAL GROWTH and ACOUSTIC CHARACTERIZATION of GALLIUM ORTHOPHOSPHATE: $GaPO_4$

D. Palmier*, M. Cochez*, R. Gohier**, C. Bonjour***, G. Marianneau***,
A. Zarembowitch**, E. Bigler***, E. Philippot*

* - L.P.M.S., CNRS, UMII, cc 003, F-34095 Montpellier cedex 05.
** - L.D.R.U., CNRS, Univ. P. et M. Curie, Tour 22, F-75252 Paris cedex 05
*** - L.P.M.O., CNRS, 32, Av. de l' Observatoire, F-25044 Besançon cedex.

ABSTRACT

The last improvements on gallium phosphate crystal growth are presented. Systematic study of growth parameters has demonstrated the feasibility of $GaPO_4$ epitaxy on large berlinite seeds in sulphuric acid with a view to further growth experiments in phosphoric medium. This latest arises to be the most suitable solvent to develop hypothetic industrial applications. Seed lengthening along the Y direction was successfully realized exhausting a manual contact-process of small Z-seeds optically joined together.

Four out of the six single crystal elastic constants of $GaPO_4$ (C_{11}, C_{33}, C_{44}, C_{14}) have been measured between 4 and 300 K applying the ultrasonic pulse superposition method. Excepted for C_{33}, the elastic constant behaviour appears to be separated in two zones, before and after 50 K, where a strong slope variation occurs. Beyond this temperature, the C_{11} and C_{33} elastic constants increase with decreasing temperature, while C_{44} and C_{14} drop monotonically.

Interdigital transducers were applied on a polished Z-cut by UV-lithography with standard AZ resist and lift-off technique. Various SAW propagation directions are investigated including orientations for which a shear-horizontal skimming bulk wave can be generated.

Finally, possible piezoelectric properties of other MXO_4 quartz like materials not yet synthesized in a form of bulk crystals were extrapolated by means of relations between structure and piezoelectric properties of available materials.

INTRODUCTION

Since a few years, quartz homeotypic gallium orthophosphate has asserted itself as a material presenting very attractive piezoelectric properties; no first order pyroelectricity, high electric resistivity, and existence of temperature compensated cuts towards the angles of maximum coupling coefficient. The crystal growth investigation of gallium phosphate is justified, at least, by two reasons. Indeed, for the same AT cut, the first results obtained for $GaPO_4$ crystals [1,2] indicate a coupling coefficient value (k≈0,14) higher than that of quartz (0.085) and berlinite (0.11). On the other hand, there is no α<–>β quartz transition (573 and 586 °C for quartz and berlinite respectively) for this material: it affords piezoelectric applications in a larger range of temperature (cristobalite transition = 933°C), for filter and sensor devices [3,4]

These investigations concerns $GaPO_4$ crystal growth in various acid media (HCl, H_3PO_4, H_2SO_4) and their mixtures, through the slow heating and vertical reverse temperature methods [3,5]. Systematic study of growth parameters has demonstrated a very strong anisotropy of V_X, V_Y and V_Z growth rates with a growth ratealong the X direction generally greater than the two others, whatever is the acid medium probed. This quite unsatisfactory growth rate disparity can be drastically improved by means of $GaPO_4$ Z-seeds, in purpose to further crystal growth in phosphoric acid medium, where V_X/V_Z ratio value can be close to one. As for α-quartz and berlinite, V_Y growth rate being always very weak, the lengthening of these crystals in such direction is gained obtained using the accurate adjustment of several joined Z-seeds.

[OH]-content, evaluated by infrared and near-infrared spectrometries, arises lower than that of berlinite, synthesized with approximatively the same crystal growth conditions.

Physical properties of gallium orthophosphate below room temperature have not yet been clearly and completely established, principally due to usual interest found to this material for upper temperature range, with the goal of industrial applications as oscillators or sensors devices [3,4]. Particularly, the overall behaviour of elastic constants of $GaPO_4$ was in our knowledge, totally unexplored.

In this work, a first elastic constant set of this material has been measured in the ultrasonic frequency range using piezoelectric transducers linked to the crystal. Pulse echo techniques and lithium-niobate transducers have been employed; the temperature range 4-300 K has been investigated for different propagation directions and polarizations.

At the present time, if experimental bulk wave measurements have allowed to confirm theoretical results concerning thermally compensated orientations [6,7], in our knowledge, only theoretical results have been published for SAW propagation on $GaPO_4$ crystals [8].

In this paper, experimental measurements acquired for surface wave propagation on Z plates of $GaPO_4$ bulk crystal are exhibited. Several delay line devices were deposited along different propagation directions of Rayleigh waves in the XY plane, and velocities extracted using the central frequency of the electrical response. In order to analyse the thermal behaviour of such devices, two oscillators in the frequency range of 70 to 80 MHz were tested and the frequency versus temperature curves measured from 0 to 100°C. A comparison with theoretical values performed from already published material constants [8,9] is discussed.

CRYSTAL GROWTH

I - Obtaining of large $GaPO_4$ seeds

For evident reasons, $GaPO_4$ crytal growth must be, at the outset, undertaken on the largest seeds available. However, due to the purely synthetic crystal growth feature of this material, no seed of such size is still existing. Therefore, preamble the continuous solid solution and the cell volume similarity with the pseudobinary system $Al_{1-x}Ga_xPO_4$, $V = 230 \pm 1 Å^3$ [10], and the large berlinite seeds already synthesized in our laboratory, $GaPO_4$ crystal growth has been contracted on berlinite seeds.

Intending to determine the crystal growth conditions of $GaPO_4$ in acid media, we have undertaken a systematic study in terms of: crystal growth temperature, Tc (°C), temperature difference between cold and hot zone, ΔT (°C), and acid concentration, M ($mol.l^{-1}$).

At the first step, some experiments have been carried out in phosphoric acid medium. Unfortunately, the dissolution kinetics of $AlPO_4$ in this solvent was faster than the growth of $GaPO_4$, so berlinite seeds were dissolved. Therefore, we have tried crystal growth in sulphuric acid medium where a much lower dissolution rate had already been demonstrated [11]. Indeed, all MXO_4 quartz-like materials present their highest solubility in this solvent. Due to too high pressure, the vertical reverse temperature gradient method (GVT) in platinum lined autoclaves, has been prefered to the slow heating method in a glass vessel.

Although the experiments are still in progress, some interesting results can be previously emphasized:

- the confirmation of the growth of $GaPO_4$ from berlinite seeds, with a growth temperature chosen greater than 200°C with the intention of reducing all the problems connected with the presence of impurities, like a strong [OH]-content.
- the confirmation of different kinetic dissolution in both phosphoric and sulphuric acid media.
- the anisotropy disparity of V_X, V_Y and V_Z growth rates leads to strongly different morphology of crystals geted from the most used X and Z berlinite seeds.

At the actual state, the Z-seeds present the greatest features; V_X being the only significant growth rate, flat crystals with a very thin epitaxy of $GaPO_4$ along the Z direction, are synthesized.

To improve the further crystal quality, these $GaPO_4$ plates are preliminary separated from the original seeds and from the small intermediate area where cracks eventually occured.

Then, such free-strain seeds can be used in other growth experiments where V_Z is significant, with one or other of the two different methods already indicated.

II - $GaPO_4$ crystal growth

Some interesting results can be emphasized from many growth experiments carried out at Tc ≤ 240°C in different acid media investigated (H_3PO_4, H_2SO_4 and HCl) and their mixtures:

- 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_3PO_4 acid concentration increases to 20 $M.l^{-1}$.
- in H_2SO_4 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 $GaPO_4$ in this solvent authorise higher growth temperatures and, so, should lead to crystals with lower OH impurity content [5].
- 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 Tc ≤ 240°C. The sulphuric acid is theoretically the best solvent for $GaPO_4$ crystal growth (Tc ≥ 200°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.***

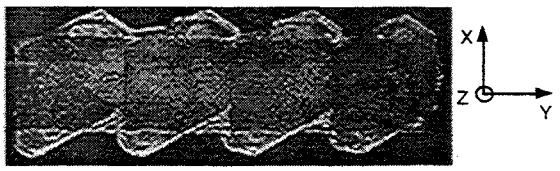
III - Seed lengthening

Excepted for the most concentrated pure phosphoric acid, any of the other acid media investigated has 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 $GaPO_4$ on large berlinite seeds [4], their very weak lengthening along the Y direction cannot be improved without a puerely 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) ± 10° oscillating angle around an horizontal axis perpendicular to the glass vessel, with growth conditions; Tc = 150°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 (Photo A-B).

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 [12,13], i.e., successive growths along both X and Y directions.



(PHOTO A)
front view

(PHOTO B)
side view

PHOTO A / B: $GaPO_4$ crystal obtained in phosphoric acid medium from splicing of 4 $GaPO_4$ seeds.

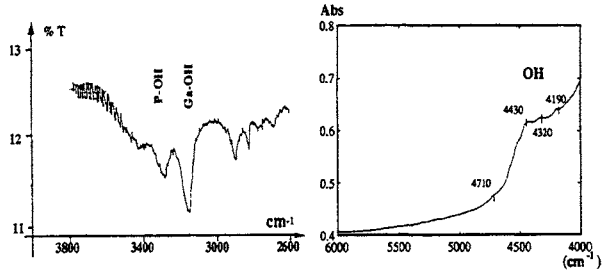
IV - [OH]-content determination by Infrared and near-Infrared spectrometries.

These methods, as for quartz and berlinite, are the most appropriate ones for the [OH]-content determination and can be related to the surtension coefficient. Figure 1a gives an idea of infrared transmission spectrum obtained for such crystals, in the range 4000-2600 cm^{-1} characteristic of [OH] absorption. The absorption coefficient, α, is calculated from the expression [14] :

$$\alpha = (1/d)\,[\log(T_{3800}/T_{3167}) + \log(T_{3800}/T_{3300})],$$

where d is the sample thickness in cm and T_{3167} and T_{3300}, the transmitance values related to gallium (Ga-OH) and phosphorous (P-OH) atoms respectively.

From a general point of view, this α value decreases with the growth temperature from α ≈ 2 for Tc ≈ 160°C to α ≈ 0.06 for Tc ≈ 230°C. There is a significant difference between berlinite, $AlPO_4$, and $GaPO_4$ behaviour. Indeed, for approximatively the same crystal growth conditions, one can observe a diminution of the α value by a factor of 10 for gallium phosphate. This result is of major interest from an industrial point of view because, in this case, high growth temperatures (Tc > 200°C) would not be needful.



(a) (b)

Figure 1

(a) IR transmission of $GaPO_4$ crystals obtained from VTG method, Tc>200°C.

(b) Near IR absorption of $GaPO_4$ crystals obtained from SHT method, 150<Tc<180°C.

On the other hand, another important difference between $AlPO_4$ and $GaPO_4$ is given by near infrared spectrometry. For low growth temperature obtained berlinite (Tc < 180°C), characteristic molecular water absorption has been noted at 5200 cm^{-1}. For gallium phosphate crystals, at the same Temperature, only weak absorption characteristic of -OH groups close to 4500 cm^{-1} is noticed, (**figure 1b**).

FIRST SET DETERMINATION OF SINGLE ELASTIC CONSTANT OF GALLIUM PHOSPHATE

I - Experiments.

The whole work has been realized on a single bulk crystal synthesized by the slow heating method [3,5] exempt from any twins or strains. The sample was placed in a temperature-controlled cryostat, and the transit time of the ultrasonic wave between parallel faces was measured at constant temperatures.

Minimum dimensions were chosen allowing several cuts and polishing with a view to get several Lithium Niobate ($LiNBO_3$) transducers deposits with a thickness of about 300 µm. Theory preconises, for quite relatively thick transducers (λ/2 > 0,1 mm), a direct pasting on the crystal, after metallisation on either face, with an epoxy resin binder. However, such process has generated significant clutters on the ultrasonic response, and so finally has been surrended.

Lastly, we have adopted an indium soldering [15], usually employed with thinner transducers, including vacuum evaporation over both pieces faces, with a first catched chromium layer (from 150 Å up to 300 Å), then gold (1500 Å), and finally indium (2500 Å). Once these metals are applied, both crystal and transducer have been get in touch and put under pressure with about 200 $kg.cm^{-2}$ during twenty or so minutes. Indium has appeared to be one of the only component able to affect expansion coefficient discrepancy between both materials (in the order of $10^{-4}K^{-1}$ for $LiNbO_3$ and $10^{-6}.K^{-1}$ for $GaPO_4$).

Measurements have been performed using a frequency synthesizer connected with an equipment labeled "MATEC" drived with a computer, and a digital oscilloscope. A high frequency electric signal is sent from the synthesizer to the MATEC, in order to be separated in different pulses. Pulse-width and repetition period are controlled by the operator. Practically, the round-trip delay of the ultrasonic pulse propagating into the sample is measured. Indeed, the transducer transmits the pulse up to the crystal, using direct piezoelectric effect.The opposited faces, between which propagation occured, have been preliminary polished in such a way echo system can set up.

As a general rule, two echoes are selected by superintensification and superimposed on a same screen, with a perfect coincidence of the arches of the two sinus curves. Therefore, the shift-time registered for the both responses, arising from the temperature drop, allow the direct achievement of the relative velocity variations.

Nevertheless, if absolute velocity resolution cannot be yielded due to the unknown delay time introduced by binder intercalation, very weak relative variations can be investigated. However, the difference of expansion variation, between transducer and sample released by indium binder, should be taken into account in the acquisition of the relative velocity variations of the sample. Therefore, the expression giving the pulse velocity in terms of sample thickness L and transit time related to the drift of either sinusoides under temperature drop effect :

$$V = 2L/\tau$$

must be rigorously differentiated in order to obtain the relative variation of velocities in terms of temperature variations :

$$1/V\ (\Delta V/\Delta T) = 1/L(\ \Delta L/\Delta T) - 1/\tau\ (\Delta\tau/\Delta T)$$

The linear expansion coefficients at room temperature, α_{11} = 5.3 $10^{-6}.K^{-1}$ and α_{33} = 1.2 $10^{-6}.K^{-1}$ [16] denote a consequent anisotropy but cannot been extropolated at lower temperature. So, in this study, $\Delta L/L$ ratio has been neglected and $\Delta V/V$ directly assimilated to the $\Delta\tau/\Delta T$ ratio.

At the present state, all of the longitudinal propagations have been successfully investigated. Concerning the thickness propagation modes, Z axis has been explored along both polarization directions, but inspection has been limited to Z polarization for the Y propagation. The X shear mode has not been yet prospected.

The complete analysis has also been realized from 4 K up to 300 K, in order to cast off any kind of doubt about the eventual presence of artefacts or irregularities that would not be intrinsic features of the crystal. Such results are not presented here with a view to preserve the whole legibility. In the same way of mind, the only first echo, the most reliable one, has been ploted on the graphs.

II - Results and discussion.

The first analysis of such results point out a quite good agreement between the absolute velocity values [9,18], and ours Table I; indeed, the error range arises to be about less then 0.6 % from our longitudinal velocity values and less then 3 % for the measured transversal modes. Therefore, despite an echo disturbance eventually generated owing to binder presence, further elastic constants acquisition will be directly readable without any important modification scale.

Table I: Compared velocities at room temperature from Ref.[9],[18] with this work.

MODE N°	N	U	MODE	Ref [9] EXPERIMENTAL VALUES ($m.s^{-1}$)	Ref [17] COMPUTED VALUES ($m.s^{-1}$)	This work (EXP.) ($m.s^{-1}$)	Discreapancy Ref[9]/Ref[17]
1	X	X	L*	4352	4344	4337	0.34/0.16
2	Y	Y	QL*	4333	/	4358	0.59/..
3	Z	Z	L*	5348	5344	5327	0.39/0.31
4	Y	Z	QS*	3228	/	3255	0.84/..
5	Z	X	S*	3238	3220	3142	2.96/2.42
6	Z	Y	S*	3247	3220	3249	0.06/0.9

(*) L = Longitudinal, S = Shear, Q = Quasi-Longitudinal, QS = Quasi-Shear

Fig.2 presents at once, measured ultrasonic round trip delay. time versus temperature for X-cut, Y-cut, Z-cut longitudinal modes and corresponding velocities. Equivalent results from Y-cut and Z-cut shear modes polarized approximately along the Z and both X and Y direction, respectively, are presented on figure 3.

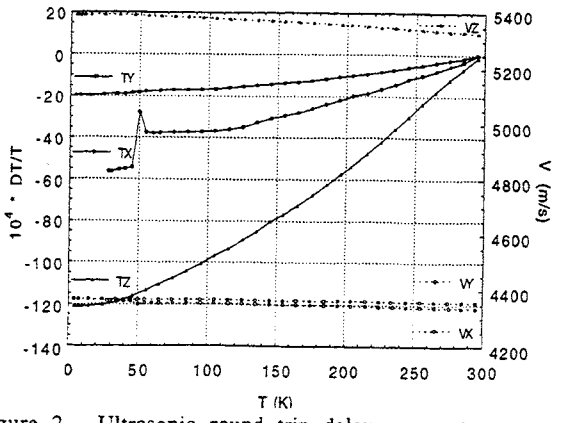


Figure 2 - Ultrasonic round trip delay versus temperature. Longitudinal modes / Propag. X, Y, Z / Frequ. 11.0, 89.0, 33.0 MHz

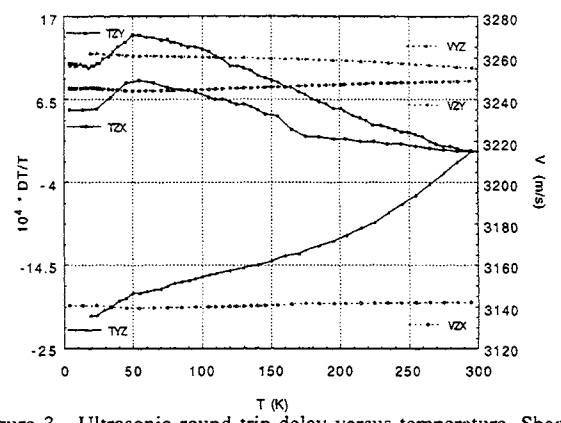


Figure 3 - Ultrasonic round trip delay versus temperature. Shear-modes / Propag. Y, Polar. Z; Propag. Z, Polar. X; Propag Z, Polar. Y / Frequ. 21.2, 18.0, 32.8 MHz

Theory generally preconises smaller network bondings with temperature decreasing ; therefore, the overall shape of such round trip delay-time is quite surprising if we consider that the shear mode with propagation along Z axis differs strongly from other curves. Indeed, it checks an increasing delay time is observed with decreasing temperature, meaning a network relaxation phenomenon for such direction.

However, the most interesting investigation rests in the two-zones settlement from each side of 50 K, except for the elongation-mode with propagation along Z-axis. As a matter of fact, at this temperature, the other modes are subject to slightly slope accentuation as Y longitudinal-mode and Y shear-mode polarization along Z-direction, or sudden reversing slope as X-longitudinal and Z transversal shear modes.

This last Z shear mode presents very closed curves from both X and Y polarizations, in complete agreement with theoreticals results which provide a degeneracy along Z-axis.

Four out of the six effective second-order elastic constants (C_{ij}) have been already acquired using the relations establishing velocity values in terms of elastic constants density.

In Table I, for the mode labeled 1, an additional piezoelectric stiffening term must be taken into account, in the form of the "squared piezoelectric stress constant e_{11} / dielectric constant for constant strain ε_{11}^{S}" ratio. The maximum value of this ratio, calculated from [16], has been evaluated in the worst case, from about less than 1.2 %.

The first set of constants, C_{11}, C_{33}, C_{44}, C_{14}, extracted respectievely from modes 1, 3, 6, 2, relays the accident arising at 50 K, figs. 5, 6, 7, except for C_{33}, fig.4. Indeed, the latest is directly proportional to the overall shape of the Z-longitudinal mode delay time, where any slope shift occured (fig. 2).

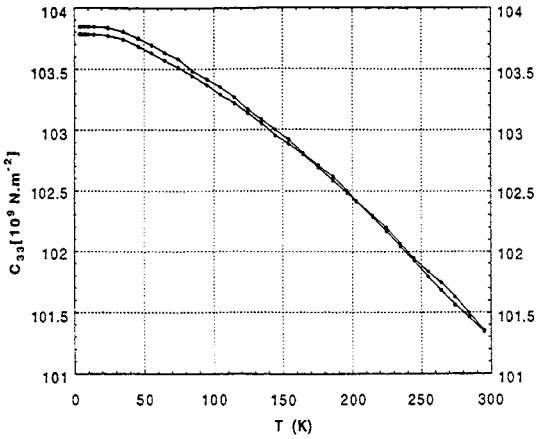


Figure 4 - Zero-field elastic constant C_{33} of $GaPO_4$ versus temperature (*). Thickness of the sample = 10.986 mm

(*) Both echos have been plotted on each graph

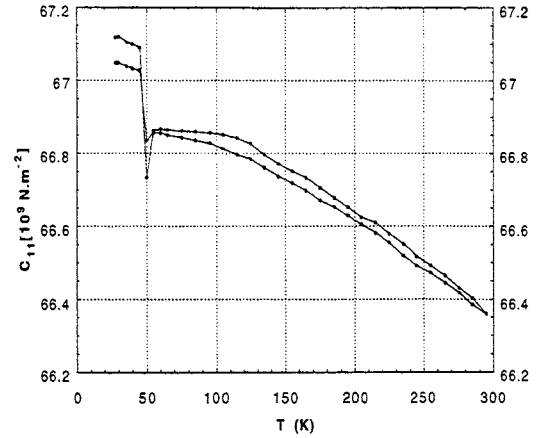


Figure 5 - Zero-field elastic constant C_{11} of $GaPO_4$ versus temperature (*). Thickness of the sample = 7.98 mm

Below this critical temperature, the independant elastic constants C_{11} and C_{33} denote a monotonical increment with decreasing temperature. The two other constants display inverse evolution. At lower temperature, in the 4-30K range, all the elastic constants register very weak variations, in accordance with the very strong stiffness of the network.

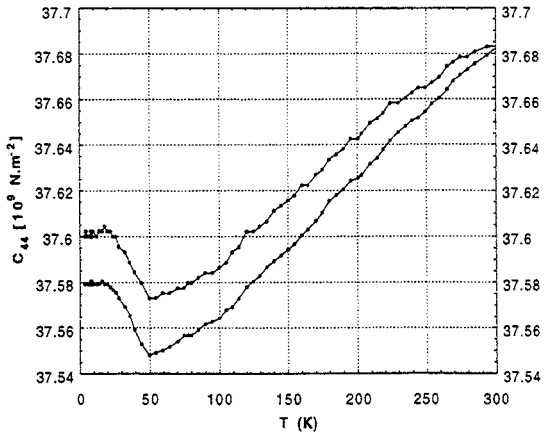


Figure 6 - Zero-field elastic constant C_{44} of $GaPO_4$ versus temperature (*). Thickness of the sample = 8.675 mm

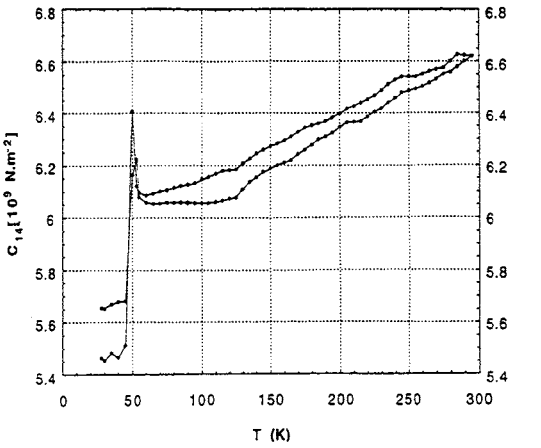


Figure 7 - Zero-field elastic constant C_{14} of $GaPO_4$ versus temperature (*).

Two main explanations can be advanced intending to explain both artefact at about 50 K and the disparity of behaviour existing between the Z-shear mode and the other modes investigated;

- a phase transition could occured, but this theory appears to be strongly criticizable according to the very weak character of the energies involved in this temperature range.
- a network relaxation, with a shortness of a certain kind of bonds and, in order to preserve a sufficient overall cohesion, a complementary lengthening of others. Such phenomenon might be put in evidence using, among other experiments, an acousto-optic subduing method.

Nevertheless, a real risk exists concerning the possibility that all of these unusual evidenced should be ascribable to some drifts of impurities, requiring very low energy contribution.

EXPERIMENTAL STUDY OF SAW PROPAGATION

I - Preparation of the samples

Fabricating SAW devices by optical contact photolithography process requires a good quality crystal with sufficent dimensions allowing to cut and polish flat plates with a minimum surface of about 1 cm^2. In this study, two crystal plates, grown using slow heating method [3,5] have been used, and SAW propagation occured in a seed-free zone, i.e. external to the middle of the crystal in the XY plane.

Two delay lines have been successively layed on the first plate, sample A (Z-surface 2 cm^2, thickness 4 mm). The second one, sample B, was cut from a bigger crystal (Z-surface 4 cm^2, thickness 1 mm) authorising the simultaneous deposition of two surface wave devices: in this ultimate case, either devices were composed with three transducers, with the purpose of compare both kinds of boundary conditions under a free or metallized surface.

In both cases, preamble the general rough rhomboedral shape of these plates, orientations of the delay lines were done optically by reference to the small diagonal parallel to the X-axis, whereas the long one according to the Y-axis.

On sample A, fig. 8, the first kind of delay line was first deposited at an angle $\Psi = 20°$, then, the crystal was cleaned, repolished and the same device was fabricated at an angle $\Psi = 30°$. On the other hand, the sample B was overlayed at the same time by two other delay lines, with propagation angles $\Psi = 10°$ and $\Psi = 30°$, fig.9 / Photo C.

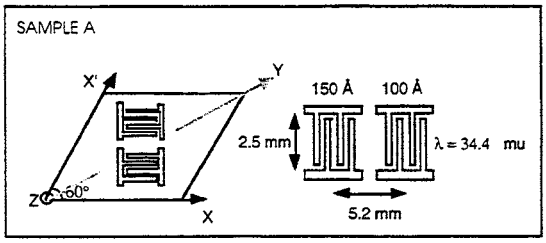
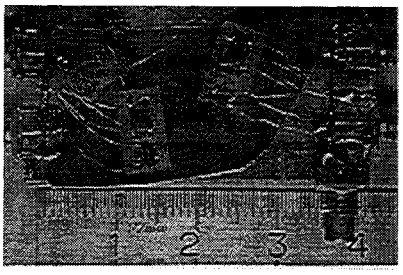


Figure 8 - SAMPLE A; 150 + 100 finger pairs, period 34.4 µm, aperture: 2.5 mm, center spacing: 5.2 mm



(PHOTO C)

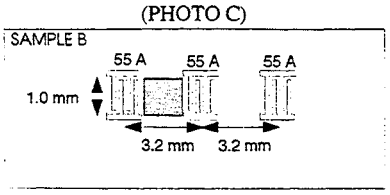


Figure 9 - SAMPLE B; 55 + 55 + 55 finger pairs, period 32 µm, aperture: 1.0 mm. center to center spacing: 3.2 mm

II - Discussion and comparison with theoretical data.

Measurements were performed with a HP 8752 A network analyzer. After calibration, the amplitude and phase of the electrical reponse of the delay lines were measured, figs. 12-13-14-15.

Electromechanical coupling factor ΔV/V behaviour for Z cut, showed in fig. 10, versus the propagation angle Ψ, denotes a maximum value of 4.10^{-4} (2,83 %): this latest, can reach five times as big value for singly rotated cuts [8]. Nevertheless, it is practically impossible to completely release the surface inconvenients. The influence of the value of the piezoelectric constant e_{14} value was checked by plotting the same curves with data from [9], but with e_{14} faced to zero.

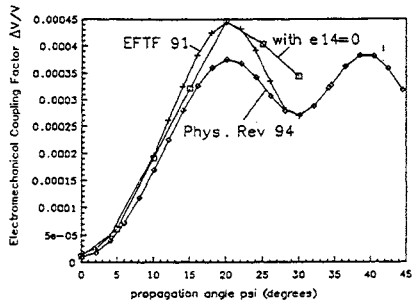


Figure 10 - Electromechanical coupling ΔV/V in the same configuration comparison between different data sets published [6,9]; influence of the value of e_{14} when closed to 0.

Velocities

If the electrical response of sample A was good enough for further experiments applications in temperature effects, very weak responses were gained for sample B.

Several explanations can be advanced like the low coupling factor which is certainly an aggravating point, but the most important one seems to be related to the nature itself of the transducer. Indeed, in order to use small flats, the number of the finger pairs of each transducer, 55, was chosen to minimize device size, but it appears that it should not be less than 100/150 for quantitative analyse on those cuts.

Furthermore, for Ψ = 30°, it is clearly evident that the classical behaviour of a sinx/x curve for Rayleigh waves is somewhat hidden by additional peaks extending beyond the expected synchronous frequency of the device. These peaks could be attributed to a coupling between Rayleigh and a shear bulk waves, the last one occured between both sides of the thin plate. Indeed, on the XY plane, the slowest shear velocities for bulk waves, i.e. 2544 m/s to the Y axis, according to material constants given in [9], correspond to the velocity range observed for Rayleigh waves, **fig.11-Table II**. So, either interactions and coupling should be taken in consideration for using this type of transducer along certain directions where velocities are very near.

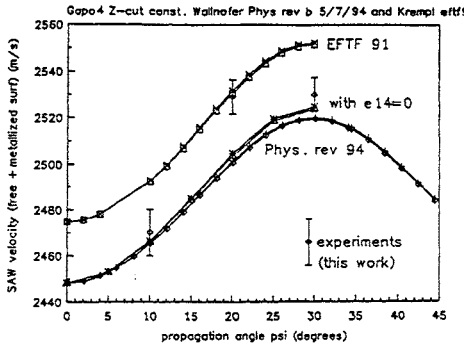


Figure 11 - Comparison of measured SAW velocities with theoretical predictions based on material constants given in [6,9]

Table II: Velocity measurements

ANGLE Ψ	SAMPLE	TRANSDUCER PERIOD	MEASURED VELOCITIES ($m.s^{-1}$)	THEORETICAL VALUE COMPUTED FROM [6] / [9]
10	B	92	2470 ± 10	2492 / 2465
20	A	34.4	2530 ± 7.5	2532 / 2500
30	A	34.4	2530 ± 7.5	2552 / 2520

However, for both plates, surface wave velocities had been computed from the central frequencies of the delay line response multiplied by the periodicity. For experimental data, the error-bars were taken, for sample A, from the - 3 dB bandwidth of the curve i.e. about 0.5 MHz, but for sample B, at Ψ = 10°, only a rough estimation was taken from the position of the strongest peak.

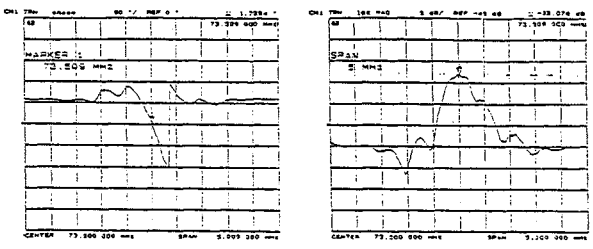


Figure 12 - Frequency response, sample A, delay line Ψ = 20°.

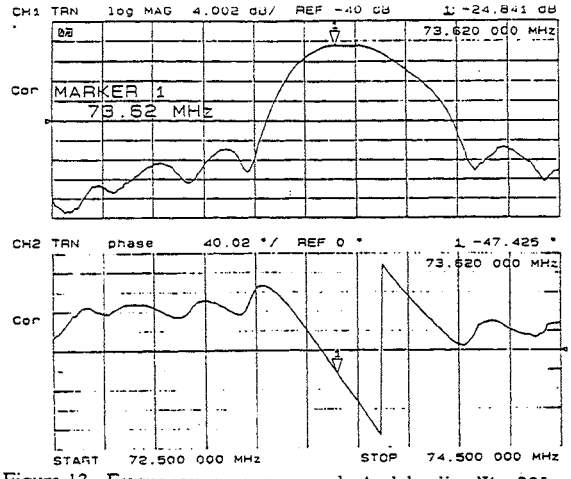
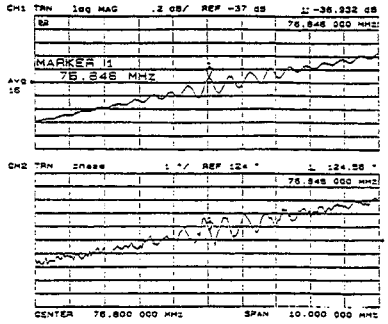
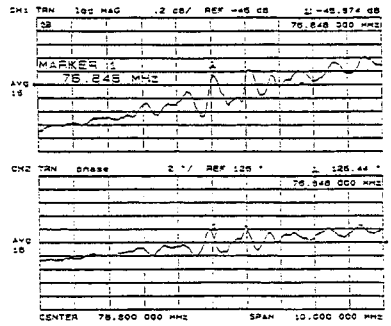


Figure 13 - Frequency response, sample A, delay line Ψ = 30°.

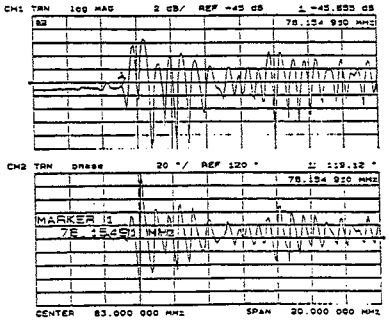


(a)

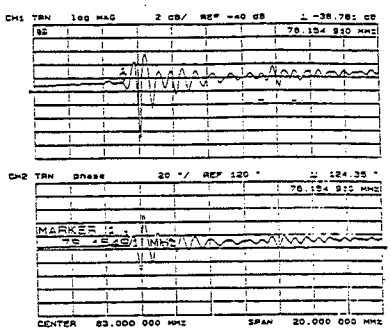


(b)

Figure 14 - Frequency response, sample B, delay line $\Psi = 10°$, metallized surface (a), free surface (b).



(a)



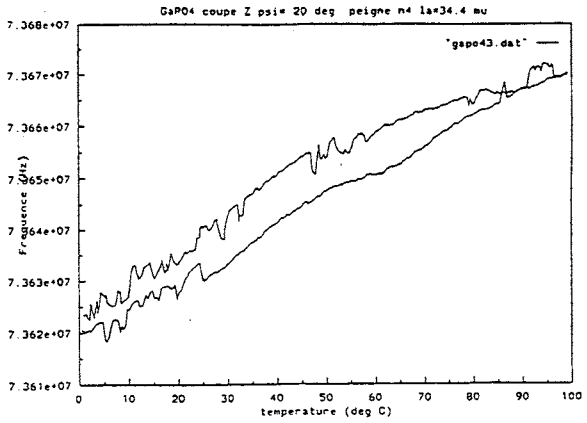
(b)

Figure 15 - Frequency response, sample B, delay line $\Psi = 30°$, metallized surface (a), free surface (b).

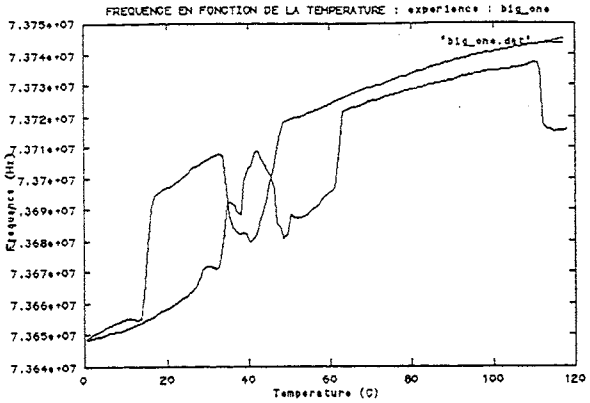
Temperature effects

The general behaviour of electrical response of sample A was good enough to allow oscillator measurements with a loop amplifier. The temperature dependance of the frequency is presented on figs. 16. Some frequency jumps at 40°C are observed for $\Psi = 30°$, but the overall behaviour of the device was still readable .

For both orientations $\Psi = 20°$ and $\Psi = 30°$, the first order temperature coefficient at room temperature had been estimated in the range of 8-10 ppm.K^{-1}.



(a)



(b)

Figure 16 - Temperature dependance of the frequency of an oscillator built with sample A, for $\Psi = 20°$ (a) and $\Psi = 30°$ (b)

In order to compare with theoretical values, the temperature coefficients of elastic constants and linear expansion coefficients given in [8] were used in a simulation program. The velocity of the surface wave is computed for several temperatures and the relative variations of velocities substracted from the relative variations of length given by the linear expansion coefficients given in [6].Theoretical results differ from the measured values. Table III.

Table III : Theoretical temperature coefficients (first order).

Angle Ψ	15	20	25	30
C.T.F. (ppm.K^{-1}) Theoretical	26.5	27.3	27.9	28.5
Experimental value		8-10		8-10

Clearly, there is a discrepancy here between theory and experiments, but there is little hope to find a Z-cut of $GaPO_4$ with a temperature compensation at room temperature. However the curves show a parabolic curvature, and an inversion point might exist above 100°C. Since GaPO4 can be used at higher temperature, this would be interesting to investigate, since large Z-cuts of $GaPO_4$ are easier to obtain than other cuts in the present method of crystal growth.

RELATION BETWEEN STRUCTURE AND PIEZOELECTRIC PROPERTIES OF QUARTZ-LIKE MATERIALS

For all the quartz-like materials, the θ limit for the α <---> β transition existence has been estimated at θ ≈ 136°, figure 18. For M-O-X angle values less than ≈ 136°, the needed energy to increase the angle up to ≈ 153° (M-O-X angle value for β-quartz) was found to be very high, and before this is reached, either another transition (i.e. cristobalite for $GaPO_4$) is observed or melting occurs. Similarly, physical properties such as coupling coefficient and AT cut angle can be directly related to structural distortions in terms of the bridging angle M-O-X. An increase in distortion intensifies the first characteristic and reduces the second one. The importance of these results allows the prediction of unknown physical properties for crystals such as $GaAsO_4$, $FePO_4$, $AlAsO_4$ structures of which have been determined.

The intrinsic properties and behaviour of all known crystals of this serie - quartz, berlinite and $GaPO_4$ - have been related to their structural distortions, and we have tried to extrapolate this important relationship between distortion and properties to other quartz-like materials which properties are not known, **fig. 18**. Structural distortions are here expressed by the M-O-X bridge angle value ([137.8°, 135.4° and 129.9° for $FePO_4$, $AlAsO_4$ and $GaAsO_4$ respectively [19]). If we extrapolate the linear variation of the coupling coefficient and the AT-cut angle values for the three known materials, $GaAsO_4$ would appear to be the most promising material (k ≈ 25 %, AT-cut ≈ 0°). This linear variation is confirmed by the extrapolation on the other side where the coupling coefficient converges to zero for θ ≈ 153°, Si-O-Si angle value for β quartz.

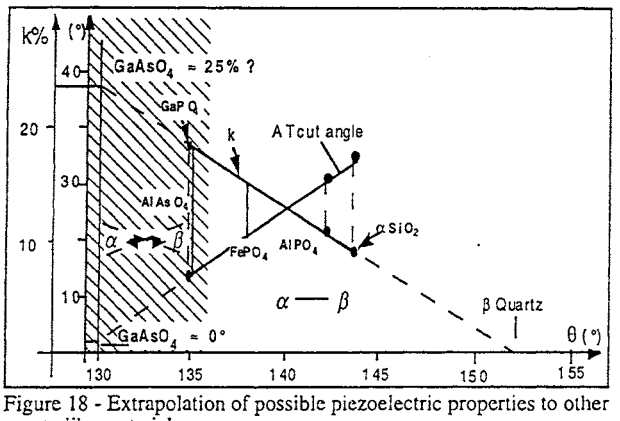


Figure 18 - Extrapolation of possible piezoelectric properties to other quartz-like materials.

CONCLUSION

From this investigation, it has been demonstrated the feasibility of $GaPO_4$ epitaxy on large berlinite seeds in sulphuric acid medium. From these large seeds, it seems that the most promising method, from an industrial point of view, is the $GaPO_4$ crystal growth in concentrated phosphoric acid medium. The OH content has been followed by infrared and near infrared spectroscopies. In comparison to berlinite, the OH content is approximatively lowered by a factor 10 for the same growth conditions.

For the forest time, the ultrasonic pulse superposition method has enabled the acquisition of a rough estimate of four out of the six single elastic constants of gallium phosphate, in the 4-300K temperature range. This first glimpse exhibits clearly a whole unexpected behaviour of the elastic constants C_{44} and C_{14}, which display drop values with increasing temperature. Such occurence is ascribable to the ultrasonic round-trip delay of the Z-shear mode, which discloses expended values with decreasing temperature.

Except for C_{33}, whatever was the propagation probed, the whole elastic constant have exhibited strong slope variation at 50 K. Unfortunately, a rigorous interpreting cannot be seriously suggested without X-ray measurements, now in progress, of the thermal expansion coefficients and an acuted analysis in the 50K zone.

The availability of Z-cut $GaPO_4$ plates with a surface greater than 1 cm^2 allowed the fabrication of SAW devices on seed-free samples. Measurements of velocities were performed and the agreement between experiments and predictions, based on experimental values of available material constants, is in the range of about 1 % which can be considered as satisfactory, considering that crystal properties are still not fully stabilized. Oscillators were built and temperature coefficients were measured, indicating that if no temperature compensation exists at room temperature on Z-cuts, higher temperature could yield a reduced thermal sensitivity.

Some of piezoelectric factors have been closely related to structural considerations and have allowed to extrapolate the physical properties of other quartz-like materials.

Acknowledgements : The authors acknowledge the DRET and the CNRS for their financial support.

References

[1] - G. Engel, P. Krempl and J. Stadler, in: *Proc. 3rd European Time and Frequency Forum, 1989, p. 50.*
[2] - G. Engel and P. Krempl, in: *Proc. 2nd European Time and Frequency Forum, 1988, p. 827.*
[3] - J.C. Jumas, A. Goiffon, B. Capelle, A. Zarka, J.C. Doukhan, J. Schwartzel, J. Détaint and E. Philippot, *J. Cryst. Growth 80 (1987) 133.*
[4] - E. Philippot, A. Ibanez, A. Goiffon, M. Cochez, A. Zarka, B. Capelle, J. Schwartzel and J. Détaint, *J. Cryst. Growth 130 (1993) 195.*
[5] - E. Philippot, A. Goiffon, M. Maurin, J. Détaint, J. Schwartzel, Y. Toudic, B. Capelle and A. Zarka, *J. Cryst. Growth 104 (1990) 713.*
[6] - P. Krempl, J. Stadler, W. Waljnöfer, W. Ellmeyer, R. Selic, in *Proc. 5th European Frequency and Time Forum, 1991, p. 143.*
[7] - E. Philippot, A. Ibanez, A. Goiffon, B. Capelle, A. Zarka, J. Schwartzel, J. Détaint, in *Proc. 6th European Frequency and Time Forum, 1992, p.383.*
[8] - W. Wallnöfer, J. Stadler, P. Krempl, in *Proc. 7th European Frequency and Time Forum, 1993, p. 653.*
[9] - W. Wallnöfer, P.Krempl, A. Asembraum, in *Phys. Rev. B, vol. 49, PP.10075-10080(1994)*
[10] - D. Cachau-Hereillat, J. Bennazha, A. Goiffon, A. Ibanez and E. Philippot, in *European J. Solid State Inorg. Chem. 29 (1992) 1295.*
[11] - O. Cambon, A. Goiffon, A. Ibanez and E. Philippot, in *European J. Solid State Inorg. Chem. 29 (1992) 547.*
[12] - A. Zarka, J. Détaint, J. Schwartzel, Y. Toudic, B. Capelle, Y.L. Zheng, E. Philippot, X. Buisson and R. Arnaud, in *French Patent 91 14007 (1991).*
[13] - A. Zarka, L. Liu and X. Buisson, in *J. Cryst. Growth 54 (1981) 59.*
[14] - R.F. Steinberg, M.K. Roy, A.K. Estes, B.H.T. Chai and R.C. Morris, in *Proc. IEEE Ultrasonics Symp., 1984, p. 279.*
[15] - J.H. Mc Fee, in *Physical Acoustics (W.P. Mason and R.N. Thurston Eds) Vol.4 A, Chap 1, New York: Academic Press (1966)*
[16] - P. Krempl, in *IEEE Ultrasonics Symposium, 1994.*
[17] - J. Detaint, J. Schwartzel, A. Zarka, B. Capelle, J. P. Denis, E. Philippot, in *Proc. 48th IEEE Frequency Control Symposium, 1994.*
[18] - E. Philippot, A. Goiffon, A. Ibanez and M. Pintard, in *J. Solid State Chem. 110 (1994) 356.*
[19] - A. Goiffon, J.C. Jumas, M. Maurin and E. Philippot, in *J. Solid State Chem., 61 (1986) 384.*

$GaPO_4$: A CRITICAL REVIEW OF MATERIAL DATA

P. W. KREMPL, F. KRISPEL, W. WALLNÖFER and G. LEUPRECHT

AVL List GmbH, Kleiststraße 48, A-8020 Graz, Austria

ABSTRACT

Due to recent progress in the growth process, mainly by epitaxial growth on quartz seeds, the quartz-homeotypic crystal $GaPO_4$ is now ready to be used in industrial products. We thus review its most important physical properties for the use in sensor and resonator applications. The piezoelectric constants were newly determined by an interferometric method. Also, new measurements of the thermal expansion coefficients of high quality material are presented. These results and their influence on the acoustic properties are compared with older values and experimental data, both from our group and from other laboratories. The suitability of $GaPO_4$ for different types of devices, specially in the field of sensorics, is discussed and compared with other piezoelectric materials. The first $GaPO_4$ application, a pressure sensor for combustion engines, is already on the market.

1. INTRODUCTION

$GaPO_4$ is one of the most interesting new piezoelectric crystal materials[1, 2, 3, 4]. Recent progress in its growth is leading to crystals of considerable size. So it seems time to give a critical review of its material constants, previously published.

Gallium orthophosphate may be regarded as the "high tech brother of quartz". The crystal belongs to the same point group as quartz (32) and has the same structure with left- or righthanded helices where silicon atoms are alternately substituted by gallium and phosphorus. The density (ρ = 3570 kg/m^3) is somewhat higher than in quartz. An α-quartz/β-quartz transition is not observed up to 930°C where a transition to a cristobalite-like phase takes place. The material has high thermal stability and only small changes in physical properties up to that temperature. The strong longitudinal piezocoefficient, twice that of quartz, and the absence of pyroelectricity make the new crystal to a very promising material for high temperature sensors (up to 900°C). The strong electromechanical coupling makes the crystal very useful in SAW and filter applications.

Melt growth is not possible; thus $GaPO_4$ is grown hydrothermally. Since $GaPO_4$ does not occur in nature, seed plates first were obtained from spontaneous crystals. But they are small and the achievable size is limited by slowly growing faces. So the repetitive enlargement of the crystals is to slow and not efficient. This problem now could be solved by a new process. The first step is to produce a thin epitaxial layer of $GaPO_4$ on quartz plates used as seeds. This is possible by admixture of fluoride to the growth solution. With this method, large crystals with good quality could be obtained.

Since thus the availability of $GaPO_4$ material of sufficient size and quality for industrial applications can be anticipated in the near future, its relevant properties are of increasing interest. We present here results of new measurements done by us and other groups and the resulting material data of $GaPO_4$.

2. PIEZOELECTRIC CONSTANTS

The piezoelectric effect is described phenomenologically with the equation

$$P_i = d_{i\mu}\, T_\mu$$

with P_i...electrical polarization ($i = 1..3$),
T_μ...mechanical stress ($\mu = 1..6$),
$d_{i\mu}$...piezoelectric tensor, which for $GaPO_4$ has the form

$$d = \begin{pmatrix} d_{11} & -d_{11} & 0 & d_{14} & 0 & 0 \\ 0 & 0 & 0 & 0 & -d_{14} & -2d_{11} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

whereby temperature T and mechanical strain S_{mn} are constant. On the other side the reverse piezoelectric effect is given by

$$S_\mu = d_{\mu k}\, E_k$$

with S_μ...mechanical strain and
E_k...electric field.

Thus the piezoelectric constants can be determined by two distinct measurement methods. Often only the first method is used to get the constants under stress with a force calibration device.

We developed a new measurement system to determine the piezoelectric constants by the inverse effect. I.e. the crystal is influenced by an electric field and the change in length is measured with a Michelson-interferometric arrangement (Fig. 1)[5].

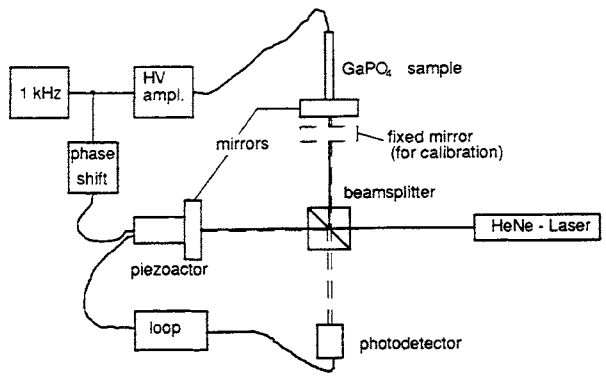


Fig. 1: Experimental arrangement for interferometric measurement of the piezoelectric constants.

The light source is a He-Ne laser at 632.8 nm. The two partial beams coming out of the beam splitter are reflected by two mirrors and interfere at the photodetector, a photo-element. The light signal in the photodetector depends on the difference of the distances between the mirrors and the beamsplitter. One mirror is coupled to a piezoactor (PA) with a high piezoconstant (PZT - ceramics, d_{PA} = 87.2 nm/V) which was calibrated by applying a voltage which changes the length by a quarter of the laser wave length. This was not possible with the $GaPO_4$ sample, since the necessary high electric field would lead to breakdown in the air near the electrodes; thus a compensating method was used. A 1 kHz signal is applied to both PA and sample; the ratio of the amplitudes and the phase is adjusted so that the 1 kHz signal in the photodetector vanishes; then both mirrors move in phase and with the same amplitude. This ratio, the PA constant and geometric factors allow determination of the relevant piezoelectric constant. After checking the apparatus with a quartz sample, measurements on gallium orthophosphate were performed. An oven was used to measure at higher temperatures.

The sample for the d_{11} measurement is a unrotated x-cut with the field along x and the elongation in y direction. To measure d_{14}, it is necessary to use a rotated crystal so that the shear deformation is transformed to an elongation. We used a y-cut rotated by 45° around the y axis.

The results are present in Fig. 2. The room temperature values for d_{11} agree for the force calibration measurements[6], theoretical calculations based on bond additivity theory[7], using atomic coordinates from Ref. 8 and 9, and the new interferometric measurements. Values for d_{14} agree very well for the experimental methods. For both constants the values measured with the interferometer decrease less with temperature than the force calibration results.

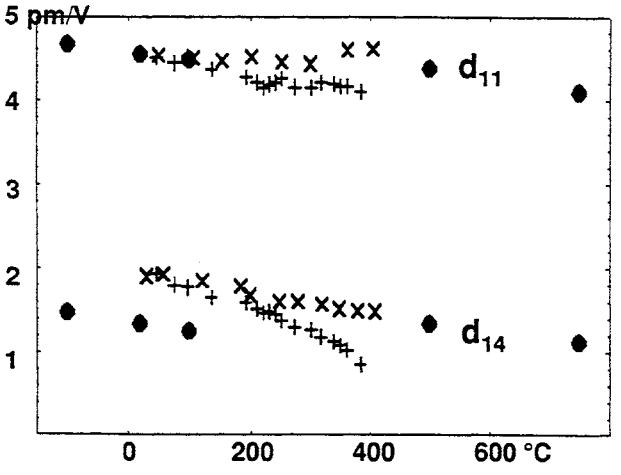


Fig. 2: Piezoelectric constants of $GaPO_4$: • : prediction by the bond additivity theory[7]; x : interferometric measurements, this work; + : force calibrator measurements[6].

A linear fit of the temperature dependences yields the values in Table 1. However, these results should be regarded as preliminary; further measurements over a wider temperature range are in progress.

Table 1: Preliminary values of the piezoelectric constants of $GaPO_4$ at room temperature and their first temperature coefficients.

	d@25°C [pm/V]	1. TC. [ppm/K]
d_{11}	4.53	44
d_{14}	1.96	-622

3. OH CONTENT

Many physical properties, e.g. the dielectric constants, depend markedly on the OH content of the material. We have to take notice of this factor although it is not an intrinsic property of $GaPO_4$. For quartz and other materials, the estimation of the OH content is used to give an idea of quality and impurities in the crystal. It is determined from the infrared absorption in the O-H stretching vibration range, which is corrected on thickness and background losses by the formula

$$\alpha = \frac{1}{d} \log\left[\frac{T_{3800}}{T_{WN[OH]}}\right] - \alpha^*$$

with α....absorbance [cm^{-1}]
d.....thickness [cm]
T.....transmission:
3800 cm^{-1}... background (no OH absorption)
WN[OH] ... wavenumber for OH measurement
α*....absorbance which is not due to O-H vibrations.

We have measured more than fifty $GaPO_4$ samples with widely varying OH contents and find that the two absorbance peaks at 3167 cm^{-1} and 3290 cm^{-1}, when subtracting the broad OH band background, are independent on OH value and only very weakly dependent on polarization (Fig. 3). Their absorbances are:

$$\alpha_{3167} = 0.689\ cm^{-1} \pm 1.2\ \%; \quad \alpha_{3290} = 0.278\ cm^{-1} \pm 2.7\ \%.$$

This finding contradicts the interpretation of these narrow peaks as due to specific O-H defects[2]. Further support is given from the IR spectrum of a very thin $GaPO_4$ sample (d = 0.00325 cm) which shows strong lattice bands at 2130 cm^{-1} and 2240 cm^{-1}, i.e. nearly 2/3 of the mentioned peaks. So we attribute these peaks to third order vibrations of the undisturbed $GaPO_4$ lattice.

We suggest the use of WN[OH] = 3400 cm^{-1} because there the OH-band absorbance is near its maximal value and the lattice absorption is very low (α^* = 0.04 cm^{-1}) and usually can be neglected.

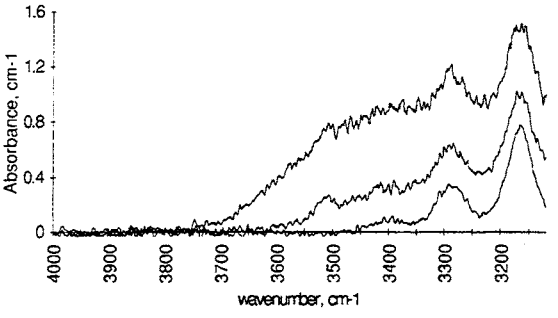


Fig. 3: Infrared absorbance spectra of 3 $GaPO_4$ samples with different OH content.

To estimate the OH content quantitatively, we suggest the expression:

$$G_W(ppm/wt) = 65\ \alpha$$

The factor 65 is derived from water in fused quartz[10].

4. DIELECTRIC CONSTANTS

Recently, a thorough study of the dielectric properties of $GaPO_4$ was published by Foulon et al.[11]. The capacitances of x and z plates were used to measure the two independent dielectric constants ε^T_{11}' and ε^T_{33}' and the dielectric losses ε_{11}'' and ε_{33}'' with varying temperature, frequency and impurity concentration (OH content), determined by infrared absorption measurements.

At elevated temperatures (above room temperature) and lower frequencies, most sample show a significant rise in all constants with frequency and temperature. Especially the dielectric loss, which is neglectable below room temperature, becomes considerable. This is due to impurities indicated by the OH-content. For the sample with the lowest OH content, there is only a low rise of the dielectric constants while the dielectric loss remains neglectable (Fig. 4).

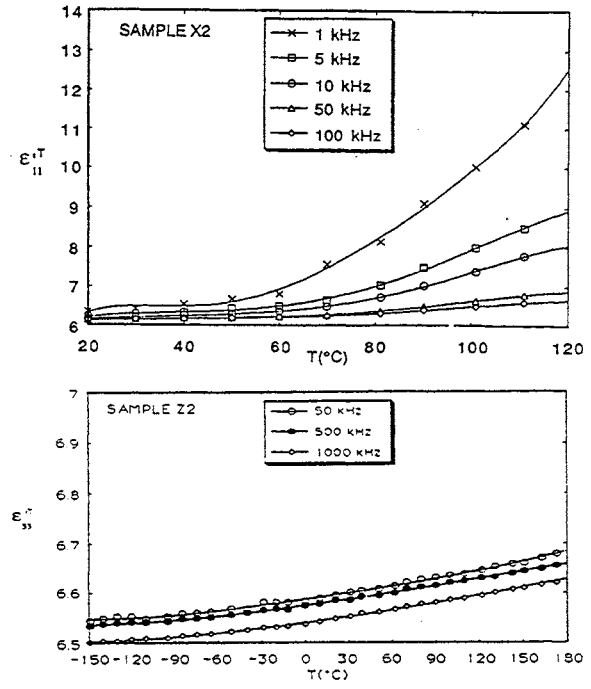


Fig. 4: Measured dielectric constants of $GaPO_4$, from Ref. 11; above: high OH content; below: low OH content.

We have extrapolated the experimental data linearly to [OH] = 0, taking into account the α values for the used lines, to get the constants of the pure $GaPO_4$. The results, displayed in Table 2, show much less temperature dependence than most of the measured samples.

Table 2: Dielectric constants of $GaPO_4$, extrapolated to OH-free material.

T	-100 °C	20 °C	100 °C
ε^T_{11}'	6.04	6.07	6.04
ε^S_{11}'	5.80	5.83	5.80
$\varepsilon^T_{33}' = \varepsilon^S_{33}'$	6.54	6.58	6.65

The superscript T indicates dielectric constants measured at constant stress, while S indicates calculated values for constant strain.

5. THERMAL EXPANSION

The thermal expansion is essential for the calculation of the temperature behaviour of acoustic devices, since it changes the density, the thickness and the effective wavelength. Thus, reliable values are needed to predict BAW and SAW properties.

The most common method for crystals is the diffractometric measurement of the lattice parameters at various temperatures. There exist 3 publications where such data are reported for $GaPO_4$: by Kosten & Arnold[9], by Goiffon et al.[8] and by Baumgartner et al.[12]. The thermal expansion values, besides being only partly comparable due to different ranges and partly large spacings, show clear discrepancies. Thus the need was seen for a new measurement in the technically interesting range.

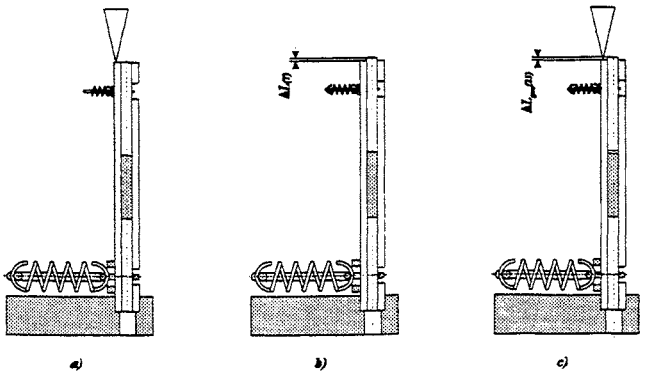


Fig. 5: Thermal expansion measurement: a) first profile measurement at RT, b) heated up in oven, c) second profile measurement at RT.

The measuring principle which was chosen is depicted in Fig. 5. In a fused quartz tube, a $GaPO_4$ sample rod is enclosed between two cylindrical fused quartz rods. The lower one is kept fixed to the tube by a strong spring, while the upper one, the indicator rod, is fixed by a very weak spring. The upper planes of indicator rod and tube are aligned and measured by a laser profile measuring device. The apparatus is then heated in an oven to the measurement temperature. Since the $GaPO_4$ sample expands more than fused quartz, it pushes the indicator rod up to a position where it is fixed during the cooling-down by the weak spring. With a second profile measurement at room temperature, the step height between the fronts of tube and indicator rod is determined. The shift, corrected with the expansion of fused quartz, gives the expansion of the $GaPO_4$ rod. The process is repeated for each temperature. A 20.67 mm x rod and a 13.99 mm z rod with OH content below 100 ppm were used. Temperatures ranged from room temperature to 400°C.

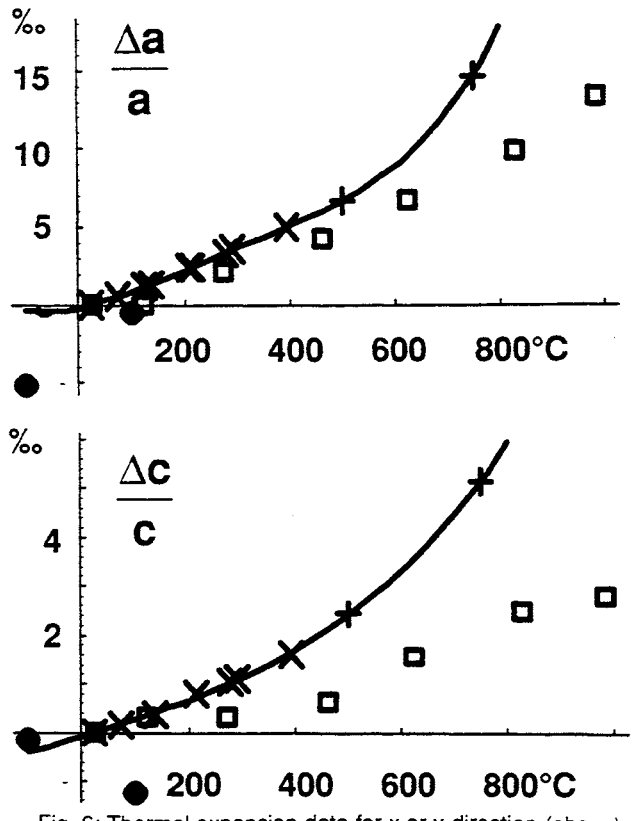


Fig. 6: Thermal expansion data for x or y direction (above) and z direction (below); x : this work; + : Ref. 12; • : Ref. 8; □: Ref. 9.

The results of these and the diffraction measurements, expressed in expansion relative to room temperature, are shown graphically in Fig. 6. The new data combine well with the two high temperature measurements by Baumgartner et al., which also supplement nicely the temperature range. Thus a polynomial least squares fit was performed to the combination of both data, delivering the following thermal expansion coefficients relative to 25°C:

	1. TC, 10^{-6} K^{-1}	2. TC, 10^{-9} K^{-2}	3. TC, 10^{-12} K^{-3}	4. TC, 10^{-15} K^{-4}
a (x or y axis)	9.02	35.4	-99.1	98.8
c (z axis)	3.38	2.0	1.0	4.5

It is recommended to use these values for the calculation of $GaPO_4$ properties, as we did in the following work.

6. THERMOELASTIC PROPERTIES

Among the most interesting possible applications of $GaPO_4$ are bulk and surface acoustic wave devices. Like quartz and berlinite, $GaPO_4$ has temperature compensated cuts for both types of waves[13]. The group at CNET[14] showed the existence of a temperature-compensated rotated Y cut with AT-like behaviour near ϑ= -13°. Its electroacoustic coupling constant k is 16%, which is almost twice the value of the quartz AT cut (k = 8.8%). The temperature behaviour of the frequency of 3 orientations near AT is shown in Fig. 7.

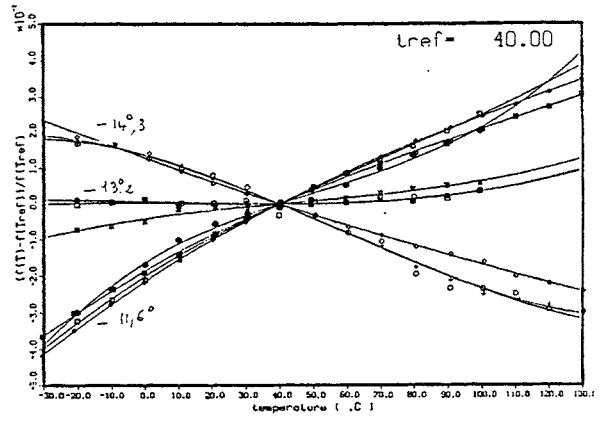


Fig. 7: Temperature dependence of frequency around 40°C of the $GaPO_4$ cuts Y-11.6°, Y-13.2° and Y-14.3° [14,15].

The only published complete set of elastic constants[16] and their temperature coefficients[13] are those derived from Brillouin measurements. While they delivered reliable room temperature values, as experimentally verified [17,4] and show the behaviour over wide temperature ranges, the details of the behaviour of the transverse waves near room temperature is given more accurately by the resonator measurements on high quality material. Thus, we used the best available resonator data and the new expansion coefficients to update the set of temperature coefficients of the elastic constants.

For the determination of the coefficients of c_{44}, c_{66} and c_{14}, besides the data by Detaint et al. shown in Fig. 7, results of 2 high quality resonators at Y and Y+53° [18] were used. c_{11}, c_{33} and c_{13} were evaluated from Brillouin data near room temperature, using the new expansion data in the evaluation. From this procedure, the following values for the first and second temperature coefficients of the elastic constants were obtained.

Table 3: First and second temperature coefficients of the elastic stiffness constants of $GaPO_4$.

	c_{11}	c_{13}	c_{14}	c_{33}	c_{44}	c_{66}
1. TC, 10^{-6} K^{-1}	-63	-180	825	-135	-62	69
2. TC, 10^{-9} K^{-2}	-88	-105	2000	-84	320	-207

7. SAW PROPERTIES

With the new material data, the SAW properties can be calculated. Table 4 compares our results on the temperature dependence of frequency with the first, and to our knowledge unique measurement published for $GaPO_4$ SAW on z cuts[17]. The comparison shows a fair agreement.

Table 4. Comparison of measured[17] and calculated mean temperature coefficients for SAWs in $GaPO_4$, in 10^{-6} K^{-1}.

Orientation	Temp. range	exp.	calc.
z cut, 20°	0°C..100°C	7	9.5
z cut, 30°	0°C..120°C	8	15

The predicted orientation for the temperature compensated ST cut is shifted to the Y-cut relative to earlier calculations[13]. The new values are: μ = 69.5°, v = 2312 m/s and k^2 = 0.35%. This would mean a even higher coupling than anticipated earlier.

8. APPLICATIONS

The material data presented here confirm the usefulness of $GaPO_4$ for a wide range of technical applications, which can be divided in the following 3 groups:

- charge-output sensors: examples are pressure or force sensors or accelerometers using the direct piezoelectric effect. The main advantages of $GaPO_4$ are the high sensitivity and its low temperature dependence up to high temperatures, shown by the d_{11} measurements, and the absence of stress-induced twinning. Utility up to 930°C is probable.
- frequency-output sensors: possible applications comprise microbalances, vibrating beam force sensors or accelerometers and various sensors for temperature, pressure or chemical quantities based on bulk or surface acoustic waves. Essential here is the high temperature stability of frequency, the possibility of pre-aging at elevated temperatures and the large useful temperature range.
- other acoustoelectric devices: these include BAW resonator filters, SAW filters and oscillators. For this field, $GaPO_4$ shows high coupling, thus high possible bandwidth, and temperature-compensated cuts with high stability.

Of all these potential applications, the first one is already on the market for over one year. It is the pressure transducer AVL-GM11D for combustion engines. The possibility to omit water cooling allows a much smaller size with an M5 thread. The sensor uses the longitudinal direct piezoelectrical effect, which is proportional to the component d_{11}. The excellent stability of this constant allows the specification that the sensitivity remains within ±2% of the room temperature value for the whole range up to 400°C.

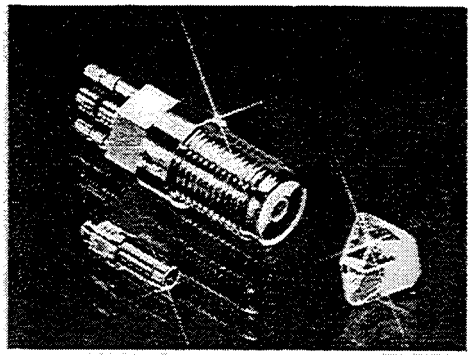


Fig. 8: $GaPO_4$ pressure sensor AVL-GM11D (in front), compared with a water-cooled quartz pressure sensor.

9. CONCLUSION

Due to epitaxial growth on quartz seeds, the availability of large $GaPO_4$ crystals can be anticipated in the near future. Thus, its physical material properties are important to assess its usefulness for sensoric and acoustoelectric applications. Preliminary results from interferometric measurements of the piezoelectric constants of $GaPO_4$ are presented. The thermal expansion between room temperature and 400°C was measured. Dielectric measurements by Foulon et al. were corrected for the OH content whose determination by IR absorption is also discussed. The thermoelastic constants are updated using resonator measurements by Détaint et al. With these new material values, more accurate calculations of the properties of devices are possible, thus improving the possibility to determine the suitability for various applications.

REFERENCES

1. P.W. Krempl, J. Stadler, W. Wallnöfer, W. Ellmeyer and R. Selic, Proc. Fifth European Frequency and Time Forum, Besancon 1991, p.143-147
2. E. Philippot, A. Ibanez, A. Goiffon, M. Cochez, A. Zarka, B. Capelle, J. Schwartzel and J. Détaint, J. Crystal Growth 130 (1993) 195-208
3. P.W. Krempl, to be published in Proc. 1994 IEEE Ultrasonics Symposium, Cannes, 1.-4. 11. 1994
4. D. Palmier, M. Cochez, R. Gohier, C. Bonjour, G. Marianneau, A. Zarembovich, E. Bigler and E. Philippot, these proceedings.
5. W. Belitsch, Diploma thesis, University Graz (1993)
6. R. Selic, AVL internal report (1991)
7. G.F. Engel and P.W. Krempl, Proceedings of the Second European Frequency and Time Forum, Neuchatel, 1988, pp.827-832
8. A. Goiffon, J.C. Jumas, M. Maurin and E. Philippot, J. Solid State Chem. 61 (1986) 384-386
9. K. Kosten and A. Arnold, Z. Kristallogr. 152 (1980) 119-133
10. A. Kats, Philips Res. Rep. 17 (1962) 133 and 201
11. J.-D. Foulon, J.-C. Giuntini and E. Philippot, Eur. J. Solid State Inorg. Chem. 31 (3) 245 (1994)
12. O. Baumgartner, A. Preisinger, P.W. Krempl and H. Mang (1984), Z. Kristallogr. 168, 83-91
13. W. Wallnöfer, J. Stadler, and P.W. Krempl, Proc. 7th European Frequency and Time Forum, Neuchatel 1993, p. 653
14. J. Schwartzel, J. Détaint, B. Capelle, A. Zarka, E. Philippot, A. Ibanez, J.P. Denis and D. Cochet-Muchy, Proceedings of the 8h European Frequency and Time Forum, München, 1994, p.231
15. J. Détaint, personal communication
16. W. Wallnöfer, P.W. Krempl and A. Asenbaum, Phys. Rev. B 49 (1994) 10075
17. E. Bigler, S. Ballandras, C. Bonjour, D. Palmier and E. Philippot, to be published in Proc. 1994 IEEE Ultrasonics Symposium, Cannes, 1.-4. 11. 1994
18. J. Stadler, AVL internal report (1992)

SESSION 4

TIME TRANSFER

Chairman : G. PETIT

IMPROVED TIME AND FREQUENCY TRANSFER OPPORTUNITIES VIA GPS

David W. Allan
Allan's TIME, Inc.

Robin Giffard, Leonard S. Cutler and Jack Kusters
Hewlett Packard Company

ABSTRACT

The number of users of GPS for time and frequency is growing at a rapid pace. These users cover a diversity of applications. Some need time stability; others need time accuracy. Some need frequency stability; others need frequency accuracy. Some are willing to make measurements and then calculate after-the-fact the relevant time and/or frequency information of interest. Others need their data in real-time. Some use the time and frequency information directly available from GPS. Others use the GPS signals as a transfer standard to communicate time and/or frequency between clocks remote to each other.

The most demanding time and frequency metrology needs come from those wishing to compare the new primary frequency standards now being developed, which have potential accuracies of about 1 x 10^{-15}. Better accuracies than these are anticipated, and future numbers for NASA's Deep Space Network (DSN) are at the 10^{-17} level. On the other hand, the telecommunication and power industries are specifying large numbers of inexpensive, very stable, time and frequency sources for network monitoring and control of their systems in real time.

Six significantly different methods of using GPS for time and frequency have been developed, and are now at different levels of maturity. With new technologies arising, new opportunities of integrating these techniques have also arisen. These sundry needs, and relevant aspects of these methods will be reviewed in the paper -- showing that some very promising opportunities are imminent for those who choose to take advantage of them.

INTRODUCTION

The civil use of GPS for time and frequency is growing rapidly, and three novel time and frequency transfer techniques have been introduced recently. [1-5] International timing stabilities at the sub-nanosecond level have been documented, and initial studies indicate that frequency transfer uncertainty approaching 1 x 10^{-16} may be reasonably achievable. [1-2]

The large cross-section of civil users of GPS timing include: the Bureau International des Poids et Mesures (BIPM), which has the responsibility for the generation of International Atomic Time (TAI) and UTC (Coordinated Universal Time); international timing centers of which there are about 46 through out the world; NASA JPL's Deep Space Network (DSN); the telecommunication and power industries; the world-wide measurement program of the spin rates of the newly discovered millisecond pulsars; the National Institute of Standards and Technology's (NIST) global time service; NASA's timing of space platforms, and numerous other industrial and calibration laboratories.

Methods of utilizing the GPS signals, and accuracy goals vary across this user set. The variety of needs and the opportunities available using GPS provide a synergistic coupling. Cross fertilization between disciplines is often fruitful.

There are at least six different methods of using GPS for time and frequency transfer. These are: GPS direct; enhanced GPS; GPS common-view; GPS used in the clock fly-over mode; GPS used as in VLBI (Very-long Baseline Interferometry); and GPS used as in geodesy. The timing accuracies and precisions vary from sub-ns to hundreds of ns across these methods. Some of the factors that determine which method is to be used are whether it provides time and frequency information in real-time or whether computations must be performed after the fact. Complexities of these methods vary considerably. In general, the simpler the better as long as the accuracy and stability goals can be met.

The enhanced GPS approach is a method of filtering the SA (Selective Availability) GPS -- degraded signals -- for time and frequency purposes. Whereas SA limits the navigation and position accuracy to a real-time accuracy of slightly better than the 100 meters, the EGPS timing technique can provide a real-time UTC(USNO MC) output approaching the accuracy and precision of the source. The technique has been tested at USNO and the rms agreement over one month's data was 1.5 nanoseconds (ns). [4]

Recently, three novel techniques have been developed: the first uses a variation of the enhanced GPS mode; the second uses a multi-channel receiver in the common-view mode; and the last uses the GPS carrier phase as in geodesy. [1, 2, 5] Fortunately, the first two techniques can be combined using one receiver. In other words, it can serve both functions at the same time -- provide a real-time estimate of UTC and provide the data for high accuracy common-view after the fact. The third of the three novel techniques has shown time stabilities of 0.6 ns over baselines as large as 4 Mm (the circumference of the earth is about 40 Mm). Given its mode of operation, it will be best suited for high accuracy frequency comparisons. Current studies indicate that the frequency measurement uncertainty should be significantly better than 1 x 10^{-15}.

Since UTC(USNO MC) is synchronized and syntonized to UTC, GPS provides an indirect source of UTC. Thus, GPS provides a world wide resource for time and frequency from UTC at heretofore unprecedented accuracies and precisions. The most cost effective method of approaching the full accuracy of UTC to date is in using the EGPS mode. The accuracy now with respect to UTC is typically less than 100 ns. A method is outlined in this paper to move this below the 10 ns level.

From a timing point of view, with the 24 satellite constellation there is a high level of redundancy, reliability, and robustness available using GPS. The advantages of this full constellation are just beginning to be appreciated. Very low cost and very stable multi-channel GPS timing receivers are now available. As supporting systems software and hardware are developed, the time-and-frequency capability available from GPS will blossom to its full potential -- 10 ns accuracy, less than 1 ns day-to-day stability and 1x10^{-16} uncertainty for comparison of frequency standards remote to each other.

It is the purpose of this paper to highlight and focus on some emerging technologies as well as some significant opportunities given the needs within the precise time and frequency user

community. As an example, the primary world reference for time and frequency is TAI and UTC. Outlined herein is a way to significantly improve the stability and accuracy of UTC and, hence, greatly advance international time and frequency comparison accuracies.

SOME IMPORTANT TIME AND FREQUENCY CONCEPTS AND TOOLS

Whereas an ideal measurement system will be limited by white-noise phase modulation (PM), the ideal cesium-beam clock is limited by white-noise frequency modulation (FM). Often overlooked are degrees of freedom given such an ideal situation. Intuitively stated, the more information obtained in a measurement, the more is known.

Specifically, the frequency difference between two frequency standards can be estimated from the slope of a linear regression to a series of time differences taken between the standards. If the residuals to the regression line have a white spectrum, then the uncertainty of the estimate is given by:

$$\sigma_{\bar{y}} = \frac{\sqrt{12}\ \sigma}{\tau_o\ n^{3/2}} \qquad (1)$$

where n is the degrees of freedom, σ is the standard deviation and τ_0 is the data spacing. Given this ideal situation, the 3/2 power in the denominator is a strong factor in reducing the uncertainty on the frequency estimate.

As a case in point, two state-of-the-art systems for comparing clocks remotely located from each other are TWSTFT (Two-way Satellite Time and Frequency Transfer) and GPS common-view (C-V). The typical σ for TWSTFT is much smaller than for GPS C-V, but because GPS is continuously available in contrast to three measurements per week currently taken for TWSTFT, the $n^{-3/2}$ factor could potentially overcome the difference in the sigmas. In other words, properly utilized GPS over a continual data stream may supply enough degrees of freedom that the frequency uncertainty in comparing remote standards would be better than TWSTFT technique used three times per week. Currently, the biggest metrology problem in both techniques is the long-term systematic difficulties which cause time measurement dispersions much larger than given by the simple white-noise PM model.

As these systems continue to mature, these systematics will be driven down. If continuous measurements were taken using TWSTFT, it clearly would be the better technique for comparison of the frequency of clocks remote to each other. However, that is not the current modus operandi. Later in the paper, given the above ideas, we will present some alternatives using GPS that could make it superior to TWSTFT for frequency comparisons.

An alternative representation of the above uncertainty is given by the following:

$$\sigma_{\bar{y}} = 2\ Mod\sigma_y(\tau) \qquad (2)$$

Even if the measurement system is not limited by white-noise PM as a model, equation (2) still yields an approximate estimate of the frequency-measurement uncertainty.

Some of the desirable aspects of a measurement system are its convenience in obtaining data and in data processing. An ideal system would be one that could be accessed anywhere in the world providing all the time and frequency accuracy and stability needed for measuring a given standard. This being impractical, one could back up a step and provide instead a very stable global frequency reference. GPS can currently provide this role to an accuracy of better than 10^{-13} and a stability of the order of 10^{-14}. This has provided a very convenient tool.

Later in the paper we will discuss some experiments which indicate an opportunity to improve this accuracy and stability by as much as an order of magnitude. Given the above considerations, we will also discuss how one may achieve better than 10^{-15} comparisons of remote clocks.

An illustration of some of the above concepts was shown once before with an orbiting hydrogen maser. In this experiment, time delay stabilities could be maintained at about the 10 ps level on a global basis. [6] However, this would be a very expensive system, and would be inconvenient to use. Our goal in this paper is to provide adequate time and frequency sources and measurement systems, which are also convenient to use.

SOME RECENT RELEVANT EXPERIMENTS

Some of the most recent relevant experiments have arisen out of a technique called Enhanced GPS (EGPS). This is a systems approach. One of the main limitations of GPS as a time and frequency source is the degradation due to the intentional modulation called Selective Availability (SA). In this approach, a proper GPS receiver is interfaced through a filtering algorithm to a stable reference clock. For a fixed location, this approach can effectively filter the SA so that the output time and frequency transfer is optimized. Time stability and the published data frequency-transfer uncertainty show 20 ns and 10^{-13} with a high quality quartz crystal oscillator as the reference and 3 ns and 10^{-14} with a high quality cesium-beam standard as the reference, respectively. [4, 5] Significant errors can arise from group delay uncertainty in receivers, from errors in atmospheric modeling, SV clock and ephemeris errors, and the limited accuracy of the broadcast GPS - UTC(USNO MC) correction. Many of these errors respond well to averaging and all are bounded over time.

Experiments have been reported in which the carrier-phase of the GPS signal is utilized as in geodesy. Using post processing, these experiments have reported phase residuals over 24 hours -- comparing two hydrogen masers remote to each other -- of 0.2 ns. [1] In a near real time carrier-phase synchronization experiment, remote clocks were compared with precisions of 0.7 ns [2]. The baselines involved in these experiments have been up to 4 Mm. In the first case, given the estimated degrees of freedom involved, the frequency comparison uncertainty may be better than 10^{-15}.

Conceptually, the power of this approach can be seen as follows: the two remote clocks are phase tracking the same clock in space. It is like phase locking to the same source. Geodesy receivers have been designed to deal with sources of error in position down to the cm level. If this could be truly translated over to timing errors, the potential for very high accuracy frequency transfer may turn out to be one of the most cost effective means now available. Clearly, this accuracy implies a very accurate SV ephemeris. This usually is not available in real-time, and has to be obtained from post-processed data.

POTENTIAL FOR A HIGH ACCURACY REAL-TIME UTC ESTIMATE:

International Atomic Time (TAI) provides the international reference for the SI (System International) second based on the frequency defined for the ground-state of the $cesium_{133}$ atom. UTC provides the international reference for time. Since UTC is calculated several weeks after the fact, it cannot be used as a real-time reference. National timing center(s) predict estimates of UTC

for local distribution in real time. More and more synchronization needs cross country boundaries. If a highly accurate estimate of UTC were available in real-time, this would greatly facilitate any country's ability to synchronize it's time distribution systems and in turn be in much better agreement with the time signals from other countries. The implementation of such a real-time UTC estimate is discussed in the following paragraphs.

Basic principles affecting the generation of a real-time UTC:

a) TAI must be generated after the fact to assure immunity against outliers and anomalous behavior of contributing clocks -- with the goal of being syntonized to a best estimate of the SI second as provided by primary frequency standards. UTC is currently generated from TAI, using leap-seconds, in the usual way, and is the official international definition of time.

b) The responsibility for UTC in any country belongs to the designated timing centers within those countries, giving rise to the UTC(k)s. The UTC(k)s have a current goal of synchronization to UTC to within 100 ns.

c) The concept, explained below, of a real-time UTC would give the timing centers the additional responsibility of providing real-time estimates of UTC. Most timing centers, to meet the synchronization goal, are already calculating such an estimate of UTC.

d) The proposed implementation of real-time estimate of UTC involves two procedures and is shown schematically in Figure 1. First, each timing center continuously uses past BIPM "Circular-T" data in an optimum manner to form a current estimate of the difference between its time-scale, UTC(k) and UTC. The performance of available cesium standards is such that the difference can be extrapolated with a precision of tens of nanoseconds over two months. Second, the local estimates of UTC, UTC'(k) are continuously compared with the GPS composite clock using EGPS. The filtered outputs of these comparisons are then sent, almost in real time, by Internet to the BIPM,where they are ensembled using known weights to give the difference between the GPS clock and the average of the UTC'(k) estimates. The GPS clock, as realized by EGPS, together with the difference calculated almost in real time by the BIPM , and available by Internet, constitutes an optimum weighted estimate of a real-time UTC, UTC'. UTC' can then be replicated by any user with EGPS and a disciplined frequency standard. The quality of the resulting time-scale will depend on the quality of the local frequency standard.

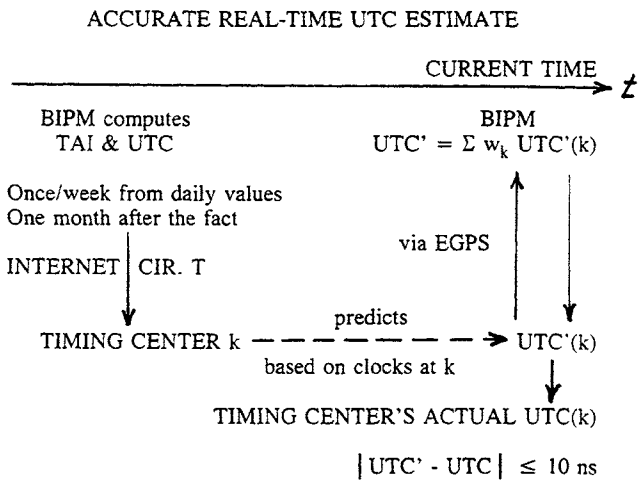


Figure 1

We now discuss how a real-time UTC might be realized in detail on the basis of some simple assumptions. Each timing center sends GPS files to BIPM every 5 days. It is assumed that TAI and UTC are calculated at a data rate of one sample each five days, and that the data is available after a 1 month delay. TAI is assumed to have a flicker floor of 1.0E-15 from five days to several months. A clock with performance comparable to that of an HP 5071A and an EGPS receiver is available at each timing center.

1) The $\sigma_y(\tau)$ characteristics of an HP 5071A clock are known and it reaches a flicker floor of 3×10^{-15} at about 80 days. A recursive, exponential filter is near optimum for estimating the frequency in this case with m = (τ value for reaching flicker floor in days x 0.6)/computation interval for TAI = 10. Computation interval, τ_o = 5 days.

2) When a new Circular T is received, the frequency estimate versus TAI is recursively updated using each point in the new Circular-T data and the following equation:

$$y_k'(t+\tau_o) = (m * y_k'(t) + y_k(t+\tau_o))/(m + 1), \quad (3)$$

where the primes "'" denote optimum estimator and the un-primed quantity is the value derived from the sequential Circular-T data. The initiation value is estimated in any reasonable way, which could be the first Circular-T value available for that clock. If other clocks exist at timing center k whose frequencies are known with respect to TAI, then a laboratory measurement of a new clock and estimating its frequency indirectly versus TAI will typically be more efficient. Any steering in TAI, of course, needs to be accounted for.

3) The last y_k' estimate is used for the prediction:

$$x_{kU'}(\text{current time}) = -x_{Uk}(\text{last Cir.T est.}) + \tau * y_k', \quad (4)$$

where τ is the prediction interval and the minus "-" sign on the first term is to reverse the sign in the Circular-T so that a positive frequency for the kth clock causes the time to run fast. If there is more than one clock at k, then they are weighted inversely proportional to $(\tau * \sigma_y(\tau))^2$ for each contributing clock. The kth clock is measured against GPS -- using, for example, EGPS. The GPS difference minus the predicted offset

$$x_{U'G}(k)@\text{current time} = x_{kG}(\text{current time}) - x_{kU'}(\text{current time}) \quad (5)$$

is communicated to the BIPM file, for example. Using this simple algorithm a new prediction is added to the file every five days as the GPS-CV data are down-loaded. The 40 day rms optimum time error of prediction for an HP 5071A is about 15 ns. Sixty of them optimally combined would yield a 40 day rms prediction error of $15/\sqrt{60} = 2$ ns assuming TAI were perfect. The instabilities in TAI obviously will add something to this number depending upon the algorithm used to compute TAI. The prediction value:

$$x_{U'G}@\text{current time} = \Sigma\, w_k\, x_{U'G}(k)@\text{current time}, \quad (4)$$

can be placed in a file accessible to all who wish to use it. The rms time prediction error of an HP 5071A over five days is about 5 ns if that were the update interval of the file. The time uncertainty will be reduced if the corrected GPS differences are ensembled more frequently than every five days.

The distribution of real-time estimate of UTC, UTC', could be made available in a file over Internet. If such a file were made generally available, then an unlimited number of users could steer to this estimate. Since GPS is acting as the transfer medium, like common-view, its time cancels out to within the measurement noise at the time of the measurement. Filtering the GPS SA with a quartz, rubidium or cesium based reference provides rms errors of about 20 ns, 10 ns and 3 ns, respectively. Over long-enough

averaging times, the corresponding frequency accuracies obtainable will be about 10^{-13}, a few x 10^{-14} and a few x 10^{-15}. This enhanced GPS approach (EGPS) has the advantage of providing a real-time estimate. Having access to the above file would allow the EGPS technique to be used as an inexpensive distribution of the real-time UTC estimate. Once one of these nodes was set up, it then, of course, could be used to distribute over fiber at about the same level of accuracy.

There are several potential error sources that would have to be dealt with in the distribution of UTC'. At the level of a few ns, uncertainties can come from the transfer via GPS. These include multipath, which can be from a few ns up to several tens of ns. mis-calibration of cable lengths and impedance matching; (with care these can be measured at the ns level in the above systems), Ionosphere, troposphere and coordinate errors can add several nanoseconds of uncertainty -- even tens of ns in some cases. Inadequate filtering of SA and improper receiver servo design can add even hundreds of ns. All of the above error sources can be kept below 10 ns in a properly designed system.

CONCLUSIONS:

Given current resources a global estimate of a real-time UTC could be achieved with an accuracy of 10 ns or better. In order to achieve this level of accuracy, all of the error sources need to be dealt with adequately and appropriate procedures followed. An accuracy of 20 ns could be achieved very cost effectively over a large number of users. This could provide an important synchronization support for the user community allowing world-wide networks to truly be synchronized. Such a system could provide a great improvement in spectrum utilization if network architecture were to take advantage of this resource. True time synchronization could become a general utility of high value helping to preserve the very precious resource of the electromagnetic spectrum.

Well over a thousand papers were published last year on using GPS for navigation or position determination. It has become a true utility. The telecom community is beginning to use it extensively as a very accurate frequency reference. The resources are now in existence to use it as a UTC time utility. It is a matter of cooperation and commitment to a reasonable set of procedures to make it become an extremely useful time utility. There may come a point where its usefulness would justify letting it be named UTC. TAI would, of course, still be calculated after-the-fact for safety and accuracy considerations, and UTC would be phase-locked in real-time to within a few ns to TAI minus the appropriate number of leap seconds. Each country's UTC(k) could provide a much more meaningful service because the accuracy and stability of each UTC(k) would be improved significantly.

With 10^{-15} accuracy imminently expected from primary frequency standards research efforts, methods to conveniently compare them are very important. Having UTC as a 10^{-15} reference will help considerably. However, having a measurement system with frequency transfer uncertainties approaching 10^{-16} would be very helpful. Candidates are TWSTFT, use of GPS carrier-phase and EGPS used in the common-view mode. In all three cases the degrees of freedom and the signal to noise are adequate to provide uncertainties approaching 10^{-16} for integration times of a few days.

Much experimental work is yet to be done to validate the above approaches. The conclusions are based on a few experiments and on theoretical extrapolations, but the evidence is that those extrapolations are sound.

ACKNOWLEDGEMENTS:

The authors are indepted to Dr. Jacques Azoubib of the Bureau International des Poids et Mesures for supplying the three years of clock data studied to develop the results presented in this paper. Mr. Mihran Miranian of USNO and Dr. Marc Weiss and Victor Zhang of NIST were very cooperative in supplying data and assisting in data acquisition.

REFERENCES

1) Dunn, C., Lichten, S., Jefferson, D., and Border, J. S., "Sub-nanosecond Clock Synchronization and Precision Deep Space Tracking," Proceedings 23rd Precise Time Time Interval Applications and Planning Meeting, Pasadena, CA, 1991. NASA Conference Publication #3159, pp. 89-.

2) Basker, S., Casewell, I., "Time and Clock Issues in WADGPS," Proceedings 1994 ION GPS-94, Sept. 20-23, 1994, Salt Lake City, Utah, pp. 1509-1518.

3) Kusters, J. A., Ho, K., Giffard, R. P., Cutler, L. S., and Allan, D. W., "A No-drift and less than 1 x 10^{-13} Long-term Stability Quartz Oscillator Using a GPS SA Filter," Proceedings IEEE 48th International Frequency Control Symposium, Boston, MA, 1994.

4) Allan, D. W., Kusters, J., and Giffard, "Civil GPS Timing Applications," Invited paper for Plenary Session, ION-GPS-94 in Salt Lake City, 20 Sept.'94, Proceedings pp 25-32.

5) Kusters, J., Giffard, R., Cutler, L., Miranian, M., and Allan, D., "A Globally Efficient Means of Distributing UTC Time and Frequency through GPS," Proceedings of 26th Annual Precise Time and Time Interval Applications and Planning Meeting, Dec. 5-8, 1994, Reston, VA.

6) Allan, D. W., Alley, C. O., Ashby, N., Decher, R., Vessot, R. F. C., and Winkler, G. M. R., "Ultra-Accurate International Time and Frequency Comparison Via an Orbiting Hydrogen-Maser Clock," Journal De Physique, Colloque C8, 395-413, 1981.

AN AUTOMATED SIGNAL DELAY MONITORING SYSTEM FOR A TWO-WAY SATELLITE TIME TRANSFER TERMINAL

D. Kirchner[1], H. Reßler[2] and R. Robnik[2]

[1] Technical University Graz, Austria
[2] Space Research Institute, Graz, Austria

ABSTRACT

A fully automated two-way time and frequency transfer system including a satellite simulator which allows to carry out signal delay measurements in conjunction with each time transfer measurement is described. First results are presented and compared with other recent results. Finally envisaged activities are mentioned.

Keywords: Two-way satellite time and frequency transfer (TWSTFT), satellite simulator, signal delay, stability.

1. INTRODUCTION

Two-way satellite time and frequency transfer (TWSTFT) via communication satellites using spread-spectrum techniques provides highest precision and has the potential of high accuracy. To exploit this accuracy potential it is necessary to correct the measurements for the signal delays in the satellite transponders, if different transponders are used for both signal directions, and for the signal delays in the earth stations. The differential delays of the earth stations involved can either be measured by a portable station or the delays of each station may be measured separately by means of a satellite simulator attached to the station as proposed by de Jong [1]. A fully automated measurement system of this type has been implemented in the dedicated VSAT-type TWSTFT earth station of the Technical University Graz (TUG) allowing frequent monitoring of the station delays. Such delay monitoring is of great importance to detect variations of the station delays and for the correction of the data possibly on a routine basis.

The correction which has to be applied for TWSTFT measurements to compensate for the non-reciprocity caused by different delays of both stations involved is given by $[(\tau_1^{TX} - \tau_1^{RX}) - (\tau_2^{TX} - \tau_2^{RX})]/2$, i.e. the difference of the differential delays of the transmit and receive parts of earth stations 1 and 2 [2]. The transmit delay is the total delay from the transmitted one pulse per second (1 pps) to the reference plane of the antenna and the receive delay is the total delay of the received 1 pps. Both delays consist of the corresponding signal delays of the earth station, the modem, related equipment, and signal delays in the connecting cables. The satellite simulator (SATSIM) employed allows to measure most of these delays apart from some remaining delays which have to be evaluated separately. The signal delays are measured by means of the spread-spectrum modem used for the time and frequency transfer measurements. The separate transmit and receive delays of a single modem can be measured by means of an oscilloscope, but only with low accuracy. With two modems the corresponding differential delays can be established with high accuracy.

2. MEASUREMENT SET-UP

The circuit diagram of the SATSIM used is given in Fig. 1. It is of the de Jong type [1,3] - this means it can measure the sum of the earth station transmit and receive delays and the receive delay, thus allowing to calculate the difference of the transmit and receive delay - but shows some modifications. The receive and transmit antennas are not simply waveguide-to-coax transitions, but horn antennas and in the receive part there is a power splitter making possible to measure power and frequency of the signal transmitted by the earth station. In the transmit part a combination of attenuators is used to obtain the same signal power as received from the satellite. The connections are by semi-rigid cables and SMA-connectors. For shielding purposes all components are in a small metallic box with the horn antennas protruding from the box. The box is mounted on the feed boom of the parabolic antenna with the horns facing the feed. The side of the box with the horns is covered with microwave absorbing material and protection from rain is achieved by a small dome (see Fig. 2).

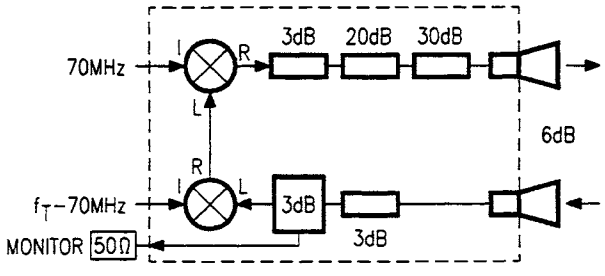


Figure 1 Circuit diagram of the SATSIM.

A simplified block diagram of the TWSTFT set-up including the SATSIM and related equipment, and the measurement and control system is shown in Fig. 3. The satellite terminal provides remote control of the transmit power in 1 dB steps and of the transmit and receive frequencies in 1 MHz steps. The 1 kHz step size required by the MITREX modem [4] or other modems not providing frequency agility is accomplished by the 70/70 MHz converter [5]. The MITREX modem is fully automated via the MITREX control system (MCS) which also provides remote control of the synthesizers used in the 70/70 MHz converter and of several switches in order to measure the different loop delays and the synthesizer frequencies. The spectrum analyzer allows the automated measurement of signal frequencies, signal levels, carrier-to-noise density ratios (C/N_O) and the storing of spectra via the PC which controls the whole system and - apart from other operational features - provides real-time graphics of the TWSTFT measurement data.

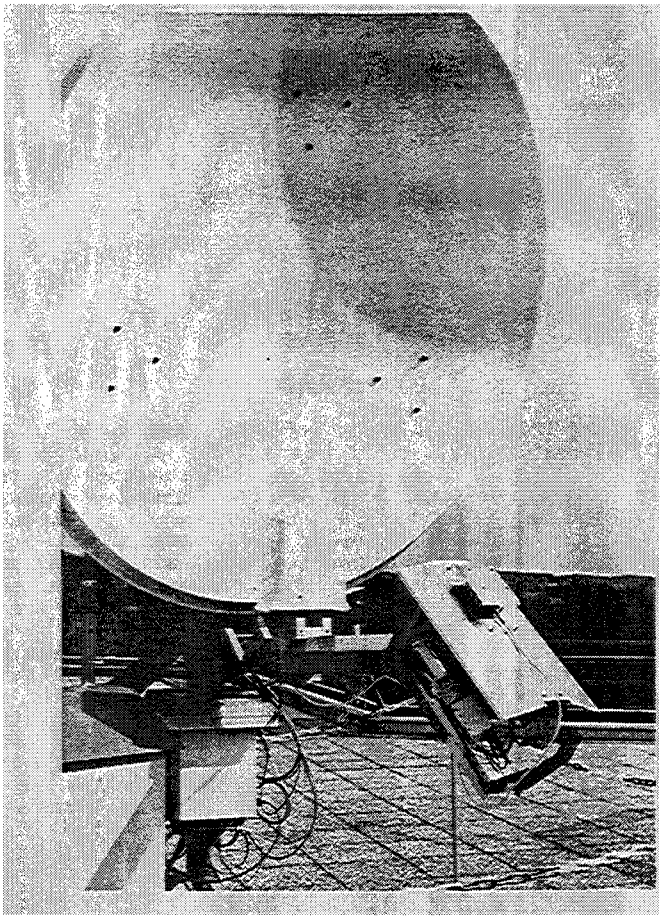


Figure 2 Satellite earth station with the SATSIM mounted on the feed boom.

Remote control (different password dependent levels) of the station is provided via telephone modem. Apart from the actual time transfer measurements a time transfer session consists of several accompanying measurements: collection of meteorological data, a counter check, the modem calibration, C/N_0 measurements of the satellite beacon, of the station carrier via the SATSIM and - if the own carrier can be received - via the satellite and - if provided - of the carriers of the partner stations, and the different loop measurements necessary for the calculation of the signal delays of the station.

A block diagram of the station from the point of view of the signal delays involved and the possible loop arrangements to measure the signal delays of interest are shown in Fig. 4. The different loops and the corresponding counter readings (REF - PPSRX + b2) are called: MOD for modem loop, ID for indoor loop (all indoor equipment is in a fully air-conditioned room), OD for outdoor loop, STR for SATSIM loop to measure the station transmit and receive delay and SR for SATSIM loop to measure the station receive delay. Together with the wanted delays other delays are measured which can only partly be eliminated by combining different measurements. These remaining delays have to be evaluated separately. CAL is the modem calibration (REF - PPSTX + b1). Each delay indicated in a square gives the signal delay between the points marked by dots, e.g. u1 is the signal delay from the modem output to the indoor switch input and so on. In order to distinguish between the transmit and receive delays mentioned above and indicated by τ the measured delays are indicated by t. The transmit delay t^{TX} (modem transmit output to satellite simulator input not including the cable connecting indoor and outdoor equipment) given by (u1 + i1 + o1 + u3 + u4) is obtained by calculating (STR - SR) and applying the correction [c1 + c2 + (s2 - s1) + (cc - uc)]. The receive delay t^{RX} (satellite simulator output to modem receive input not including the cable connecting outdoor and indoor equipment) given by (d4 + d3 + o2 + i2 + d1) is obtained by calculating (SR - MOD - OD + ID) and applying the correction [-(c1 + c2) - s2 - (cc - uc) - (m1 + m2 - m3) - (i3 - o3 -i1 - i2)]. The differential delay $(t^{TX} - t^{RX})/2$ is obtained by calculating (STR - 2*SR + MOD + OD - ID)/2 and applying the correction [(c1 + c2) + (s2 - s1/2) + (cc - uc) + (m1 + m2 - m3)/2 + (i3 - o3 - i1 - i2)/2]. Apart from (c1 + c2) all other terms are about zero or smaller than 1 ns. The delay c1 is 14.55 ns and c2 is about 20 ns. The cables connecting the indoor and outdoor equipment (delays: cc, uc, dc) are parts of equal lengths (approx. 30 m) of one cable and in the same duct. Therefore signal delay variations are assumed to be equal for all of them. The sum of uc and dc and related delays (OD - ID) is measured for each session and the three individual delays (cc, uc, dc) are measured occasionally like other delays which in the present set-up cannot be measured in an automated mode. In the following the respective delays t are given without the above mentioned corrections and are designated by T.

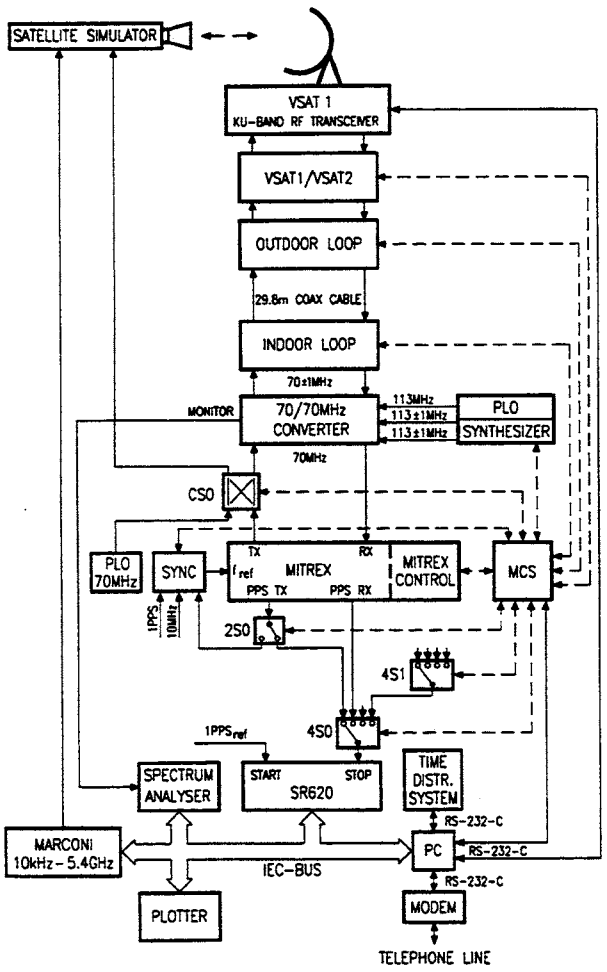


Figure 3 Simplified block diagram of the TWSTFT set-up.

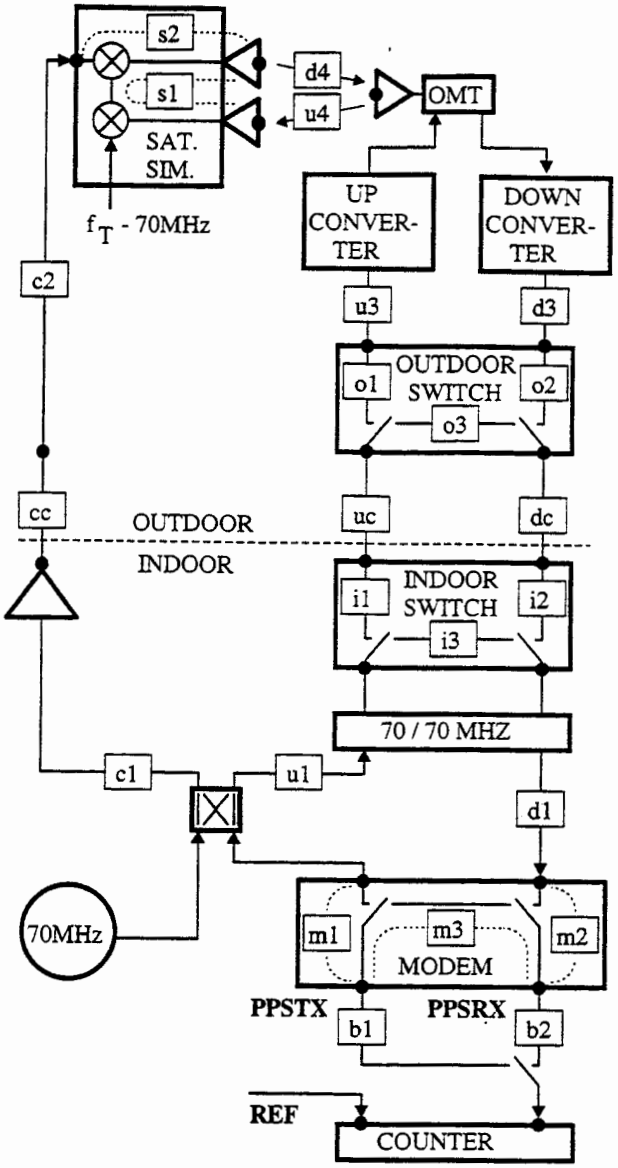


Figure 4 Station diagram showing the different signal delays and measurement loops.

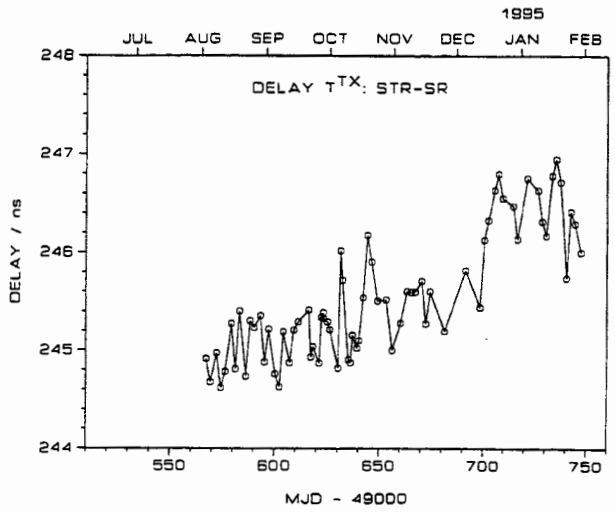


Figure 5 Transmit delay T^{TX}.

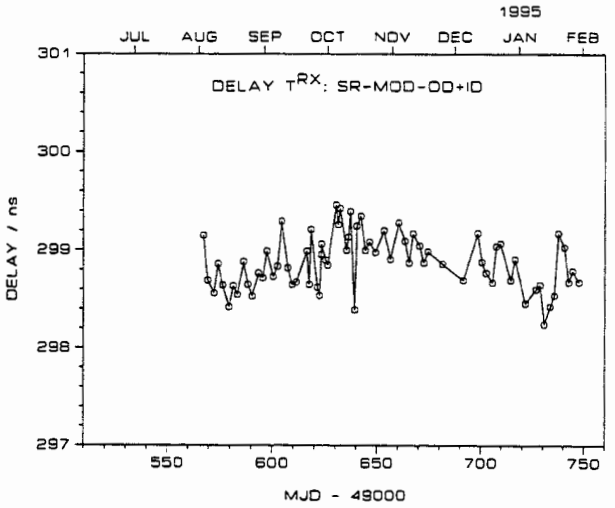


Figure 6 Receive delay T^{RX}.

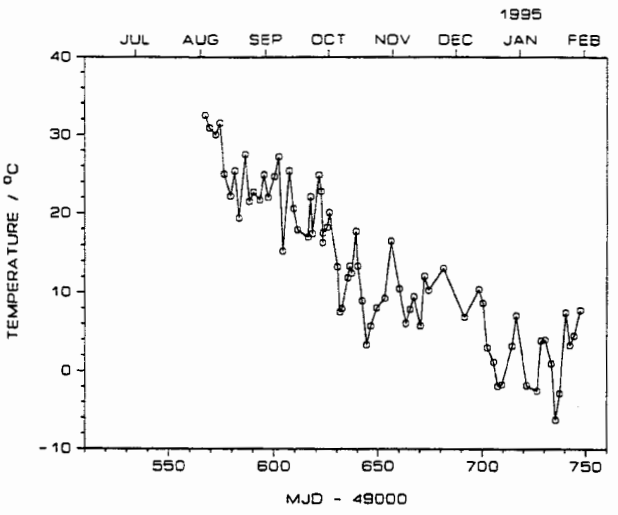


Figure 7 Outside temperature.

3. MEASUREMENTS AND RESULTS

TWSTFT measurements are carried out between two laboratories in the USA and six laboratories in Europe [6,7]. Since summer 1994 in connection with each session all loop measurements necessary to calculate the different delays are performed, and in the following the delay measurements accompanying the European sessions are presented. Each single loop measurement is the mean of 100 measurements with one measurement per second. The completion of all loop measurements takes about fifteen minutes. Figs. 5 and 6 show the transmit delay T^{TX} and the receive delay T^{RX}, respectively. The outdoor temperature and relative humidity is given in Figs. 7 and 8. There is an obvious correlation between the delays and the temperature and humidity. As an example the temperature dependent variation of T^{TX} is plotted in Fig. 9. The total delay variation over a temperature range of about 37°C is smaller than 2.5 ns. Polynomial fits - as indicated in Fig. 9 for the T^{TX} temperature behaviour - were used to model the different temperature and humidity dependencies in order to correct for them thus removing systematic effects. The delay T^{TX} corrected accordingly is plotted in Fig. 10. The differential delay $(T^{TX} - T^{RX})/2$

computed from the delays as measured and that computed from the delays corrected for the modelled temperature and humidity behaviour is given in Figs. 11 and 12, respectively. The measurement error estimated by an error budget taking into account the errors contributed by the single measurements is smaller than 50 ps. The total variation of the uncorrected differential delay is about 1.5 ns and that of the corrected is below 1 ns.

4. DISCUSSION OF RESULTS

The stability [8] of the differential signal delay $(T^{TX} - T^{RX})/2$ as measured and the corrected one, together with stability data of two common clock experiments carried out in the USA at the National Institute of Standards and Technology (NIST) [9,10], are given in Fig. 13, which is a composite time and frequency stability plot [11]. In a diagram of the time stability $\sigma_x(\tau)$ lines of Mod.$\sigma_y(\tau)$ is constant are plotted making use of the relationship $\sigma_x(\tau) = \text{Mod.}\sigma_y(\tau)\,\tau/\sqrt{3}$. Thus from only one graph the time and frequency transfer capability of a system can be estimated. For averaging times up to some minutes TWSTFT usually shows white-noise PM behaviour as indicated in Fig. 13 with the noise level

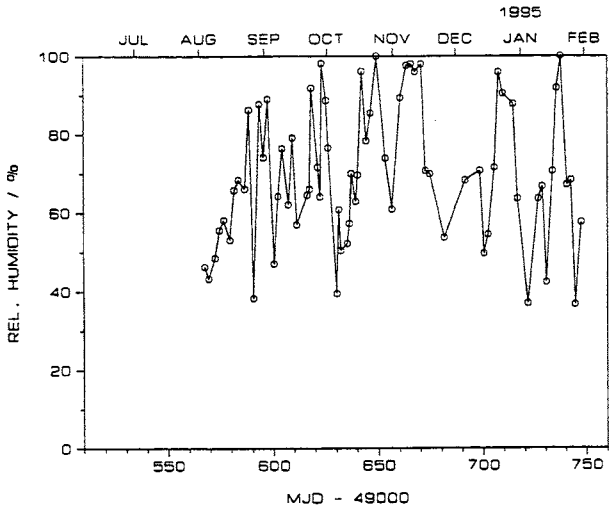


Figure 8 Outside relative humidity.

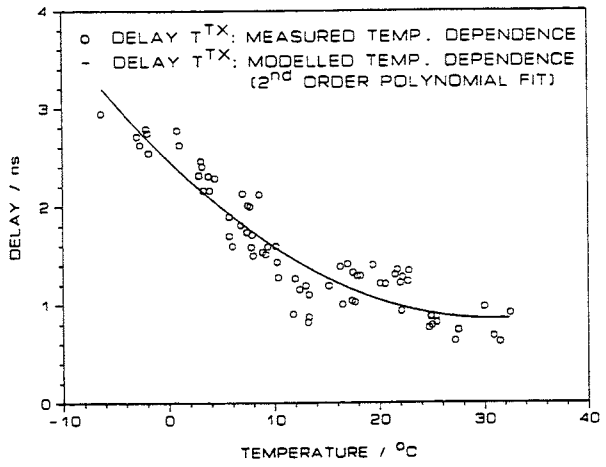


Figure 9 Temperature dependence of the delay T^{TX} together with the polynomial fit used for modelling of the temperature dependence.

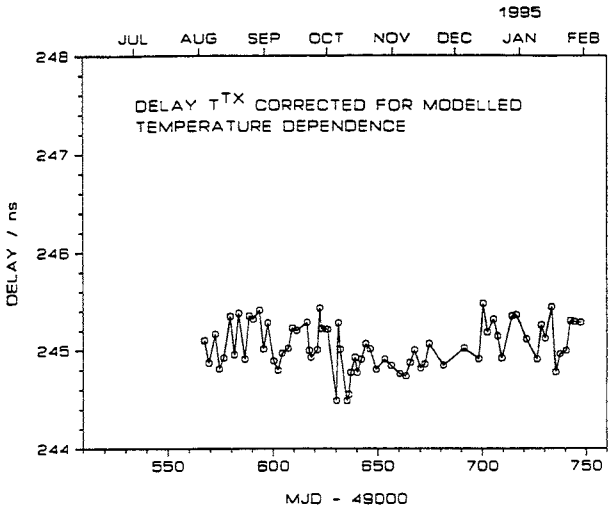


Figure 10 Transmit delay T^{TX} corrected for modelled temperature dependence.

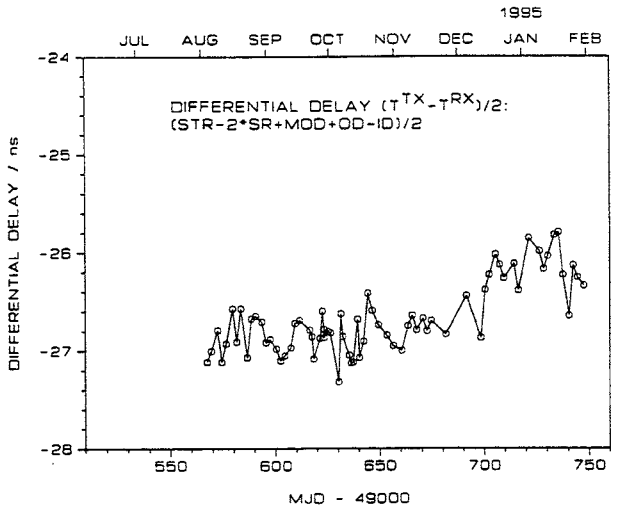


Figure 11 Differential delay $(T^{TX} - T^{RX})/2$ computed from T^{TX} and T^{RX} as measured.

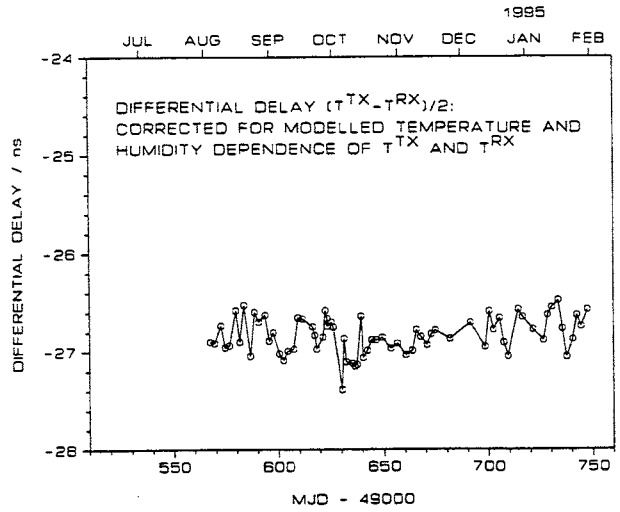


Figure 12 Differential delay $(T^{TX} - T^{RX})/2$ computed from T^{TX} and T^{RX} corrected for modelled temperature and humidity dependence of T^{TX} and T^{RX}.

depending on the actual C/N_0. Studying the stability of TWSTFT for longer averaging times - in order not to be limited by the stability of the clocks employed - common clock arrangements are used. The results of two such experiments carried out at the NIST are given for comparison. In one experiment (1) two earth stations of VSAT-type together with a satellite simulator on top of a mountain were employed [9]. Means of sets of hundred measurements, one measurement per second, were used to calculate the stability which for averaging times longer than about thousand seconds reveals flicker-noise PM at a level of about 20 ps. This stability limit seems to come from the MITREX modem itself as experienced at the TUG by investigating one of these modems. Station delay variations caused by temperature and humidity effects may have cancelled because of the similarities in the earth stations. In the other experiment (2) the measurements were done via a satellite using a VSAT-type and a 3.7 m earth station [10]. Measurements were done three times per week during sessions of five minutes duration and the results were interpolated to one-day intervals for the stability calculation revealing a flicker-noise PM level of about 200 ps. The measurements of the differential signal delay at the TUG were also done three times per week and interpolated to one-day intervals. They show a somewhat lower noise level than (2) for the data as measured (a) and as one can expect an even lower level (below 100 ps) for the data corrected by modelling the temperature and humidity behaviour of the transmit and receive delays (b). This result indicates that performing TWSTFT between stations both equipped with a SATSIM and correcting the TWSTFT data of each station by the measured signal delay variations could considerable improve the TWSTFT stability.

5. CONCLUSION AND ENVISAGED ACTIVITIES

In a measurement set-up designed for fully automated measurements a SATSIM can easily be included and operated. Because during the SATSIM operation the earth station is transmitting, about five minutes extra satellite time is needed. The data obtained can immediately be used to correct the TWSTFT data for possible variations of the differential signal delay of the station. At the TUG a second earth station will soon be available allowing common clock experiments between two stations both equipped with a SATSIM thus being able to obtain some more information on the stability of TWSTFT systems. Furthermore the second station will be used for detailed investigations concerning temperature and humidity dependence of the signal delays in order to optimize the station design. Still a problem is the availability of modems appropriate for this work.

ACKNOWLEDGEMENTS

The contribution of INTELSAT providing free of charge satellite transponder time and the help of W. Schladowsky and H. Peintinger (ÖPTV) in administrative and technical matters is gratefully acknowledged. The loan of a MITREX modem initially by the U.S. Naval Observatory and later on by Prof. Hartl, University of Stuttgart, is deeply appreciated. The work was supported by the Austrian Academy of Sciences and the Jubilee Fund of the Austrian National Bank.

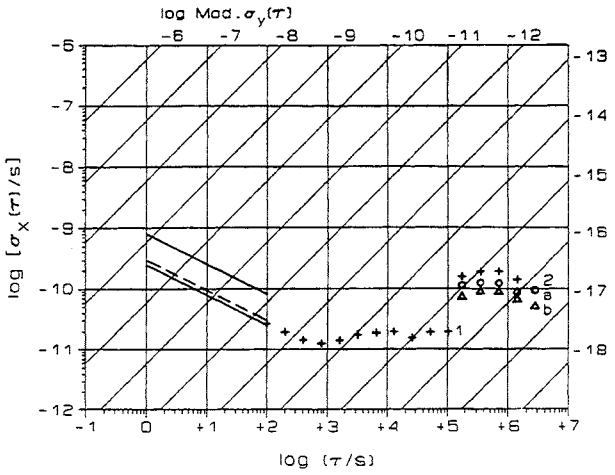


Figure 13 Stability of differential delays $(T^{TX} - T^{RX})/2$ as measured (a) and corrected for temperature and humidity (b), and of common clock TWSTFT experiments employing stations of the same type and a satellite simulator (1) and stations of different type and a satellite (2).

REFERENCES

[1] G. de Jong, "Accurate delay calibration of satellite ground stations for two-way time transfer", Proc. 3rd European Frequency and Time Forum, pp. 198-203, 1989.

[2] D. Kirchner, "Two-way satellite time transfer via communication satellites", Proc. IEEE, vol. 79, pp. 983-990, 1991.

[3] G. de Jong, "Automated delay measurement system for an earth station for two-way satellite time and frequency transfer", Proc. 26th Annual PTTI Meeting, 1994, (in press).

[4] Ph. Hartl et al., "Spread spectrum time transfer experiment via INTELSAT", Proc. 15th Annual PTTI Meeting pp. 331-356, 1983.

[5] D. Kirchner, "Design considerations for Ku-band two-way satellite time transfer terminals", Proc. 4th European Frequency and Time Forum, pp. 631-637, 1990.

[6] J.A. Davis et al., "Two-way satellite time transfer experiments between six European laboratories using the INTELSAT (VA-F13) satellite", Proc. 8th European Frequency and Time Forum, Vol.I, pp. 296-314, 1994.

[7] J.A. DeYoung et al., "Preliminary results of the 1994 international transatlantic two-way satellite time and frequency transfer experiment", Proc. 26th Annual PTTI Meeting, 1994, (in press).

[8] D.W. Allan and A. Lepek, "Trends in international timing", Proc. 7th European Frequency and Time Forum, pp. 221-227, 1993.

[9] ITU, International Telecommunication Union - Radiocommunication Study Groups, Doc. 7-2/TEMP/5-E, Nov. 1994.

[10] ITU, International Telecommunication Union - Radiocommunication Study Groups, Doc. 7-2/TEMP/3-E, Nov. 1994.

[11] D. Kirchner, "Two-way satellite time and frequency transfer: Principle, implementation and practical operational requirements and performance", PTTI Tutorial (Understanding and getting the most from time transfer: An applications tutorial), 25th Annual PTTI Meeting, 1993.

Session Time transfer Number 4

GLONASS COMPUTER CONTROLLED RECEIVER: FIRST EXPERIMENTS ON TIME TRANSFER

N. Koshelyaevsky[1], A. Zorin[2]

[1] GP "VNIIFTRI", Mendeleevo, Moscow Region, Russia

[2] Moscow State Institute for Electronic Engineering, Moscow, Russia

GLONASS manually controlled receiver of A-724M type is well known for time transfer application in Russia [1,2]. Basing on a specially designed interface card we have succeeded for the first time to demonstrate automatic operation of this type receiver for time transfer experiments. The format of the output data meets requirements of TECHNICAL DIRECTIVES FOR STANDARDIZATION OF GPS TIME TRANSFER RECEIVER SOFTWARE [3].

References

1. P. Daly, N.B. Koshelyaevsky, W. Lewandowski, G. Petit and C. Thomas, Comparison of GLONASS and GPS Time Transfer, Metrologia, 1993, 30, 89-94.

2. A. Al'shina, N. B. Koshelyaevsky, S. Pushkin and P. Daly, Russian National Time Scale Long Term Stability Estimation Using GLONASS Time Link, Proceedings of the Seventh European Frequency and Time Forum, Neuchâtel, March 16 - 18, 1993.

3. Rapport BIPM-93/6 TECHNICAL DIRECTIVES FOR STANDARDIZATION OF GPS TIME TRANSFER RECEIVER SOFTWARE (TD SGPS TTRS) Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92310 Sèvres, France, October 1993.

CLOSING ERRORS OBSERVED DURING TWO-WAY TIME TRANSFER EXPERIMENTS

J A Davis, P R Pearce and K A Bartle

National Physical Laboratory, Queens Road, Teddington, Middlesex, UK.

P Hetzel

Physikalisch-Technische Bundesanstalt, Braunschweig, Germany.

D Kirchner

Technical University Graz, Graz, Austria.

H Ressler

Space Research Institute, Graz, Austria.

W Powell, A McKinley, J A De Young, W Klepczynski

United States Naval Observatory, Washington, DC, USA.

G de Jong

NMi, Van Swinden Laboratories, Delft, The Netherlands.

A Söring

Forschungs- und Technologiezentrum, DBP Telekom, Darmstadt, Germany.

F Baumont, P Claudon, P Grudler

Observatoire de la Cote d'Azur, Grasse, France.

C Hackman

National Institute of Standards and Technology, Boulder, Colorado, USA.

L Veenstra

RSI, Comsat World Systems, 6560 Rock Spring Drive, Bethesda, Marylands, USA.

ABSTRACT

The results of a study of closing errors observed during an international Two Way Satellite Time and Frequency Transfer (TWSTFT) experiment are reported. Closing errors have been examined during regular TWSTFT experiments performed within Europe and between Europe and North America. The characteristics of these closing errors, including their stability over periods of several months, are examined.

The origins of the closing errors are discussed in terms of the asymmetries in the delays within each earth station's instrumentation. Experiments are performed using co-located satellite earth stations. These experiments have demonstrated that at the nanosecond level, the value of the receive chain delay within the earth station instrumentation is dependent on the characteristics of the up-link signal and hence on the choice of transmitting station. A direct consequence of this result is that for calibration purposes the up-link delay of a transmitting earth station and the down-link delay of the partner station must be considered together rather than as individual elements.

keywords
Time transfer, satellite communications, microwave measurement, delay measurement.

1) INTRODUCTION

A Two-Way Satellite Time and Frequency Transfer (TWSTFT) experiment has been recently performed between six European and two North American laboratories [1]. This experiment has shown that the two-way system is easy to operate on a regular basis, and offers both high measurement precision and ultimately high accuracy. The usefulness of the TWSTFT method for regular comparisons of clocks at primary timing laboratories was clearly demonstrated.

Future developments of TWSTFT systems may be aimed at high precision frequency comparisons of new ultra stable frequency standards, where frequency comparisons of 1 part in 10^{15} are required with averaging times of less than one day. Recent experiments from co-located TWSTFT stations suggest that the stability of the TWSTFT instrumentation may be sufficient for this application [2].

A second use for the TWSTFT method is for the international comparison of time scales contributing to international co-ordinated time (UTC). These comparisons are at present being made using the method of common view of Global Positioning System (GPS) satellite signals. TWSTFT offers several advantages over the GPS method, both in the precision obtainable from the measurements, shorter averaging time required, and ultimately higher accuracy. The TWSTFT method may also be used as a check on the performance of the GPS common view method.

Recent experiments have involved the use of a portable X-band TWSTFT earth station which was transported between different GPS installations [3]. The use of this X-band station with hemispherical satellite beams will enable the delay asymmetries in transatlantic Ku-band transponders of the INTELSAT (VA-F13) satellite to be calibrated, enabling accurate transatlantic TWSTFT's to be performed using the INTELSAT satellite at Ku band.

Detailed analysis of the results of the international 'field trial' experiment has revealed systematic closing errors of up to several nanoseconds within a group of time transfers [4]. The presence of these closing errors may limit the ultimate accuracy obtainable from the TWSTFT system. In order to understand further the characteristics and origin of these closing errors a detailed study has been undertaken. The results of this study are presented here.

2) METHOD

Eight laboratories participated in this experiments. These were the Technical University of Graz, Austria (TUG), National Physical Laboratory, Teddington, Middlesex, United Kingdom (NPL), NMi, Van Swinden Laboratories, Delft, Netherlands, Forschungs- und Technologiezentrum, DBP Telekom,

Darmstadt, Germany (FTZ), Physikalisch-Technische Bundesanstalt, Braunschweig, Germany (PTB), Observatoire de la Cote d'Azur, Grasse, France (OCA), National Institute of Standards and Technology, Boulder, Colorado, USA (NIST), and the United States Naval Observatory, Washington, DC, USA (USNO). Cost free satellite time was provided for a year by INTELSAT on their (VA-F13) satellite, which enabled the experiment to proceed.

The arrangement and characteristics of the earth stations used in the international TWSTFT experiment have been reported previously [1], [5]. A summary of the characteristics of the European stations is given in Table 1 of [5].

Regular time transfers sessions were performed three times per week over a period of one year. Each session consisted of a series of individual time transfers each with a measurement period of 5 minutes. Up to three time transfers were performed simultaneously, a complete measurement session lasted approximately 40 minutes.

3) CLOSING ERRORS

3.1) Definition

Consider the time transfers made between clocks located at earth stations A, B, and C. The time transfer between clocks at two distant locations may be determined using equation (1) of [6]. The time transfers $(T_A - T_B)$, $(T_B - T_C)$, and $(T_C - T_A)$ may be added to form a zero time transfer closure. Following the analysis given in [4]:

$$(T_A - T_B) + (T_B - T_C) + (T_C - T_A)$$
$$= 0.5(\Delta T_{AB} - \Delta T_{BA}) + 0.5(\Delta T_{BC} - \Delta T_{CB}) + 0.5(\Delta T_{CA} - \Delta T_{AC}) \quad (1)$$

where ΔT_{AB} is the counter timer readings made at station A when receiving signals from a partner station B. The above combination of counter timer measurements were found not to sum to zero. The discrepancy from zero closure was defined as the closing error C_{ABC} where:

$$C_{ABC} = 0.5(\Delta T_{AB} - \Delta T_{BA}) + 0.5(\Delta T_{BC} - \Delta T_{CB}) + 0.5(\Delta T_{CA} - \Delta T_{AC}) \quad (2)$$

3.2) Closing errors determined from regular measurements.

Closing errors may be easily calculated from the regular time transfer measurements. Examples of closing errors determined from both the European and Transatlantic time transfers are shown in Figures (1a,b,c), these have been calculated over a period of 350 days. The mean values and standard deviations of European, transatlantic, and mixed closing errors, are given in Table 1.

Discrete changes in the closing error occurred on MJD 49480 for the PTB-USNO-NPL-NIST closing error, shown in Figure 1b. These changes may be correlated with the replacement of the time transfer modems operating at USNO. Discrete changes were observed in some but not all of the combinations of closing errors calculated from the time transfers involving USNO. Unexplained discrete changes have occurred in the closing errors, for example the NPL-NIST-PTB closing error on MJD 49448, shown in Figure 1c. Closing errors may change over a period of several months as is shown in Figure 1a.

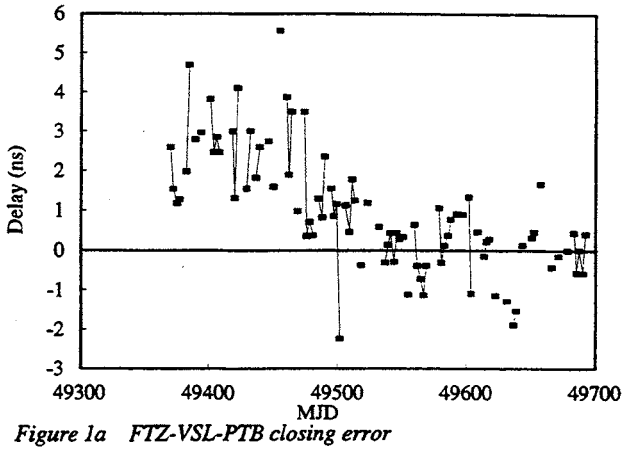


Figure 1a FTZ-VSL-PTB closing error

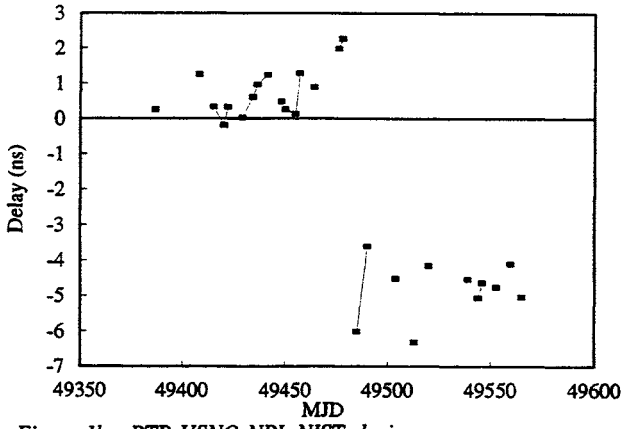


Figure 1b PTB-USNO-NPL-NIST closing error

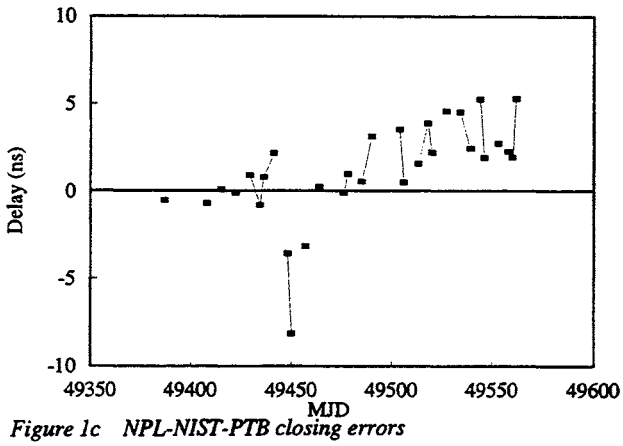


Figure 1c NPL-NIST-PTB closing errors

The data scatter varied significantly between the measurements of different closing errors formed from the European time transfers. The lowest values occurred for the TUG-PTB-NPL closing error (Table 1). Similar results were obtained with the transatlantic closing errors, the scatter on the measurement of transatlantic and European closing errors was in general comparable. The scatter on the measurement of mixed closing errors was noticeably higher. The data scatter is due both to measurement noise, and also to clock noise and instrumentation delay changes occurring within the two-way system between the successive measurement sessions. Approximately 3 hours occurred between the European and transatlantic measurements. During this period the settings of both the frequency and level of the Tx and Rx signals have been changed.

Table 1 Closing errors determined from regular measurements

Closing Error	NPL-TUG TUG-PTB PTB-NPL	PTB-NPL NPL-FTZ FTZ-PTB	VSL-TUG TUG-FTZ FTZ-VSL	NPL-TUG TUG-FTZ FTZ-NPL
Mean	-0.72 ns	2.21 ns	-0.50 ns	-0.67 ns
Standard Deviation	0.36 ns	0.78 ns	1.53 ns	0.79 ns

a) European closing errors (MJD 49440 - MJD 49660)

Closing Error	PTB-USNO USNO-NPL NPL-NIST NIST-PTB	VSL-USNO USNO-TUG TUG-NIST NIST-VSL	PTB-USNO USNO-FTZ FTZ-NIST NIST-PTB	FTZ-USNO USNO-NPL NPL-NIST NIST-FTZ
Mean	-4.79 ns	0.16 ns	-3.33 ns	-1.56 ns
Standard Deviation	0.76 ns	0.78 ns	1.04 ns	1.09 ns

b) Transatlantic closing errors (MJD 49480 - MJD 49565)

Closing Error	TUG-USNO USNO-VSL VSL-TUG	TUG-USNO USNO-NPL NPL-TUG	NPL-USNO USNO-VSL VSL-NPL	VSL-NIST NIST-FTZ FTZ-VSL
Mean	0.27 ns	-5.28 ns	-5.61 ns	-5.60 ns
Standard Deviation	1.00 ns	1.01 ns	3.70 ns	6.14 ns

c) Mixed European / transatlantic closing errors (MJD 49480 - MJD 49565)

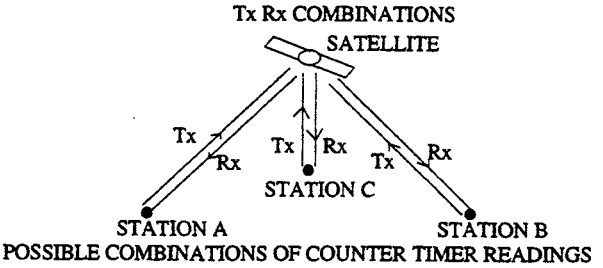
A systematic error may occur in the determination of the closing errors if there is a large difference in the rates of the clocks at two locations. The largest rate difference occurred for time transfers made using clocks at TUG, this was typically 8ns/day, this resulted in a maximum systematic error of 300ps, which was small in relation to the closing errors.

3.3) Specific experiments

A series of experiments have been performed to determine the origin of the closing errors. Experiments were performed to study the effects on the closing error of changing the modem triggering, so that all of the counter timers would similarly trigger on either the rising or falling edge of the 1pps modem output signal. Experiments were performed to study the effects of residual satellite motion, incorrectly set system clocks, and changing the modem codes. In all of these experiments, no changes were observed in the closing errors that were beyond measurement uncertainty. An experiment to determine the effect of reducing the levels of up-link EIRP was inconclusive as the reduced signal levels resulted in an increased data spread which increased the uncertainty in the closing error determination. The value of the closing errors appear insensitive to changes of many of the earth station operating parameters.

3.4) Closing error determined from combinations of satellite links.

The method of forming closing errors from regular measurements suffers from the disadvantage that the individual time transfers must be performed in sequence. This introduces uncertainties in the result which are dependent both on clock noise and instabilities of the earth station instrumentation delays. Experiments may be performed in which the closing errors are determine without this requirement to 'flywheel' on the local atomic clocks.



POSSIBLE COMBINATIONS OF COUNTER TIMER READINGS

(a) $(\Delta T_{AA} + \Delta T_{BB} + \Delta T_{CC})$ (b) $(\Delta T_{AA} + \Delta T_{BC} + \Delta T_{CB})$

(c) $(\Delta T_{AB} + \Delta T_{BC} + \Delta T_{CA})$ (d) $(\Delta T_{BB} + \Delta T_{AC} + \Delta T_{CA})$

(e) $(\Delta T_{AC} + \Delta T_{CB} + \Delta T_{BA})$ (f) $(\Delta T_{CC} + \Delta T_{AB} + \Delta T_{BA})$

Figure 2 Combinations of PN coded satellite links

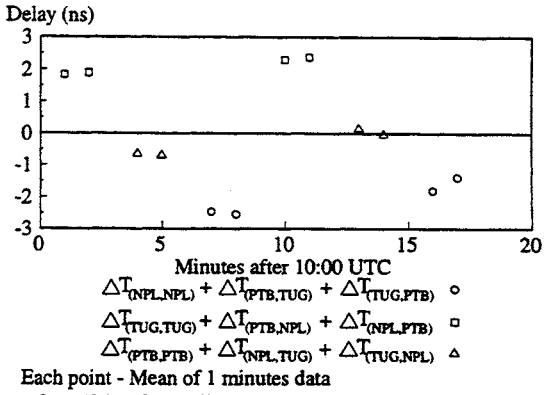


Figure 3a 'Mixed' satellite links

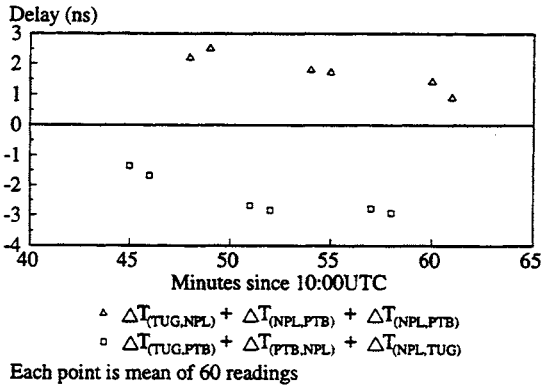


Figure 3b 'Loop' satellite links

Combination of PN coded satellite links may be constructed as shown in Figure 2. ΔT_{AB} represents the counter timer reading recorded at A when receiving signals from B. Assume that the residual motion of the geostationary satellite may be neglected and that delays in the two way system remain constant. The sum of the counter timer readings made at all three locations should then be equal for all of the six combinations given in Figure 2. Some of the satellite link combinations shown in Figure 2 have been implemented using the earth stations at TUG, NPL and PTB (Figure 3). Experiments have been performed where the satellite links were repeatedly switched between two different configurations every three minutes. The sum of the counter timer measurements were then calculated, a second order regression analysis was used to remove the residual motion of the geostationary satellite from the data. Examples of the residuals are shown in Figures 3a and 3b. Delay changes are clearly observed. The NPL-PTB-TUG closing error may be determined from the difference between combinations (c) and (e) given in Figure 2 since:

$$0.5(\Delta T_{AB} + \Delta T_{BC} + \Delta T_{CA}) - 0.5(\Delta T_{AC} + \Delta T_{CB} + \Delta T_{BA})$$
$$= (\Delta T_{AB} - \Delta T_{BA}) + (\Delta T_{BC} - \Delta T_{CB}) + (\Delta T_{CA} - \Delta T_{AC}) = C_{ABC} \quad (3)$$

There was very good agreement between the TUG-NPL-PTB closing errors calculated directly from the time transfer and from the satellite links.

The advantage of the satellite link method is that combinations may be formed and 'closing errors' calculated, that are not possible to determine directly from the time transfers. 'Closing errors' may be determined by calculating the differences between any of counter timer combinations given in Figure 2.

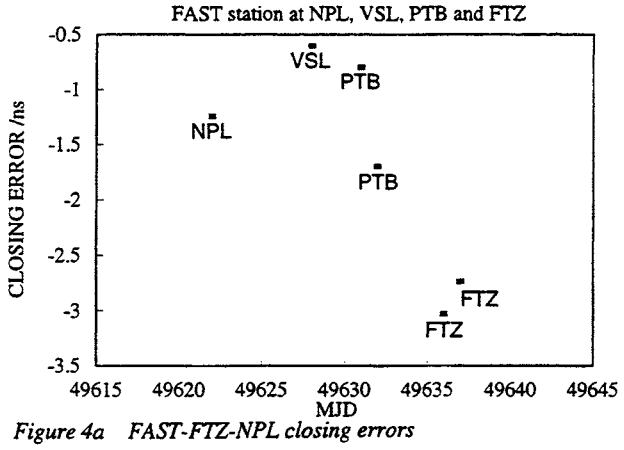


Figure 4a FAST-FTZ-NPL closing errors

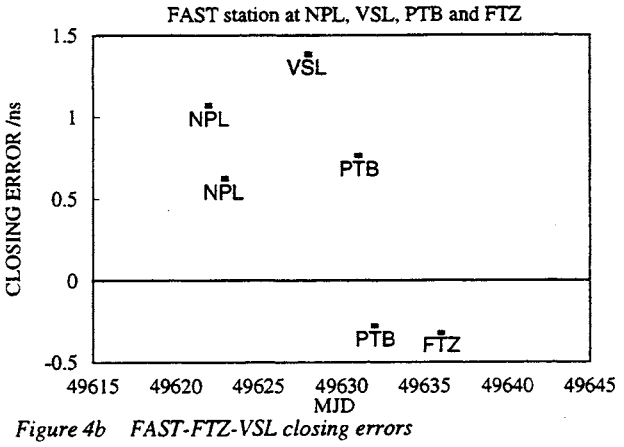


Figure 4b FAST-FTZ-VSL closing errors

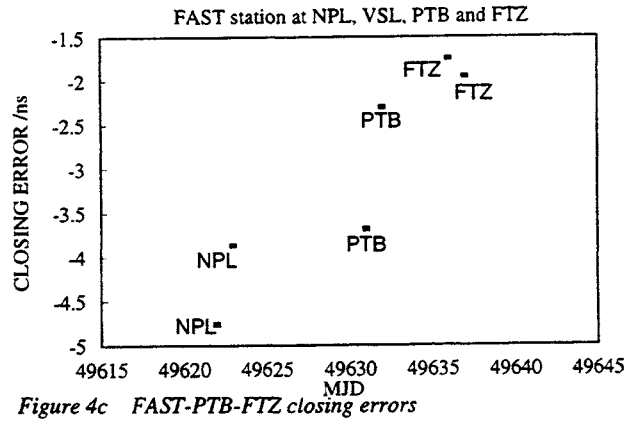


Figure 4c FAST-PTB-FTZ closing errors

4) EXPERIMENTS USING A PORTABLE EARTH STATION

4.1) Introduction

During September and October 1994 a calibration of the TWSTFT earth stations in Europe took place using a portable FlyAway satellite terminal (FAST) station [7]. The FAST station consisted of A Very Small Aperture Terminal (VSAT) terminal containing a 1.8m diameter dish which operated using an old MITREX modem [7]. The FAST station was transported in turn to each of the European laboratories. The calibration trip was used not only to calibrate the absolute delay asymmetry of the earth stations but also to calibrate the closing errors.

Closing Error	FAST-VSL VSL-TUG TUG-FAST	FAST-PTB PTB-NPL NPL-FAST	FAST-PTB PTB-TUG TUG-FAST	FAST-FTZ FTZ-VSL VSL-FAST
Mean	-0.21 ns	-1.06 ns	-1.05 ns	0.36 ns
Standard deviation	0.42 ns	1.14 ns	1.11 ns	0.58 ns

Table 2 Fast station closing errors

4.2) Closing errors determined from time transfers involving the FAST station.

Closing errors were calculated from time transfers made using the FAST station. The FAST-FTZ-NPL, FAST-FTZ-VSL and FAST-PTB-FTZ closing errors are shown in Figure (4a,b and c), calculated over the several weeks of the FAST calibration trip. The location of the FAST station at the time of each measurement is shown in the figures. The general characteristics of these closing errors are given in Table [2]. There were generally longer times between the measurement periods when compared with the regular European time transfers. The use of a common clock with both the local and the FAST station resulted in lower clock noise in the closing error determination. These results are interesting because the FAST station has been dismantled and reassembled at a different location between measurement sessions. The satellite footprint observed by the FAST station will also be different at each location. The transportation of the FAST station may introduce changes in the closing errors, especially in the FAST-PTB-FTZ closing error shown in Figure 4c.

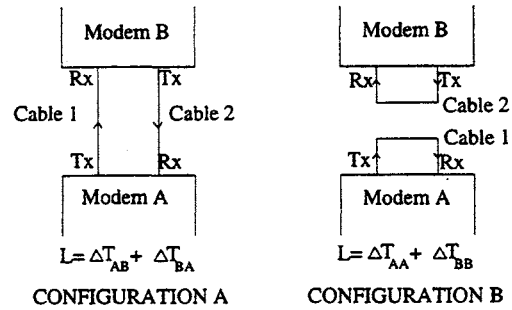


Figure 5 Modem loop configurations

4.3) Study of co-located modems and earth stations at NPL.

A closing error has been formed between two co-located modems. Using the same connecting cables, a time transfer was performed between the modems as shown in configuration A of Figure 5. A closed loop experiment was then performed as shown in configuration B. The sum of the counter timer readings were calculated as shown in Figure 5. In both configurations the sum of the instrumentation delays are equal to L. A discrepancy of 100ps was obtained between the two measurements, which was a negligible closing error.

A second experiment was performed, this time between the two satellite earth stations. A schedule was followed where both stations ranged the satellite for several minutes, followed by several minutes of time transfer experiments using a common clock. This schedule was repeated several times. Neglecting the effects of residual satellite motion, the sum of the ranging and time transfer data should be the same. A second order regression analysis was applied to the data sum to remove the effects of the residual satellite motion. Examination of the residuals revealed that 1ns delay steps occurred when changing from the ranging to time transfer configurations. Similar reverse steps were observed on returning to the ranging configuration. These results are most interesting as a closing error of 1ns was obtained when performing experiments between co-located earth stations but not when performing the similar experiments using only the modems. This suggests that the process of transmitting and receiving the time transfer signals via the earth stations and satellite is contributing towards the closing errors.

4.4) Common view experiments using side by side stations.
A possible explanation of closing error may arise from the changes in the delays within the receive chain of the earth station instrumentation. The following common view experiment was devised to study further this possible cause of the closing errors. Consider the experimental arrangement shown in Figure (6).

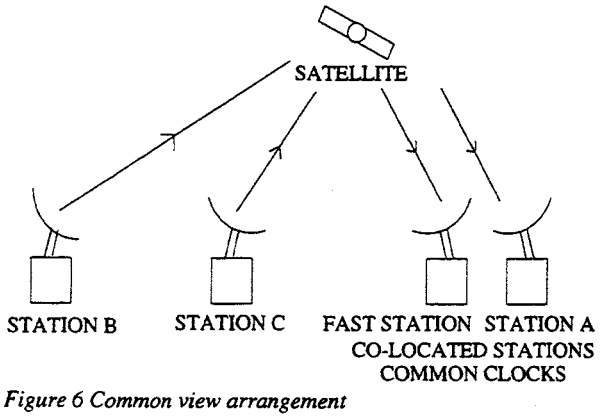


Figure 6 Common view arrangement

Both the FAST and local station (A) simultaneously receive signals from Earth station (B), while operating using a common clock. The difference between the measurements taken by the FAST station and local station (A) is then calculated ($\Delta T_{FB} - \Delta T_{AB}$). The FAST and local station (A) now switch to simultaneously receive signals from a second transmitting station (C), and the difference between the measurement made with the side by side stations is again calculated ($\Delta T_{FC} - \Delta T_{AC}$). Any discrepancies between these two differences would demonstrate that the receive delays in either the FAST station, local station, or both stations had changed on switching the transmitting station.

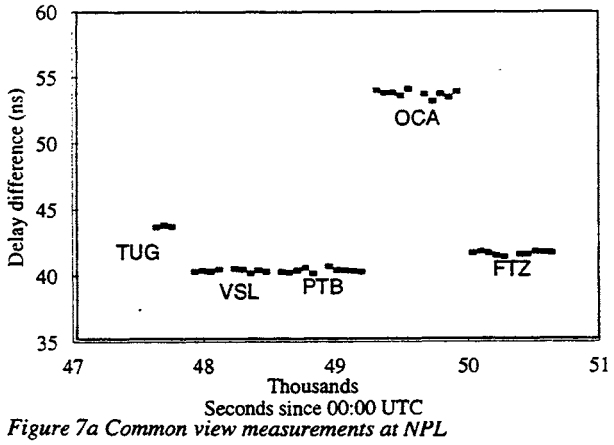


Figure 7a Common view measurements at NPL

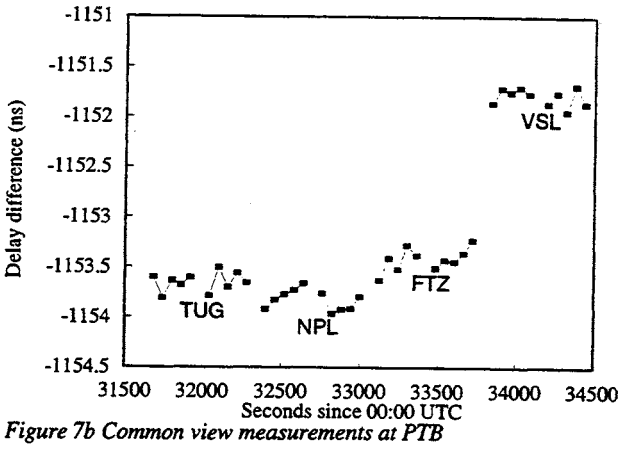


Figure 7b Common view measurements at PTB

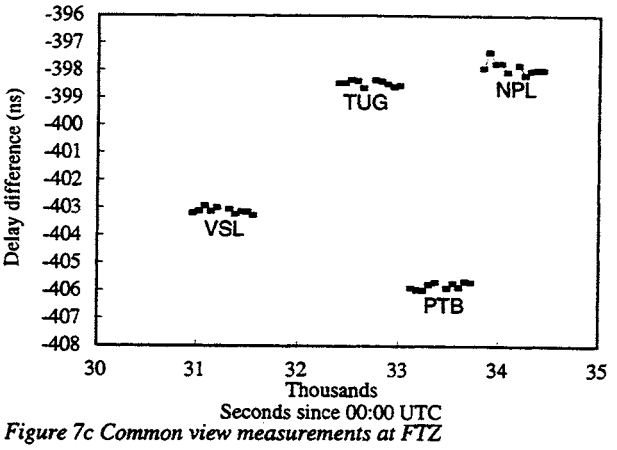


Figure 7c Common view measurements at FTZ

4.5) Characteristics of the common view experiments.
A schedule of measurements was undertaken within Europe where each remote station transmitted a signal in turn, this signal was simultaneously received by the co-located FAST and local stations, operating using common clocks. Results are shown in Figure (7a,b,c) for comparisons made at NPL, PTB, and FTZ respectively. In all examples changes of several nanoseconds were observed, clearly demonstrating that the instrumentation delay of the receive chain of the earth stations were dependent on the choice of transmitting station.

5) ORIGIN OF CLOSING ERRORS

5.1) Introduction
The presence of relatively large closing errors of several nanoseconds indicates that the theory underlying the TWSTFT method is incomplete. TWSTFT earth stations are usually calibrated using a portable earth station as a transfer standard [6]. The delay asymmetries of the two stations must be calculated, for a time transfer between stations 1 and 2. This asymmetry term is given by equation (1) of [6] as:

$$0.5\{(\tau_1^{Tx} + \tau_2^{Rx}) - (\tau_2^{Tx} + \tau_1^{Rx})\} \qquad (4)$$

In this calculation the up-link and down-link earth station instrumentation delays are assumed to be independent of the characteristics of the partner station. It has been shown here that at the nanosecond level the delay within the earth station receive chain is dependent on the characteristics of the signal being transmitted, and hence on the transmitting earth station. In this case a re-formulation of the equations underlying TWSTFT is required, which should then account for the closing errors. This re-formulation has been undertaken in [8].

It is important when considering the delays of the down-link instrumentation, to consider also which partner station is making the transmissions. In the convention used here τ_{AB}^{Rx} is the receive delay of station B when receiving signals from station A, the two subscripts give firstly the transmit and secondly the receive station. In general τ_{AB}^{Rx} and τ_{CB}^{Rx} are not equal. Similarly τ_{AB}^{Tx} and τ_{AB}^{S} represent the transmit and satellite transponder delays of a signal transmitted from station A and received by station B.

5.2) Determination of closing errors from common view measurements.
The closing errors which were determined from routine time transfers may also be calculated from the common view measurements. It is shown in [8] that the closing error C_{123} is given by:

$$C_{123} = 0.5\{(\tau_{21}^{Tx}+\tau_{21}^{S}+\tau_{21}^{Rx})-(\tau_{12}^{Tx}+\tau_{12}^{S}+\tau_{12}^{Rx})\}+0.5\{(\tau_{32}^{Tx}+\tau_{32}^{S}+\tau_{32}^{Rx})-(\tau_{23}^{Tx}+\tau_{23}^{S}+\tau_{23}^{Rx})\}+0.5\{(\tau_{13}^{Tx}+\tau_{13}^{S}+\tau_{13}^{Rx})-(\tau_{31}^{Tx}+\tau_{31}^{S}+\tau_{31}^{Rx})\} \quad (5)$$

Closing errors determined both from common view measurements and directly from time transfers are shown in Table [3]. The two methods of determining closing errors agree within measurement error.

TABLE 3

Closing Error	Calculation from common view measurements	Calculation from regular measurements
(NPL-FTZ)+(FTZ-PTB) +(PTB-NPL)	3.43 ns	2.20 ns
(VSL-FTZ)+(FTZ-PTB) +(PTB-VSL)	1.43 ns	0.54 ns

Table 3 Closing errors calculated from both common view and regular measurements.

5.3) Errors in the common view method of determining closing errors
Errors occur due to the variation of the common view measurements during a co-located measurement session. This is due both to measurement noise and earth station delay instabilities. Typical values of several hundred picoseconds are obtained for this uncertainty. A second uncertainty aries from the instrument delay changes within the FAST station during the period between the measurements being made at different locations. It must be remembered that the determinations of a closing error directly via time transfers, and by using the FAST station are measuring slightly different quantities. The direct measure of closing errors gives a 'snapshot' estimate of the closing error whereas the error determined by the FAST station essentially determines a separate component of the error, one during its operation at each earth station.

5.4) Correction term for absolute delay calibrations
The standard method of calculating the delay asymmetry between two earth stations using a portable station is given by equation (4) of [6]. It has been shown [8] that in the presence of closing errors the calculation given in [6] is in error. A correction term must be added which is given by equations (18a,b) of [8] as:

$$E = 0.5(\Delta T_{BA}^{1}-\Delta T_{AB}^{1})+0.5(\Delta T_{AF}^{1}-\Delta T_{FA}^{1})+(\Delta T_{FB}^{1}-\Delta T_{BF}^{1}) \quad (6a)$$

$$E = 0.5(\Delta T_{BA}^{2}-\Delta T_{AB}^{2})+0.5(\Delta T_{AF}^{2}-\Delta T_{FA}^{2})+(\Delta T_{FB}^{2}-\Delta T_{BF}^{2}) \quad (6b)$$

The superscripts give the location of the FAST station. This correction term is given by the FAST-A-B closing error. The two equations give this closing error determined on different occasions. Values are shown in Table 2.

5.5) Origins of the delay changes responsible for closing errors.
A possible cause of the delay changes responsible for closing errors may be due to signal distortion of the transmitted signals when it is amplified through earth station up-links operating close to saturation. The up-link filters of each of the earth stations may possess slightly different characteristics, which may result in a different spectral distribution for each signal reaching the satellite. The receive chain filters may themselves have some asymmetries in their pass-band and may not be centred exactly at 70MHz. A combination of different spectral distributions and slightly different transmit frequencies may then result in different delays in the receive chain filters when different signals are received. Extensive further investigation is however required.

6) CONCLUSIONS

Closing errors of several nanoseconds have been observed in both European and Transatlantic two-way time transfers. The value of these closing errors was found to be insensitive to changes in many of the earth station operating parameters, but strongly dependent on the modem used. Alternative methods were developed to study closing errors which were independent of clock noise. Common view measurements made using side by side stations demonstrated that the closing errors were due to the receive chain delay of the earth station instrumentation being dependent on the choice of transmitting station. Further investigation is required into the mechanism producing these delay changes.

ACKNOWLEDGEMENTS

This work has been supported under the UK National Measurement System Program for Time and Frequency, by the Austrian Academy of Sciences and the Jubilee Fund of the Austrian National Bank. The authors thank INTELSAT for the provision of cost free satellite time.

REFERENCES

[1] J A DeYoung, W J Klepczynski, A Davis, W Powell, P Hetzel, A Bauch, J A Davis, P R Pearce, C Hackman, D Kirchner, G de Jong, A Söring, P Grudler, F Baumont, H Ressler and L Veenstra, 'Preliminary results of the International two way time and frequency transfer experiment', Proc PTTI94.

[2] Delay stability of satellite earth stations as used in two-way transfer, CCIR Radiocommunications study groups, document 7-2/7-E.

[3] P Wheeler, D Chalmers, A Davis, P Koppang, A Kubik, and W Powell 'High accuracy time transfer synchronization' Proc 26th PTTI meetings, December 1994.

[4] 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. (To be published April 1995)

[5] J A Davis, P R Pearce, D Kirchner, H Ressler, P Hetzel, A Söring, G de Jong, F Baumont, P Grudler, L Veenstra, Proc EFTF94, pp.296-314, March 1994.

[6] D. Kirchner 'Two way time transfer via communication satellites' IEEE Proc. No.79, pp.983-990, 1991.

[7] W Klepczynski, J A Davis, D Kirchner, G De Jong, F Baumont, P Hetzel, A Söring, C Hackman, M Granveaud and W Lewandowski, 'Comparisons of two-way satellite time transfer and GPS time transfer during the Intelsat field trial' Proc 26th PTTI meeting, December 1994.

[8] J A Davis 'Theoretical treatment of closing errors occurring in two-way satellite time transfers' Proc EFTF95 (these proceedings).

SESSION 5

OSCILLATORS

Chairman : B. NEUBIG

9 TH EUROPEAN FREQUENCY AND TIME FORUM - BESANÇON - 8 - 9 - 10 MARCH 1995

Session Oscillators Number 5

RECENT ADVANCES IN COMPOSITE HEATED QUARTZ RESONATORS INTENDED FOR HIGH STABILITY MOBILE OCXOs

Igor V.Abramson

Omsk Piezoengineering Company Ltd.

Application of the internal heated quartz resonators (IHQR) containing in a single vacuum volume the crystal plate and electronic components to stabilize its temperature provides simultaneously short warm-up time, low power consumption and compact sizes of the OCXO. One of the most promising IHR design - composite heated resonator (CHR) is based on idea of combination of direct heating the crystal plate with the film heaters and indirect heating it with the power transistor.

Since the first presentation of the composite heated resonator by the author at 1992 IEEE AFCS the investigations have been directed to further improvement of the CHR characteristics via new ideas of building of such a resonator.

As a result of essential modification of the CHR design and technology the warm-up time, the power consumption, the temperature stability and the acceleration sensitivity have been improved considerably.

Application of the CHR novel design in 10 MHz OCXO has lead to unique combination of its properties to be very attractive for modern variety of high stability mobile systems: temperature stability $3 \cdot 10^{-8}$ over a temperature range $(-60\ +70)^{o}C$; warm-up time 15 s to $1 \cdot 10^{-7}$ and 30 s to $2 \cdot 10^{-8}$; consumption power 150 mW; phase-noise level about -160 dBc at offset over 1 kHz; acceleration sensitivity about $5 \cdot 10^{-10}$ 1/g; volume - 45 cm^3.

NONLINEAR SIMULATION OF QUARTZ CRYSTAL OSCILLATORS

N. Ratier*, R. Brendel*, T. Blin*, G. Marianneau*, 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 actual state of a computer program especially designed to accurately simulate the behaviour of quartz crystal oscillators. The program is based on the full nonlinear Barkhausen criterion which consists in deriving the oscillation condition under the form of a nonlinear complex characteristic polynomial the solution of which depends on both the amplitude and the frequency of the signal. A set of simulations using SPICE performed at run time enables the program to calculate the nonlinear behaviour of the transistor. The software initially limited to the oscillator structures collected into an oscillator library is being completed by a topological analysis module. This module is intended to automatically derive and code the oscillation condition of an oscillating circuit described by a netlist.

Introduction

Up to now, the development of high performance oscillators was, in a large extent, based on the skill of the designer. In the same time, new applications such as orbitography or atomic clock processing need more and more stable quartz crystal oscillators so that the usual design process no longer permits to reach the required performance.

Some computer aided design tools have been proposed to help the designer in his work. Because of the large Q-factor of quartz crystal oscillator circuits, the general purpose electrical simulation programs are unable to give useful informations on the signal characteristics in an acceptable computer time. Because of this limitation, dedicated softwares have been developed [1,2].

In most of these programs, the oscillation frequency is obtained by using the zero phase condition of the oscillation loop, that is the linear Barkhausen criterion. Often the start-up condition is also checked on, but nonlinear effects such as amplitude-frequency effect [3] or resonator's isochronism defect are not taken into account.

Nevertheless, by using the full Barkhausen criterion, this technique can be expanded so as to take into account all nonlinear effects. Most of them come from the active device i.e. the transistor. Then, it is essential to know its nonlinear behaviour which is obtained at run time by a set of transient analysis with increasing amplitudes performed by using an electrical simulator (SPICE) [4]. Then the transistor is represented by a nonlinear *y*-parameter two-port network whose components depend on the amplitude and, of course, on bias and temperature conditions [5]. The oscillation condition derived under the form of a nonlinear complex characteristic polynomial is eventually solved for the frequency and amplitude of the signal [6].

Oscillation condition derivation

By using simple admittance and impedance transforms like series or parallel combination or Kennelly's transforms, it can be shown that the dynamical equivalent circuit of any oscillator using a single transistor can be reduced to either one of the equivalent T- or Π- reduced forms represented in Fig. 1 [7].

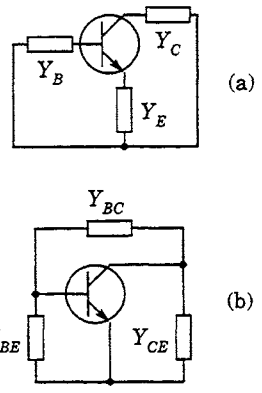


Fig. 1 : *T- and Π-reduced forms of an oscillator*
(a) : T-reduced form
(b) : Π-reduced form

The Π-reduced form is obtained from the T-reduced form by using Kennelly's transforms

$$Y_{BE} = \frac{Y_B \cdot Y_E}{\Sigma_y}, \quad Y_{CE} = \frac{Y_C \cdot Y_E}{\Sigma_y}, \quad Y_{BC} = \frac{Y_B \cdot Y_C}{\Sigma_y} \tag{1}$$

$$\Sigma_y = Y_B + Y_E + Y_C$$

All the admittances Y_B, Y_C, Y_E are in fact rational functions of the circuit components like resistances, capacitances and inductances in the Laplace's variable *s*.

The transistor itself 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. 2) [8]

$$y_n = g_n + s \cdot C_n \quad (n = 1, f, r, o) \tag{2}$$

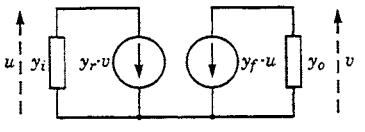


Fig. 2 : *y-parameter representation of a transistor*

From their definition, it results that y_i and y_f depend on the input voltage u while y_r and y_o depend on the output voltage v.

Figure 3a shows the dynamical equivalent Π-form of an oscillator obtained by replacing the transistor of figure 1a by its y-parameter representation (Fig. 2).

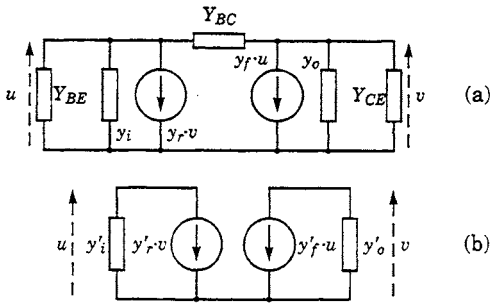


Fig. 3 : *Dynamical representation of an oscillator*
(a) Dynamical equivalent Π form
(b) Dynamical reduced form

By using simple transforms (3), the oscillator circuit is eventually reduced to the dynamical reduced form represented in Fig. 3b.

$$\begin{aligned} y'_i &= y_i + Y_{BE} + Y_{BC} & y'_r &= y_r - Y_{BC} \\ y'_f &= y_f - Y_{BC} & y'_o &= y_o + Y_{CE} + Y_{BC} \end{aligned} \tag{3}$$

Under this form, the oscillation condition can be obtained by applying Kirchhoff's law to both input and output part of the circuit

$$\begin{aligned} y'_i u + y'_r v &= 0 \\ y'_f u + y'_o v &= 0 \end{aligned} \tag{4}$$

The system (4) admits a nontrivial solution only if its determinant is nul, this gives the oscillation condition of the circuit [9]

$$y'_i y'_o - y'_r y'_f = 0 \tag{5}$$

Substituting y' in (5) by their expression (3) leads to the following form of the oscillation condition

$$\begin{aligned} &\Delta + \sigma . Y_{BC} + y_i Y_{CE} + y_o Y_{BE} + Y_{CE} Y_{BC} \\ &+ Y_{BE} Y_{BC} + Y_{BE} Y_{CE} = 0 \end{aligned} \tag{6}$$

Another equivalent form can be obtained by substituting the admittances Y_{BE}, Y_{CE} *and* Y_{BC} by their expression (1)

$$\begin{aligned} &\Delta \Sigma_y + \sigma . Y_B . Y_C + y_i . Y_C . Y_E \\ &+ y_o . Y_B . Y_E + Y_B . Y_C . Y_E = 0 \end{aligned} \tag{7}$$

with

$$\begin{aligned} \Delta &= y_i\ y_o - y_r\ y_f \\ \sigma &= y_i + y_o + y_f + y_r \end{aligned} \tag{8}$$

Do not forget that the transistor y-parameters depend on the input or output voltage u or v.

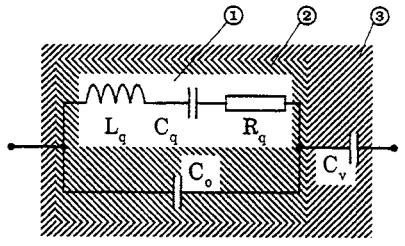
Characteristic polynomial, oscillation frequency and amplitude

Replacing the admittances in equations (6) or (7) by their expressions which are rational functions of the circuit components in the Laplace's variable s leads to the oscillator characteristic polynomial. This polynomial has the general form (9)

$$\sum_{k=0}^{K} a_k s^k = 0 \tag{9}$$

Each coefficient of this polynomial is expressed as a function of the component value of the circuit. The degree and the complexity of the characteristic polynomial are, of course, strongly depending on the component number. The full development of the polynomial coefficients often results in intricate expressions hardly manageable without error. Fortunately, by using computer aided symbolic calculation it has been possible to partially automate the derivation and coding in high level language of the polynomial coefficients so as to minimize the risk of error. Usually, the analysis of an oscillator structure starts with the expression of the circuit admittances, in a second step some admittances are combined and expressed under the form of a rational function whose the numerator and denominator are polynomials. The coefficients of these polynomials which are function of the circuit parameters are stored and are given a symbolic name so as to get a simpler rational function of same degree.

This new symbolic form is used in the subsequent admittance combination steps until the T- or Π-reduced form of the circuit is obtained. An example of this encapsulation process is demonstrated in Fig. 4. The last step of the derivation consists in substituting the reduced symbolic admittances in either one of the oscillation condition (6) or (7) so as to obtain the characteristic polynomial (9). At each step of the reduction process, the symbolic coefficients are expressed as a function of the coefficients defined in the previous step and coded in high level language so that they can be integrated into a subroutine almost without manual intervention.



$$Y_1 = \frac{a_{11} p}{b_{10} + b_{11} p + b_{12} p^2}, \quad a_{11} = C_q \quad b_{10} = 1 \quad b_{11} = C_q R_q \quad b_{12} = L_q C_q$$

$$Y_2 = \frac{a_{21} p + a_{22} p^2 + a_{23} p^3}{b_{10} + b_{11} p + b_{12} p^2}, \quad a_{21} = a_{11} + b_{10} C_o, \quad a_{22} = b_{11} C_o, \quad a_{23} = b_{12} C_o$$

$$Y_3 = \frac{a_{31} p + a_{32} p^2 + a_{33} p^3}{b_{30} + b_{31} p + b_{32} p^2}, \quad a_{31} = a_{21} C_v \quad b_{30} = b_{10} C_v + a_{21}, \quad a_{32} = a_{22} C_v \quad b_{31} = b_{11} C_v + a_{22}, \quad a_{33} = a_{23} C_v \quad b_{32} = b_{12} C_v + a_{23}$$

Fig. 4 : *Example of encapsulation process showing three reduction steps*

When the characteristic polynomial coefficients are defined, it is possible to obtain the steady state frequency and amplitude of the oscillator. To this end, the Laplace's variable s is changed into the harmonic variable $j\omega$, this splits equation (9) into two equations, one for the real part, the other one for the imaginary part (Eq. 10)

$$\sum_{m=0}^{M} a_m(u).\omega^m \quad , \quad \sum_{n=0}^{N} \beta_n(u).\omega^n = 0 \qquad (10)$$

Because some of the polynomial coefficients α_m and β_n in Eq. 10 depends on the amplitude through the nonlinearities of the transistor admittances, the problem amounts to find the amplitude u_0 and the frequency ω_0 satisfying both equations (10). Moreover this technique permits to take into account other nonlinear effects such as frequency-amplitude effect of the resonator by describing the resonant frequency under a polynomial in the drive level variable P

$$\omega_q(P) = \omega_q(0)\,(1 + \alpha_1 P + \alpha_2 P^2) \qquad (11)$$

where $\omega_q(0)$ represents the resonator frequency at low drive level, α_1 and α_2 are the coefficients required by the program to describe the frequency-amplitude effect.

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 behavior 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.

Problem arising with the automation of the oscillator analysis

Although the oscillation condition is derived with the help of a symbolic calculation program, it still needs a large part of manual intervention so as to write down the complete module describing the oscillator behaviour. Furthermore, the calculation of the resonator frequency-amplitude effect needs a special analysis which cannot be easily automated from the T- or Π-reduced form because some transforms performed on the circuit, during the encapsulation process result in a scattering of the resonator admittance among the three remaining ones (Fig. 1). The complete automation of the oscillator analysis starting from the netlist up to the generation of the code describing the oscillation condition and the resonator drive level needs to restate the circuit analysis. First, the resonator has to be kept outside the encapsulation process so as to permit the drive level calculation, second, the transistor itself must be kept apart because its nonlinear behavior is expressed as functions of the input or output voltage u or v (Fig. 2).

Reduced form and oscillation condition

So as to meet the requirements described in the previous section, the oscillator circuit should be reduced under the form shown in Fig. 5a.

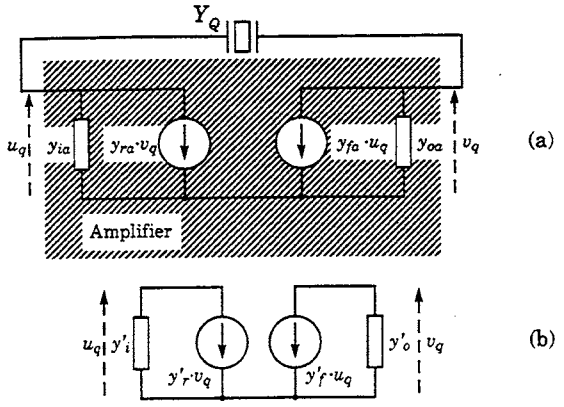


Fig. 5 : *Reduced form of an oscillator*

The resonator always appears connected between the input and the output of an amplifier represented by a y-parameter tow-port circuit. The admittance of this two-port circuit consists in the nonlinear transistor y-parameters to which the other circuit admittances have been associated as described in the next section.

Under this form, the oscillation condition can be derived by reducing the circuit in Fig. 5a to the one shown in Fig. 5b similar to Fig. 3b using the transforms (12)

$$\begin{aligned} y'_i &= y_{ia} + Y_Q \quad & y'_r &= y_{ra} - Y_Q \\ y'_f &= y_{fa} - Y_Q \quad & y'_o &= y_{oa} + Y_Q \end{aligned} \qquad (12)$$

So that the oscillation condition (5) still applies. Substituting y' in (5) by their expression (12) gives a new form of the oscillation condition

$$\Delta_a + \sigma_a \cdot Y_Q = 0 \qquad (13)$$

where

$$\begin{aligned} \Delta_a &= y_{ia}\, y_{oa} - y_{fa}\, y_{ra} \\ \sigma_a &= y_{ia} + y_{oa} + y_{fa} + y_{ra} \end{aligned} \qquad (14)$$

Because the resonator admittance is separated from the amplifier circuit, the drive level is quite simple to express.

Elementary transformation

The amplifier can be reduced to the form represented in Fig. 5 by using a limited number of elementary transformations on the circuit admittances like series or parallel combinations, Kennelly's transforms, etc. When a transformation involves the transistor itself, the circuit is reduced to a new y-parameter two-port equivalent circuit whose components are expressed in terms of the initial circuit components. In addition, when new input and output variables are introduced by the transformation, they are expressed in terms of the former ones so as to keep track of the transistor input and output variables.

As an example, Fig. 6 shows one of the elementary transformation.

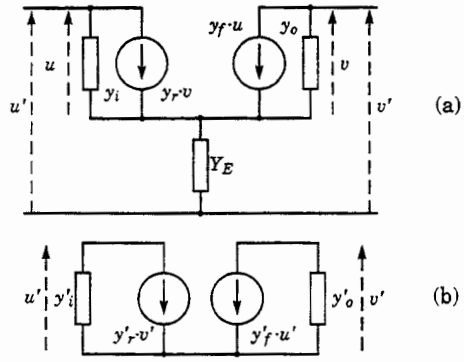


Fig. 6 : *Elementary transformation of the circuit*

Equations (15) and (16) represents the associated relationships

$$y'_i = \frac{Y_E\, y_i + \Delta}{Y_E + \sigma} \qquad y'_r = \frac{Y_E\, y_r - \Delta}{Y_E + \sigma}$$
$$y'_f = \frac{Y_E\, y_f - \Delta}{Y_E + \sigma} \qquad y'_o = \frac{Y_E\, y_o + \Delta}{Y_E + \sigma} \tag{15}$$

$$u' = \frac{y_i + y_f + Y_E}{Y_E}.u + \frac{y_r + y_o}{Y_E}.v$$
$$v' = \frac{y_i + y_f}{Y_E}.u + \frac{y_r + y_o + Y_E}{Y_E}.v \tag{16}$$

Δ and σ have the same meaning as in Eq. (8).

Figure 3 and Eq. (3) represent another example of elementary transformation where the input and output variables do not change.

In the case where the emitter is not the common lead it is necessary to transform the common emitter representation in the common base or common collector one as shown in Fig. 7.

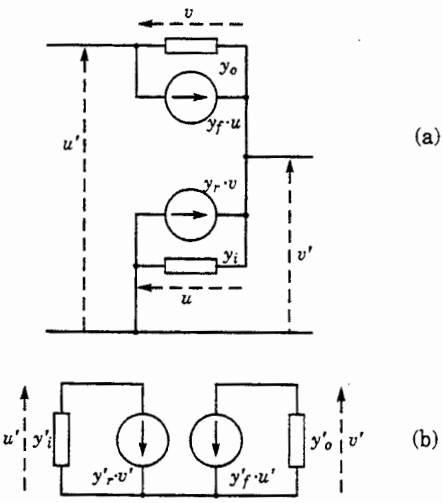


Fig. 7 : *Common emittor to common base transformation*

The associated relationships are shown in Eqs. (17) and (18)

$$y'_i = y_o \qquad y'_r = -(y_f + y_o)$$
$$y'_f = -(y_r + y_o) \qquad y'_o = y_i + y_f + y_r + y_o \tag{17}$$

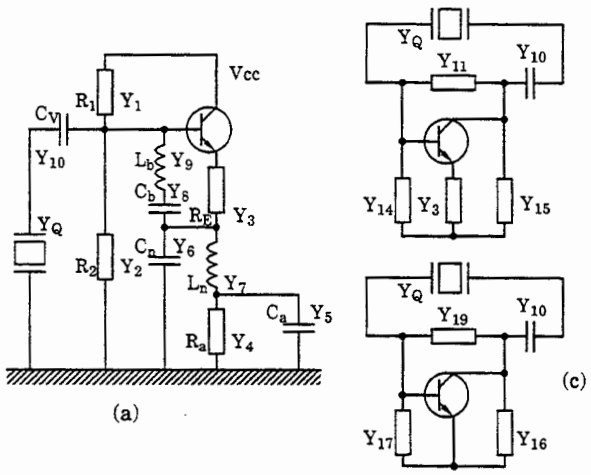
$$u' = v - u$$
$$v' = -u \tag{18}$$

Each time an elementary transformation occurs, the admittances of the new reduced two-port circuit are expressed under the form of a symbolic rational function in the variable s by using the encapsulation process previously described. The process is repeated until the final reduced form shown in Fig. 5 is obtained so that the oscillation condition (13) can be applied.

Code generation

In the present state, the oscillator circuit is described in an input file under the form of a SPICE [10] netlist. The main program parses the input file, identifies and performs the transformations on the admittances until all possible associations are done. Each time an association is performed, the FORTRAN code of the rational function coefficients is automatically generated and stored and the coefficients are given a new symbolic name [11].

Then the elementary transforms involving the transistor are performed in a similar manner, the reduced circuit parameter coefficients are generated, coded and stored as well as the input/output variable expression. At the conclusion of the process, the final reduced form is obtained and all necessary relations are coded in a file. Figure 8 shows an example of the reduction and coding process.



$$Y_3 = \frac{1}{A_7} \qquad Y_{10} = A_8 \cdot s$$
$$Y_{11} = \frac{A_{11}}{A_{12}} \qquad Y_{14} = \frac{A_9 \cdot s + 1}{A_{10} \cdot s^2 + 1}$$

$$Y_{15} = \frac{A_3 \cdot s^3 + A_2 \cdot s^2 + A_1 \cdot s + 1}{A_6 \cdot s^2 + A_5 \cdot s + A_4}$$

$$Y_{16} = \frac{A_{18} \cdot s^5 + A_{17} \cdot s^4 + A_{16} \cdot s^3 + A_{15} \cdot s^2 + A_{14} \cdot s + 1}{A_{24} \cdot s^5 + A_{23} \cdot s^4 + A_{22} \cdot s^3 + A_{21} \cdot s^2 + A_{20} \cdot s + A_{19}}$$

$$Y_{17} = \frac{A_{39} \cdot s^3 + A_{38} \cdot s^2 + A_{37} \cdot s}{A_{24} \cdot s^5 + A_{23} \cdot s^4 + A_{22} \cdot s^3 + A_{21} \cdot s^2 + A_{20} \cdot s + A_{19}}$$

$$Y_{19} = \frac{A_{30} \cdot s^5 + A_{29} \cdot s^4 + A_{28} \cdot s^3 + A_{27} \cdot s^2 + A_{26} \cdot s + A_{25}}{A_{36} \cdot s^5 + A_{35} \cdot s^4 + A_{34} \cdot s^3 + A_{33} \cdot s^2 + A_{32} \cdot s + A_{31}}$$

```
A1 = Ra*Ca+Cn*Ra
A2 = Cn*Ln
A3 = Cn*Ln*Ra*Ca
A4 = Ra
A5 = Ln
A6 = Ln*Ra*Ca
A7 = Re
A8 = Cv
A9 = Cb
A10 = Cb*Lb
A11 = R1+R2
A12 = R1*R2

A13 = A8
A14 = A1
A15 = A2+A10
A16 = A3+A10*A1
A17 = A2*A10
A18 = A3*A10
A19 = A4+A7
A20 = A7*A9*A4+A5+A7*A1
A21 = A7*A9*A5+A6+A4*A10+A7*A2+A10*A7
A22 = A7*A9*A6+A10*A5+A7*A3+A7*A1*A10
A23 = A6*A10+A7*A2*A10
A24 = A7*A3*A10
A25 = A11*A4+A11*A7
A26 = A12*A9*A7+A11*A7*A9*A4+A11*A5+A11*A7*A1
A27 = A12*A9*A7*A1+A11*A7*A9*A5+A11*A6+A11*A4*A10+A11*A7*A2+A11*A7*A10
A28 = A12*A9*A7*A2+A11*A7*A9*A6+A11*A5*A10+A11*A7*A3+A11*A7*A1*A10
A29 = A12*A9*A7*A3+A11*A6*A10+A11*A7*A2*A10
A30 = A11*A7*A3*A10
A31 = A12*A4+A12*A7
A32 = A12*A7*A9*A4+A12*A5+A12*A7*A1
A33 = A12*A7*A9*A5+A12*A6+A12*A4*A10+A12*A7*A2+A12*A7*A10
A34 = A12*A7*A9*A6+A12*A5*A10+A12*A7*A3+A12*A7*A1*A10
A35 = A12*A6*A10+A12*A7*A2*A10
A36 = A12*A7*A3*A10
A37 = A9*A4
A38 = A9*A5
A39 = A9*A6
```

Fig. 8 : *An example of circuit reduction and coding process*

At this stage, the oscillation condition (13) can be applied since all terms are expressed under the form of rational functions of known degree whose coefficients are themselves expressed as a function of the circuit parameters.

The FORTRAN code of the characteristic polynomial coefficients is then edited as well as the resonator drive level calculation [12].

All the FORTRAN statements coded during the reduction process are collected into a file. The subroutine is then generated almost without any human intervention thus minimizing the risk of error.

When the code generation is completed, it finds a place in the oscillator library and can be called by the root-finder program so as to calculate the oscillation characteristics : amplitude, frequency, resonator drive level, sensitivity, etc. as described in ref. [6].

Conclusion

The development of a topological analysis program permits to automatically generate a FORTRAN module describing the oscillator behaviour starting from a SPICE netlist. The oscillator module consists in a complex characteristic polynomial coefficient the solution of which gives the oscillation frequency and amplitude.

Because the polynomial coefficients are literal expression of the circuit parameter, it is possible to accurately analyze the influence of anyone of them on the oscillator behaviour.

Some additional effort has still to be spent so as the topological analysis module rises to its operating level.

Up to now only the steady state oscillation features are calculated by the program. Nevertheless, the characteristic polynomial can also be used to describe the transient oscillator behaviour. In this case, the polynomial is used in the time domain as a nonlinear differential equation which can be solved for the transient analysis of the oscillation amplitude and frequency. In addition, this differential equation should be used for the analysis of noise mechanisms in the oscillation loop.

Acknowledgements

This work has been supported by CNES and DGA under contracts : # 962/CNES/91/1476/00, # 962/CNES/94/1230/00.

References

[1] B. Parzen, A. Ballato, "Design of crystal and other harmonic oscillators", John Wiley & Sons (1983).

[2] A. Benjaminson, "The design and analysis of VHF/UHF crystal oscillator", Proc. 41st AFCS (1987).

[3] J.-J. Gagnepain, R. Besson, "Nonlinear effect in piezoelectric quartz crystals", Physical Acoustics, vol. XI, Academic Press (1975).

[4] L.W. Nagel, "SPICE 2 : A computer program to simulate semiconductor circuits", Memorandum ERL-M520, Univ. of California, Berkeley (1975).

[5] M.E. Frerking, "Crystal oscillator design and temperature compensation", Van Nostrand Reinhold Co (1978).

[6] R. Brendel, F. Djian, E. Robert, "High precision nonlinear computer modelling techniques for quartz crystal oscillators", Proc. 45th AFCS (1991).

[7] R. Brendel, T. Blin, G. Marianneau, "Modélisation non linéaire des oscillateurs ultrastables pour applications aérospatiales",
Final report contract 962/CNES/91/1476/00 (1994).

[8] 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 (1974).

[9] R. Brendel, G. Marianneau, T. Blin, M. Brunet, "Computer aided design of quartz crystal oscillators", Proc. of 1994 IEEE IFCS (1994).

[10] P. Tuinenga, "SPICE - A guide to circuit simulation and analysis using PSpice", 2nd edition, Prentice Hall, 1992.

[11] B. Char, K. Geddes, G. Gonnet, B. Leong, M. Monagan, S. Watt, "Maple V Language Reference Manual", Springer Verlag, 1991.

[12] G. Gomez, "Macrofort : a Fortran code generator in Maple", INRIA Report 119 (1990).

NEW DEVELOPMENTS IN CAPACITANCE SWITCH ARRAY FOR DIGITAL COMMAND OF CRYSTAL OSCILLATORS

E.Girardet*, J.Marguier*, G.Marianneau, W.Daniau****

** AR Electronique*
Valparc - 14, rue de Franche Comté - F25048 Besançon Cedex .
*** Laboratoire de Physique et Metrologie des Oscillateurs, C.N.R.S.*
32, av. de l'Observatoire - F25044 Besançon Cedex .

This work was supported by the French Ministry of Research under grant number : 92S0715 and 92S0716 .

ABSTRACT

On classical DTCXOs, the digital temperature compensation is applied through a Digital-Analogic Converter and a varactor diode . The performances in terms of stability, the power and supply voltage requirements are particularly limited by these component characteristics .

An interesting alternative is the use of a capacitance switch array (CSA) which is directly digitally addressable. The faisability has been prooved in precedent studies . One of the major inconvenient was the need of a unitary adjustment of the capacitance array, which is not affordable for industrial applications .

A.R Electronique and the LPMO, with the support of the French Ministry of Research, have realized a pre-industrial study of a CSA for piezoelectric applications . This work is based on a theoretical simulation and the interactive optimization of both capacitance array and switching unit . The electrical characterization was done by using different oscillators as transducers .

Capacitance ranges of 15 pF to 100 pF, with unitary capacitance steps of 0.1 pF has been obtained . The experimental arrays are tightly connected with the expected calculations . The average dispersion values of about 5 % is compatible with industrial processes .

This capacitance switch array, which is to be used with a micro-controller, opens large application fields for the digital command of crystal oscillator's frequency .

I - TECHNICAL CONTEXT

On classical crystal oscillators (XO, TCXO, VCXO, OCXO), the pullability of the nominal frequency is obtained by use of capacitor-trimmers or varactor diodes .
On temperature compensated oscillators (TCXO or DTCXO), the frequency versus temperature drift, the ageing and the quartz crystal calibration are compensated by adjustment of the voltage applied on the varactor diode, wich is mounted in line with the crystal (Fig 1).
It is possible to summerize the advantages, disadvantages and limitations as follows :
On TCXOs, the analogic compensation by thermistor network gives a continuous frequency versus temperature curve, without any steps .On the other hand, the thermal network cannot compensate frequency vs temperature characteristics that are not typical 3rd order curves . For wide temperature ranges (-40 to +100°C), it is difficult to reach frequency stabilities better than 1.10-6 .
In addition, both the dynamic voltage range of the compensation network and the capacitance range of the varactor diode decrease with low power voltages (5V or lower). It limits consequently the frequency compensation capability .
On DTCXOs, the digital compensation allows to compensate any kind of frequency vs temperature curves . The compensation principle is well adapted to an industrial process . As disadvantages, the frequency compensation is now a compensation by steps, depending on the Digital/Analogic Converter (DAC) dimension . As for an analogic TCXO, the dynamic of the varactor diode decreases with lower voltage range, reducing the compensation capability as well . This phenomena is more critical for ultra low power voltages (down to 3V) (future needs) .

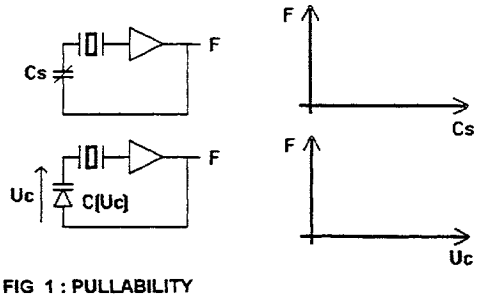


FIG 1 : PULLABILITY

For both TCXO and DTCXO, the limited varactor's characteristic dC/dVr (capacity versus reverse voltage) requires the adaptation of both quartz resonator and oscillator parameters, in order to provide a frequency sensitivity sufficient for the frequency drift versus temperature and ageing compensation . This sensitivity can be expressed versus the serial capacitance (ppm / pF) or versus the varactor's reverse voltage (ppm / V) . This is achieved by the use of high sensitivity resonators (low inductance value L1, low Q), working at low serial capacitance value (~ 10 pF) . The varactors diodes are often abrupt or hyperabrupt diodes, combined if necessary with serial inductors . Consequently, the Q factor of the oscillator is degraded, wich is not optimal for the frequency stability .

The ***capacitance switch array (CSA)*** principle consists of switching elementary capacitances mounted in parallel, through a binary command word . This fully integrated component is used instead the usual compensation system . Both varactor diode and D.A.Converter are eliminated, reducing the power requirement . The CSA is directly driven by the micro processor (Fig 2) .

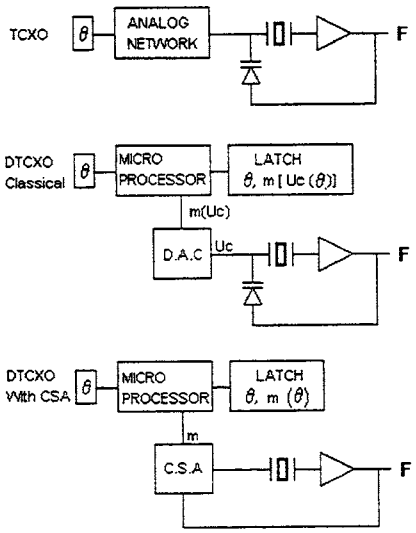


FIG 2 : FUNCTIONNAL SCHEMES OF TCXO

The original idea has been presented in 1983 by Takehiko Uno and Yoshio Shimoda [1]. Then, a study has been initiated by Daniau et al [2,3], in order to demonstrate the practical interest of such a DTCXO using a CSA. The faisability has been shown and the following characteristics have been obtained :

- Frequency stability vs temperature range (0 to 80°C) : +/- 2.10-7
- Power consumption : less than 25mW
- Supply voltage : +5V .

This experimental first approach has comfirmed the high interest of such a component, wich provides a wide capacity range independently with the supply voltage, associated with low consumption . These advantages are especially positive for severe environmental conditions (wide temperature range, low supply voltage, low consumption) .

A new study was launched with the support of the ***French Ministry of Research***, with ***L.P.M.O*** and ***A.R Electronique*** as partners . The main goal of this study constists first to improve the electrical characteristics, combined with a pre-industrial approach . This work is based on simulation calculations, experimental tests, and successive runs of realizations . The second goal is the developpment and the characterisation of DTCXO using the CSA.

The following paragraphs present the fundamentals of the CSA, the technology of capacitance array design, the switching technic, the main electrical and technical requirements that have to be taken into account, the CSA achieved characteristics, and the application on DTCXO .

II - FUNDAMENTALS

The capacitance switch array (CSA) is composed of parallel capacitors, each placed in line with a digitally controlled analog switch, as shown on figure 3 .

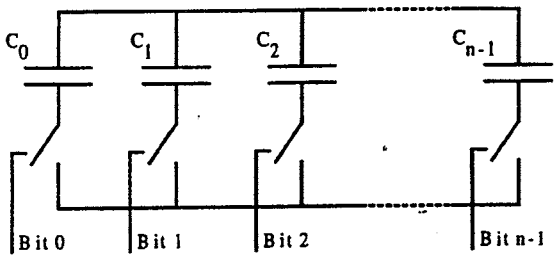


FIG 3 : CAPACITANCE SWITCH ARRAY (CSA)

Ideal component :

If the *N* capacitor's values are binary ponderated, with *Co* elementary capacitance value of the array, and if the switches are perfect commutators, then the application of a binary command word *n* will give a practical capacitance of *nCo* .

- The minimum capacitance value is 0
- The maximum capacitance value is NCo
- The absolute capacitance deviation is DC = Cmax - Cmin = NCo
- The elementary capacitance step (binary word i to i + 1) is dC = Co (constant)
- This component is characterized by its capacitance versus command word curve, given on figure 4 .

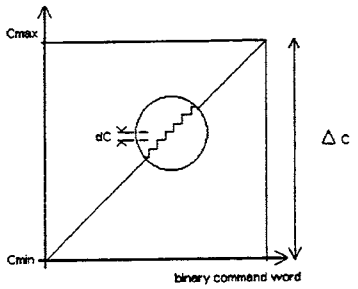


FIG 4 : CAPACITANCE vs COMMAND WORD (IDEAL CSA)

Practical component :

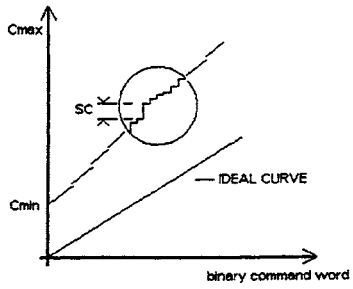
Let's consider a manufactured CSA .It is now necessary to take into account the practical characteristics of the two tightly connected components : capacitors array and switches . These characteristics are of course not ideal :

- The binary ponderation of the elementary Ci values is affected by a dispersion .
- In addition, parasitic capacitors and resistors complicate the ideal scheme of the switch .

All these parameters induce degradations on the capacitance versus binary command word characteristic, as shown on figure 5 .

As consequences, the minimum capacitance value Cmin is increased . In addition the elementary steps values dC are more important, especially for high order binary command words .

FIG 5 : CAP. vs COMMAND WORD (EXPERIMENTAL CSA)



Electrical requirements for temperature compensation applications :

The CSA device is used in order to compensate both the frequency drift versus temperature and the ageing, in a given temperature range, within a given tolerance .

The target values are as follows :

- overall frequency stability to be compensated :
  +/- 40.10-6 (temp. range -50 to +100°C)
- expected compensation :
  +/- 1.10-7
- minimal capacitance value (binary command word=0):
  as low as possible

full capacitance range DC = Cmax - Cmin :
  > 100 pF

- maximum step value dC :
  < 30 fF (frequency steps < 5.10-8)
- number of digits N = DC / dC :
  12 bits / 4000 command words .

Characterisation tool :

Because of the low capacitance steps to be measured, the use of a standard capacimeter is difficult . (steps of a few tenth of fF) .

The measurement system consists of a capacitance to frequency converter, by use of a high stability OCXO .

If the frequency versus serial capacitance characteristic F(C) of the OCXO is recorded, the oscillator's frequency measurement allows to compute the CSA capacitance value . The scheme of the measurment system is shown on figure 6 .

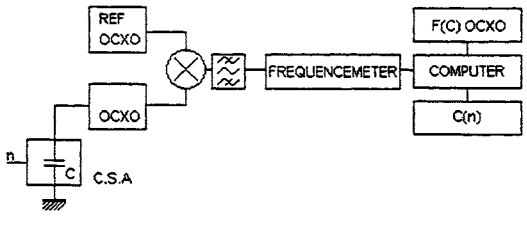


FIG 6 : CAPACITANCE MEASUREMENT SYSTEM

The optimisation of the CSA needs simultaneously an improvment of capacitor array manufacturing, a carefull switch selection, and as it will be demonstrated, an adaptation beetween both constitutive components .

III - TECHNOLOGY OF CAPACITANCE ARRAY

Some preliminary experimentations have shown the practical impossibility to use discrete capacitors for a CSA :The lower capacitance values, wich define the elementary steps are not available (some tenth of fF) . In addition, the computed capacitance Ci are non standard values . More, the tolerances on commercial components are not compatible with the reliability targets .

It has been decided to realize all the capacitors simultaneously, with a thin film technology.

An adapted design allows to realize on the same substrate all the elementary capacitors . The dimensions are compatible for a small size integration . The dispersion on the capacitance values are now correlated, wich is very influent for the global performance . In addition, this technology is compatible with mass production techniques, wich gives access to low cost products .

As shown on figure 7, all the capacitors are made with three layers . The first aluminium layer (150 nm thick) contains the individual armatures of the capacitors and the connections to the swiches . The second layer is a PECVD silicon nitride layer with a thickness of 400 nm used as dielectric of the capacitors . The third layer is another 150 nm aluminium layer wich consists in the common armature of the capacitors .

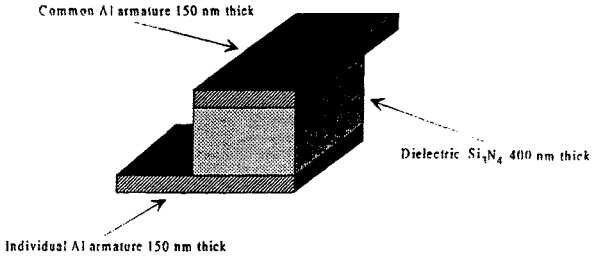


FIG 7 : CAPACITOR'S STRUCTURE

The choice of the substrate is very critical : it is necessary to minimize the parasitic capacitances . After different successive tests,a dielectric substrate has been choosen to reduce parasitic capacitance effects . (a two inches glass wafer with good polishing).

In order to qualify the manufacturing process, a first capacitor array has been designed . It consists of 12 binary pondareted values, from some .1 fF to 75 pF

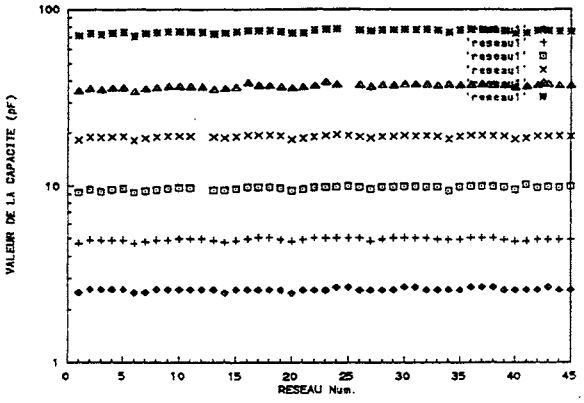
Two wafers have been processed, containing each 80 CSA.

The calculation of the capacitance values has been verified . Then the capacitance values dispersions have been characterized in two ways :

- Capacitance dispersion from one capacitance array to another, on the same wafer
- Capacitance dispersion from two wafers .

Fig 8 shows the high level of correlation on the capacitors values dispersions on a same CSA (< 5%). As we will see later, this point is very important for the global characteristics of the complete CSA .

FIG 8 : CAPACITOR'S VALUES DISPERSION



IV - SWITCH SELECTION

The available switches are of course not perfect . They can be modeled as shown in figure 9 . *Ron* is the switch resistance value (state closed) . *Coff* is the parasitic capacitance of the switch . Both parameters greatly influence the resulting response of the CSA. A simulation method has been developped, wich allows to investigate the sensitivity of these parameters, and helps for the definition of choice criteria . A comparison between experimental data the simulation results has validated this approach .

When connected with a binary ponderated capacitance array, the equivalent capacitance of the complete CSA can be written as follows :

(Ai : logic state of the switch i : Ai=0 / open, Ai=1 / closed)

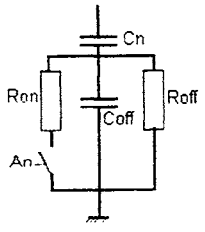


FIG 9 : SWITCH EQUIVALENT SCHEME

Equivalent capacitance C :

$$C = \sum_{i=0}^{i=n} \left(A_i C_i \frac{(1+R_{on}^2 C_{off}^2 \omega^2)(1+R_{on}^2 C_{off} \omega^2 (C_i+C_{off}))}{R_{on}^2 C_{off} \omega^2 + (1+R_{on}^2 C_{off} \omega^2 (C_i+C_{off}))^2} \right) + \sum (A_i \frac{C_i C_{off}}{C_i + C_{off}})$$

Maximal capacitance value (b.c.w = 2^n-1) :

$$C_{max} = \sum_{i=0}^{i=n} \left(A_i C_i \frac{(1+R_{on}^2 C_{off}^2 \omega^2)(1+R_{on}^2 C_{off} \omega^2 (C_i+C_{off}))}{R_{on}^2 C_{off}^2 \omega^2 + (1+R_{on}^2 C_{off} \omega^2 (C_i+C_{off}))^2} \right)$$

Parasitic capacitance Cmin (binary command word = 0) :

$$C_{min} = \sum_{i=0}^{i=n} \left(\frac{C_i C_{off}}{C_i + C_{off}} \right)$$

the switch parameters effects on the resulting response of the CSA are first a higher minimal value Cmin, and secondly big capacitance steps, especially when binary power values are switched (Fig 5) .

The *Coff* value, wich cannot be as small as desired, will determine the minimum capacitance value of the array, for low power command words (lsb) . For high power command words (msb) (high capacitance values), the *Ron* parameter as well as the working frequency are to be taken into account (as small as possible) .

A large investigation has been realized, on different switch transistors of different technologies and manufacturers . They have been compared regarding their electrical parameters first, but also their technology, their cost and availability . The DMOS transistors that have been selected are of course not perfect . It's now necessary to redefine the associated capacitance array in order to optimize the global resulting response .

V- CSA OPTIMIZATION

The simulation tool that has been developped has demonstrated, with a good correlation comparing with experimental results, the high influence of the switch parameters when connected to a binary ponderated capacitance array (Fig 10) .

This same simulation programm can be used in order to adapt the elementary capacitance values of the array to the switch parameters and to minimize the capacitance steps effects .

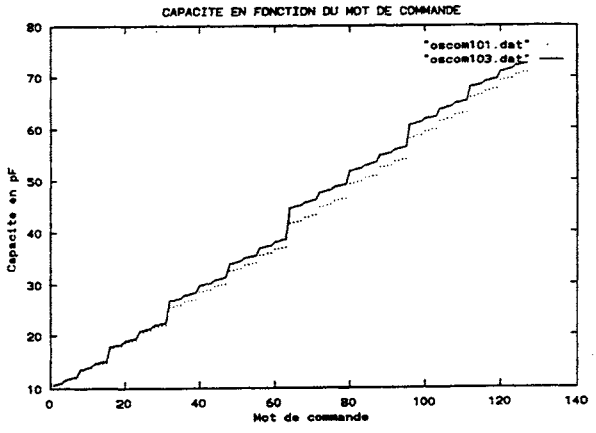


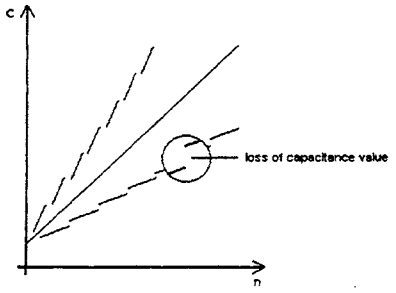
FIG 10 : CSA (EXPERIMENTAL / SIMULATION)

A first experimental approach has been done by Daniau et Al. by adjusting the capacitance values with a laser etching process . This method involves an important limitation of the global capacitance range . But the most important problem is that this optimization is realized individually, for each CSA . This heavy concept is not adapted for mass production .

A second way consists in redefining the elementary capacitance values of the array . It is theoretically possible to optimize them, and then the elementary capacitance steps will be minimum and constant .

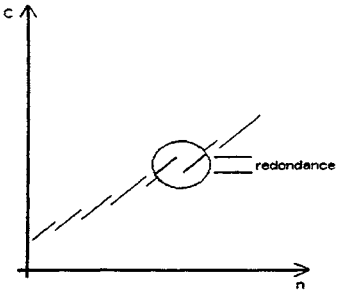
In practice, this optimization will be degraded because of the dispersion on the capacitors values and the parasitic parameters of the switch . In particular, this allways will induce big steps at binary powers . (Fig 11).

FIG 11 : CAPACITANCE vs COMMAND WORD : INFLUENCE OF DISPERSION



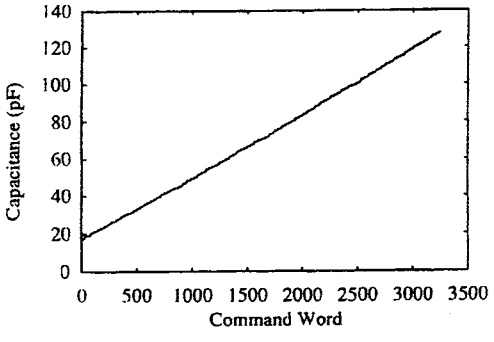
An improvement of this result consists of adjusting the elementary capacitor values so that the sign of these steps is allways negative, taking into account the maximal dispersion of both switch parasitic parameters and elementary capacitances of the array (Fig 12) .

FIG 12 : OPTIMIZED CHARACTERISTIC (BEFORE CLASSIFICATION)



In this way, any loss of capacitance value is avoided . Since a digital treatment is performed behind the CSA, it will be possible to make a classification of the values in a table, in order to obtain a quite linear response (Fig 13) .

FIG 13 : CAPACITANCE vs COMMAND WORD (AFTER CLASSIFICATION)



FINAL DEVICE :

Using the new fonctionnal scheme (Fig 14), it was possible to realize a twelve bit device without increasing the total variation of capacitance .

It provides a large capacitance range, from Cmin =20 pF to Cmax = 130 pF, with maximum capacitance steps lower than 100 fF .

The device has been fully integrated in a 20 pins dual in line package, wich contains the CSA and the D-MOS switches (photograph Fig 15) .

FIG 14 : NEW FUNCTIONNAL SCHEME (12 BIT)

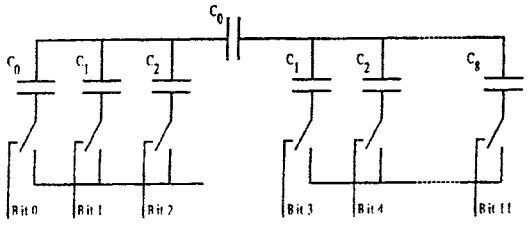
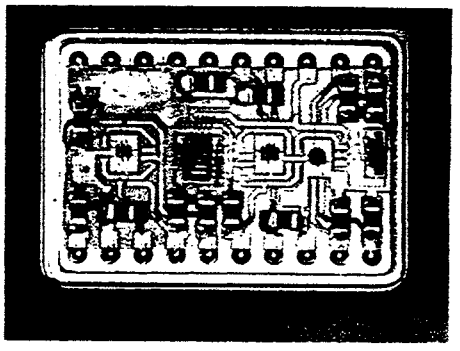


FIG 15 : VIEW OF THE INTEGRATED CSA



VI - APPLICATION FOR DIGITAL COMPENSATION

The final CSA device has been specially designed for digital temperature compensation on crystal oscillators . The wide capacitance range, combined with elementary steps lower than 100 fF, allows the use for frequency compensation on wide temperature ranges (-50 to +120 °C) . It is important to notice that these performances are quite independant of the power supply . This is particularly interesting for low supply voltage applications, where capacitance ranges of classical varactor diodes are limited (Fig 16) .

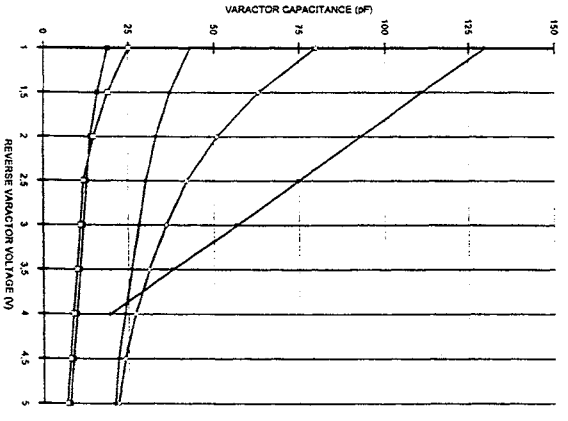


FIG 16 : COMPARISON VARACTOR DIODE / CSA

For practical use, the CSA device is ponderated with two capacitances CL, CP, as shown on figure 16 . It is important to use the whole capacity range of the CSA, in order to minimize the frequency steps corresponding to the capacitance steps .

CL an CP capacitors are optimized so that the complete range of the command word on the CSA induces the expected frequency deviation .

Regarding the wide capacitance range available, it is possible to optimize oscillator's design in different ways :
- For standard fundamental resonators, the nominal value of the serie capacitance can be increased . At lower serial capacitance values, the need of inductors can be minimized or suppressed . Consenquently, the Q factor of the oscillator is sligtlly improved .
- Another interesting way is the use of high stability resonators (3 rd overtone, high Q) . The lower sensitivity (ppm / pF) can be compensated by the wide capacitance range of the CSA . This way is especially interesting when a high stability level is required .

Some first experiments have been done . Fig 17,18 and 19 show some results of temperature compensation, for the following oscillators :

- *TCXO 16 MHz, ATcut, fund. crystal, 5V, Cl = 22 pF*
  temperature range : -50, +100°C
  frequency stability : < +/- 1.10-7
  maximal frequency steps : 1.10-7

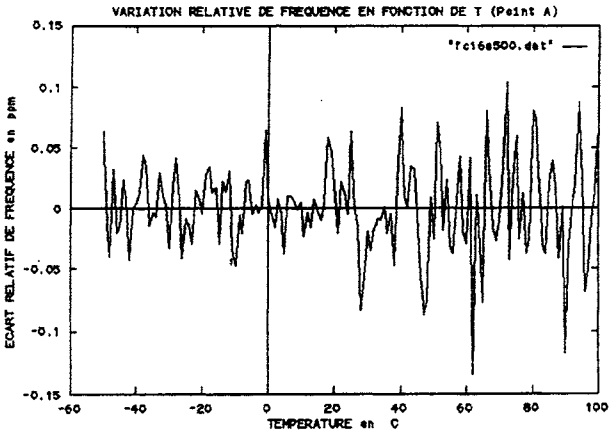


FIG 17 : F(T) COMPENSATED CURVE
16 MHz, AT CUT, FUND. CRYSTAL, CL = 22pF

- *TCXO 16 MHz, ATcut, fund. crystal, 5V, Cl = 90 pF*
  temperature range : -50, +100°C
  frequency stability : < +/- 6.10-8
  maximal frequency steps : 6.10-8

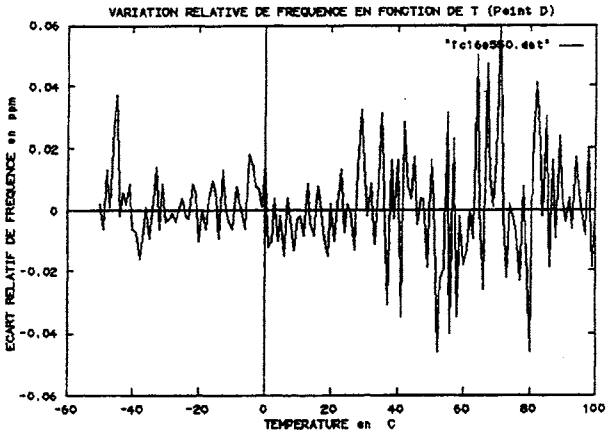


FIG 18 : F(T) COMPENSATED CURVE
16 MHz, AT CUT, FUND. CRYSTAL, CL = 90pF

- *TCXO 16 MHz, AT cut, 3 rd O.T, 5V , Cl = 25 pF*
  temperature range : -50, +100°C
  frequency stability : +/- 5.10-8
  maximal frequency steps : 5.10-8

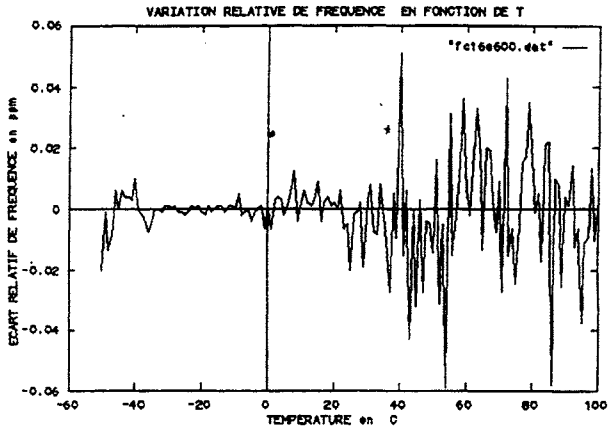


FIG 19 : F(T) COMPENSATED CURVE
16 MHz, AT CUT, 3rd O.T CRYSTAL, CL =25pF

CONCLUSION

Significative improvments have been obtained on the capacitance switch array device, especially regarding the low capacitance steps values combined with a wide capacitance range .
It opens new investigation fields for temperature compensation on oscillators, especially for wider temperature ranges, low consumption applications, as well as for high stability compensated oscillators .

BIBLIOGRAPHY

[1] Takehiko Uno, Yoshio Shimoda, " A nex digital TCXO circuit using a capacitor switch array " Proc. 37 th AFCS pp 434-441, 1983 .

[2] W. Daniau, " TCXO Numérique à réseau de capacité programmable ", Thèse de docteur de l'Université de Franche-Comté, Faculté des Sciences et des Techniques. Université de Besançon, 1991 .

[3] W. Daniau, G.Marianneau, " OSCCAR, Oscillator Compensated with a Capacitance Array " ; Proc. 6 th EFTF (European Frequency and time Forum) pp 165-167, 1992 .

9 TH EUROPEAN FREQUENCY AND TIME FORUM - BESANÇON - 8 - 9 - 10 MARCH 1995

Session Oscillators Number 5

VOLTAGE CONTROLLED CRYSTAL OSCILLATORS USING STRONG PIEZOELECTRIC CRYSTAL UNIDS

V. M. Aksyonov, V.V. Bezdelkin

Joint Stock Company "Phonon", 44 Krasnobogatyrskaya, Moscow, Russia

The designing of miniature oscillators using crystal units made of strong piezoelectrics (SP) is a perspective way in developing VCXO types having wide retuning ranges (0,1 % and more). Theoretical and experimental investigations have been carried out with the aim in view of designing such oscillators at 5 and 35 Mc. For this purpose the circuit of quartz VCXO was optimized with regard of characteristics of crystals units, temperature - frequency characteristics (TFC) were analyzed as well as characteristics of outup voltage control and characteristics outup voltage dependence upon frequency control voltage for $LiTaO_3$ (NL) and $La_3Ga_5SiO_{14}$ (LGS) crystall units. Simultaneous objects of the investigations were the problems of enlarging retuning ranges, desreasing nonlinearity, increasing the steepness of regulation characteristic and improving TFC.

SESSION 6

ATOMIC STANDARDS I

Chairman : L. CUTLER

A Cesium fountain frequency standard : present status and future improvements.

S.Ghezali, P.Laurent, S.Lea, G.Santarelli, M.Bahoura, K.Szymaniec and A.Clairon.
LPTF/NBNM,Observatoire de Paris,61 avenue de l'observatoire, F-75014 Paris.

ABSTRACT:

Neutral atom cooling has opened up the way to the next generation of cesium frequency standards. Frequency stability and accuracy can be considerably improved by using laser cooled atoms. The feasibility of an atomic fountain using cold atoms was demonstrated several years ago. Cesium cooling and atomic manipulation techniques are now relatively simple. A laser cooled Cs atomic fountain has been in operation at the LPTF for a year. Ramsey fringes as narrow as 700 mHz are obtained, two orders of magnitude narrower than in thermal atomic beam devices. At present the short term stability, measured against an H maser, is $2.10^{-13}.\tau^{-1/2}$. Most of the sources of inaccuracy are considerably reduced to the 10^{-16} level, when cold atoms and long interrogation times are used. A new effect which arises in a pulsed fountain is the spin exchange frequency shift due to collisions between cold atoms. The measurements we have performed showed that it is possible simply to reduce this shift to less than 10^{-15}(by using optical molasses) while preserving a potential short term stability of $\sim 10^{-13}.\tau^{-1/2}$. Further reduction of this shift can be obtained by ejecting from the fountain those atoms which contribute to the shift but not to the signal, by using 2-D velocity-selective Raman transitions. The present limitations of our fountain frequency standard and future improvements will be discussed at the conference.

1-INTRODUCTION:

Frequency standards are based on the Ramsey separated oscillatory field method [1] which leads to a resonance line width inversely proportional to the transit time between the two Ramsey interactions:

the longer this time, the sharper the line. In conventional atomic frequency standards, which use thermal beams of a temperature of about 300 K , the width of the resonance is around 100Hz.

With laser cooling, the transit time can be lengthened to 1s leading to a sub-Hertz linewidth.

2-DESCRIPTION OF THE DEVICE AND ITS OPERATION:

The device shown in figure1 is the cesium atomic fountain operating at the LPTF already described in CPEM 95 [3].

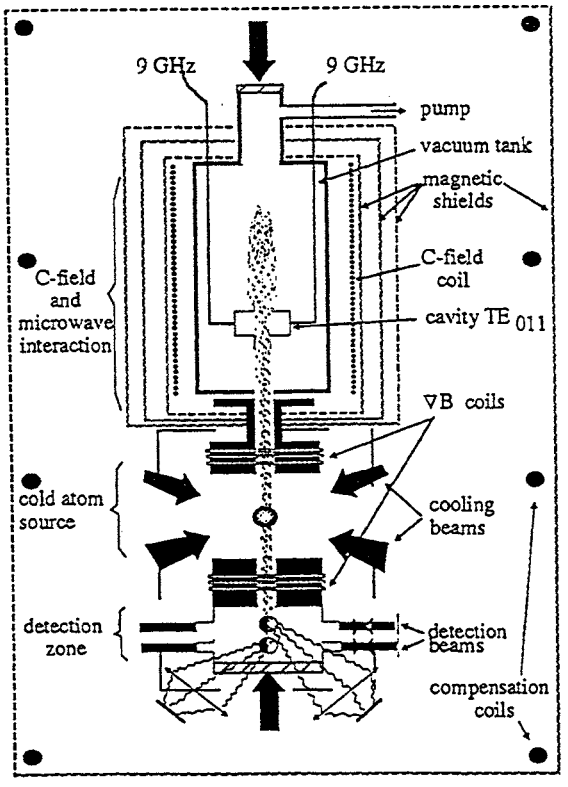


Fig.1:The LPTF fountain prototype.

Six independant laser beams tuned to the red of the cyclic transition F=4--->F'=5 capture and cool atoms to few μK in a low pressure cesium vapor cell (10^{-8}torr) at room temperature. A repumping laser locked to the F=3--->F'=4 transition prevents from atomic losses due to optical pumping to F=3.

From 10^5 up to few 10^8 atoms are cooled down to 2.5 μK in a lin-per-lin optical molasses or a $\sigma^+ - \sigma^-$ magneto optical trap [4]. In the MOT, the gradient magnetic field, produced by two anti-Helmholtz coils, is around 12G/cm. The fountain operates in a pulsed mode. The cycle is about 1.2s. There is, first, a precooling step in a MOT where the six beams are at the maximum intensity ($7mW/cm^2$) and at the optimum detuning δ (-3Γ) during about 300ms up to 1s. The current in the coils is switched off during the last 80 ms of this step to let the gradient magnetic field deacrease to zero. This allows to avoid dramatic heating and an eventual mixture of the Zeeman populations. The loading step is followed by launching which is performed using a moving optical molasses [5] , [6]: when the vertical beams are detuned symetrically of a given quantity $\Delta\nu$, one creates a standing wave moving at the velocity $v_z = \Delta\nu\ \lambda$ in the laboratory frame(horizontal beams can be on, 1D-, or off, 3D-launching). This realizes a static optical molasses in the frame moving at v_z.

The accelerations obtained in the 1D case are at least a factor of two larger than those for the 3D case.

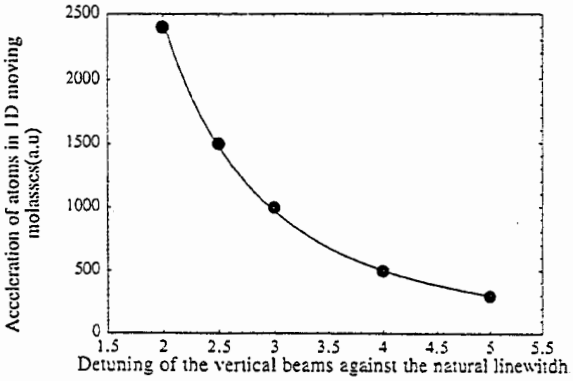


Fig.2:Doppler launching in one dimension.

However, the acceleration should not be too high in order to preserve the initial temperature during the acceleration phase. When the moving molasses is realized by switching immediately the frequencies of the vertical beams, the acceleration varies typically as $1/\delta^2$ (fig.2); this shows that atoms experience a pure Doppler force in 1D launching. As the launching is more efficient for low detunings and high intensities which correspond to rather high temperatures, one must apply to the atoms an additional sub-Doppler cooling for around 2ms where δ is swept from -3Γ to -11Γ in the red of the resonance and Γ is the natural linewidth of the excited state. The intensity is decreased to the fifth part of the maximum during the last 0.25ms. The intensity and the frequency of the laser beams are controlled via acousto-optical modulators. All the beams are cut off by mechanical shutters to avoid eventual light shifts.

The atoms, with an initial velocity of 3.4m/s in our fountain, will perform a ballistic flight in the dark under the effect of gravity . They will pass through a high Q cylindrical TE_{011} copper cavity where they undergo a microwave $\pi/2$ interaction, will continue their flight in a highly homogeneous magnetic field (C-field) produced by a solenoid and compensation coils. These compensation coils provide the continuity of the magnetic field along the atoms' path. The C-field and the microwave region are surrounded by a set of three μ-metal magnetic shields in order to reduce the axial magnetic fluctuations. The axial shielding factor is about 10^3. The whole set-up is surrounded by a forth μ-metal shield. In the upper part of the device, the pressure is kept low (5.10^{-10}torr) with the help of two ion pumps of 25 l/s in order to reduce collisions with residual background gases. Graphite tubes between the different regions reduce the Cs pressure in the C field and microwave region and the detection region.

Due to the aperture of the microwave cavity (1cm diameter), only 10% of the atoms will undergo a second microwave $\pi/2$ interaction. They will be detected in the detection area by a time of flight technique [7] which allows to measure the longitudinal

velocity disribution and thus the axial atomic temperature. A first standing wave tuned slightly below the closed transition F=4--->F'=5 detects atoms in F=4 by fluorescence. Each atom scatters about 10^4 photons. These atoms are then pushed away by a radiation pressure with a travelling wave situated just below. The remaining atoms in F = 3 are first optically pumped in F = 4 by a standing wave tuned to the F = 3--->F' = 4 transition and then detected as described above. The Ramsey signal is equal to $N_{F=4} / (N_{F=3} + N_{F=4})$, where $N_{F=4}$ and $N_{F=3}$ are the number of atoms respectively in F = 4 and F = 3. This normalisation permits to reduce the signal's variations due to shot to shot atom number fluctuations which are of the order of 1%.

3:RESULTS:

The frequency stability can be evaluated in terms of the two sample Allan variance given by :

$$\sigma_y(\tau)=\Delta\nu/(\pi\nu_0).N/S.(T_c/\tau)^{1/2}$$

where T_c is the cycle's fountain time, τ the integration time, ν_0 the central frequency and $\Delta\nu$ the linewidth of the Ramsey fringe.

The signal to noise ratio S/N measured on the top of the fringe is around 700. The expected stability is about $7.10^{-14}.\tau^{-1/2}$ whereas the measured one is only $2.10^{-13}.\ \tau^{-1/2}$ (fig.3). This degradation is expected to be due to the local frequency noise oscillator [8].

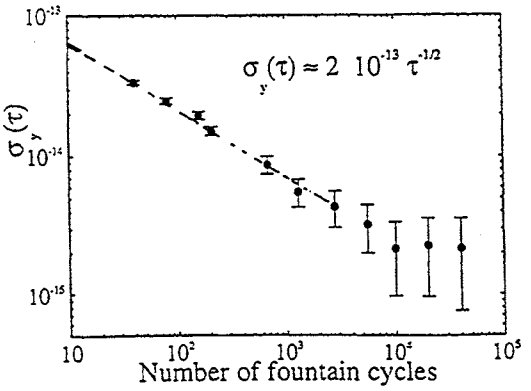


Fig.3:The frequency stability of the fountain.

For 10^4 s integration time, the frequency stability reaches that of the H-maser.

To compare the frequency of the central fringe to the one of an H-maser, a microwave chain multiplies a 10MHz quartz (BVA) phase locked to the maser, up to 9.18 GHz. This frequency is mixed to the one of a dielectric resonator oscillator (DRO) leading to a signal which is mixed again with a synthesizer at 12.63177 MHz. The beat note controlls the DRO which drives the Ramsey cavity. A personal computer digitizes the fluorescence signals from the two detectors giving rise to an error signal used to control the synthesizer.

Figure 4 shows typical Ramsey fringes obtained for a launching height of 60cm. The Rabi envelope is about 60Hz wide corresponding to an interaction time of 15ms in the cavity.

The high number of fringes is due to the narrow velocity distribution. The relative width of the velocity distribution of the atoms is only 1%; this induces a small variation of the fringes' contrast.

The width of the central fringe (fig.5) is less than 700mHz, a factor 100 less than the one obtained in the best thermal beam

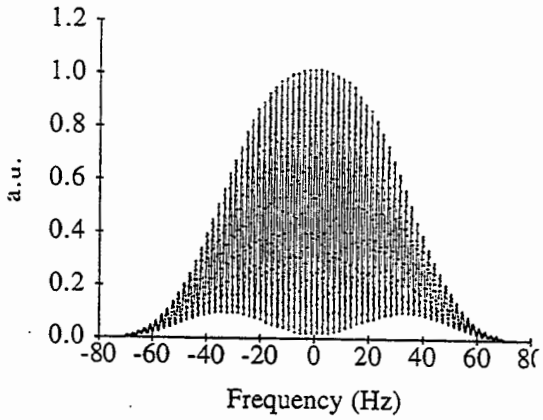


Fig.4: Ramsey fringes obtained by sweeping the microwave frequency around the hyperfine frequency.

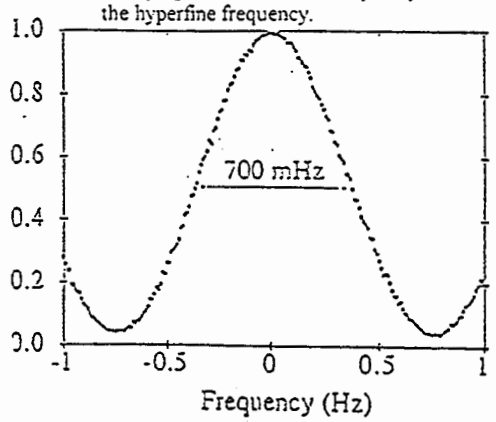


Fig.5:The fringe width is less than 700mHz.

4-FREQUENCY SHIFTS:

A number of major frequency shifts limits the accuracy of any atomic frequency standard. Many of them are considerably reduced when using laser cooled atoms. We will give their order of magnitude and describe the measurements of two of them .

The relavistic frequency shifts are reduced to less than 10^{-17} [9]. At temperatures of 300 K, most of the energy of the black body radiation field is at frequencies much higher than the hyperfine frequency 9.19263177 GHz The shift due an ac Stark electric field perturbation, for a thermal radiation, is about 2.10^{-14}.

All the cavity effects are reduced : the cavity pulling shift can be calculated as in a passive cesium tube as we are far from the maser oscillation due to the rather low atom number[10]. For a cavity quality factor of 30000 and a Q of the atomic line of 10^{10}, the cavity pulling is about 3.10^{-16} for a cavity detuning of 300 kHz.

The residual first order Doppler shift due to the transverse distribution of the microwave field's phase in the cavity is mainly proportional to the distributed phase shift $\delta\phi$. De Marchi and coworkers [11] have estimated, for a high Q cylindrical TE_{011} cavity, $\delta\phi_{max}$ to be 20 µrad leading to the first order Doppler shift of $7\ 10^{-16}$.

As a consequence of the narrow linewidth, the C- field can be reduced in a fountain leading to a reduction of the sensitivity to magnetic field fluctuations and

inhomogeneities. A field of 1.9mG (several tens of mG in a conventional standard) is sufficient to separate the seven microwave transitions $\Delta m_F = 0$ by 1.4 kHz (several tens of mG in a conventional standard). In these conditions, the overlap of neighbouring transitions is completely negligible (fig.6). The further advantage of a fountain is the possibility to map the axial static magnetic field by launching atoms to different heights.

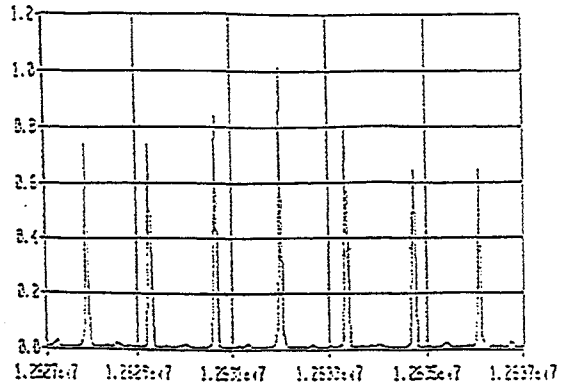


Fig.6:The seven $\Delta m_F=0$ microwave transitions separated by 1.4kHz.

A first method is to measure the difference between a linear magnetic field dependant transition(the $m_{F=3}=1 \rightarrow m_{F=4}=1$ for example,figure7) and the one that does not depend on the magnetic field in the first order ($m_{F=3}=0 \rightarrow m_{F=4}=0$). This gives the time average of the DC magnetic field experienced by the atoms above the cavity. Thus, we have to deduce from a set of measurements the magnetic field as a function of the position above the cavity allowing the calculation of the time average of the quadratic magnetic field.

The Ramsey fringes are not centered into the Rabi envelope of the (1-1) transition (fig.7)because of the significant difference (few tens of μG) between the mean magnetic fields inside the cavity and in the region above it(fig.8). This is due to a parasitic magnetic element inside the cavity. The figure7 shows that Ramsey fringes have a contrast better than 50% : this confirms that the average magnetic field differs by less than 2μG for different atomic trajectories. The central fringe has the highest contrast.

For heights above 40cm, the number of fringes is so large that the central fringe is difficult to identify with an uncertainty less than around 2 fringes.

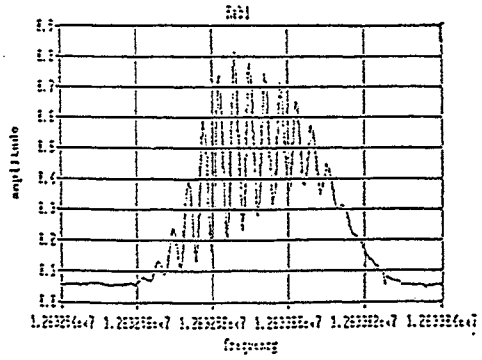


Fig.7:The $m_F=1 \rightarrow m_F=1$ transition.

That is why we finally opted for a second method in which we apply a microwave pulse when the atoms are at their apogee (v = 0) in order to minimise the Doppler effect. We use a 10ms pulse duration to resolve completely the micro-wave transitions. We measured for different

launching heights the frequency difference between two symetric microwave transitions in order to eliminate the residual first Doppler shift. This technique gives directly the measurement of the local magnetic field with a resolution better than 1μG and a spatial resolution less than 1 cm^3.

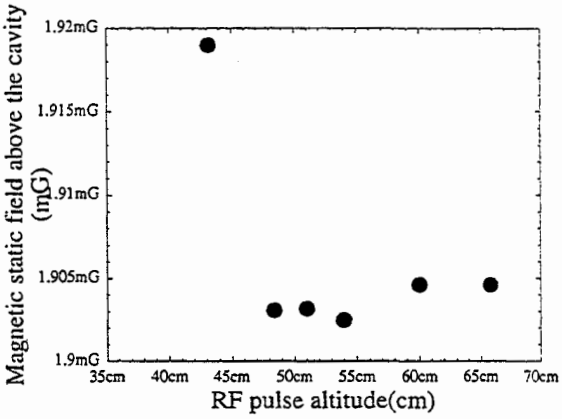


Fig.8:The magnetic field map.The arm of the Ramsey cavity arrives at 38cm.

As shown in figure8, the axial magnetic field variations above the cavity are less than 2 μG. This measurement allows us to calculate the time average of the quadratic magnetic field with an equivalent frequency uncertainty of about 5.10^{-16}.

With the new cavity, the magnetic perturbation will be suppressed and the techniques described precedently will allow to determine the static magnetic field with less than 1μG of uncertainty. We have observed also a degradation of the magnetic field homogeneity when using a magneto optical trap.

The Rabi and Ramsey shifts are estimated to be negligeable at 10^{-14}. These two shifts can be eliminated by using atoms in m_F=0 [12].

A new effect which arises in a laser cooled atomic standard is the spin exchange frequency shift. It has not been taken into account for conventional devices because of the large temperature.

At μK temperatures, the atomic De Broglie wavelength λ_{DB} become unusually large, much larger than the scale of the interatomic potential. Tiesinga et al [14] have performed a quantum calculation where they predict that, for small relative velocities, mainly "s" scattering waves contribute to cold collisions leading to crosssections as large as λ^2_{DB} / π. For s waves, the collisional frequency shift is thus independant on the velocity and is mainly proportional to the density averaged over the free flight above the cavity n_{ave}. It is given by:

$$\Delta\nu/\nu \sim N_i. / 2. (2.\Pi)^{3/2}. \sigma_I^2. \sigma_F = n_{ave}$$

where N_i is the initial atom number, σ_I and σ_F are the initial and final r.m.s atom cloud sizes in the cavity.We have loaded atoms in a MOT to have average densities of the order of 10^9at/cm^3 above the cavity. The

MOT allows an increase of the atomic density of a factor more than 10 compared to that obtained in an optical molasses. Thus, cold collisions can be investigated for a large domain of densities.

Measurements have been first performed for atoms almost equally distributed among all Zeeman sublevels of the hyperfine level F=4 of the ground state (fig.6). We have measured a relative collisional frequency shift of $\sim -1.10^{-21} n_{ave}$ (fig.9). In this regime, the collisional shift is found independant of the temperature for low temperatures. Two measurements have been performed with a loading in an optical molasses ($\sim$ few 10^6at/cm^3). We vary the density by varying the loading time or the cooling laser detuning.

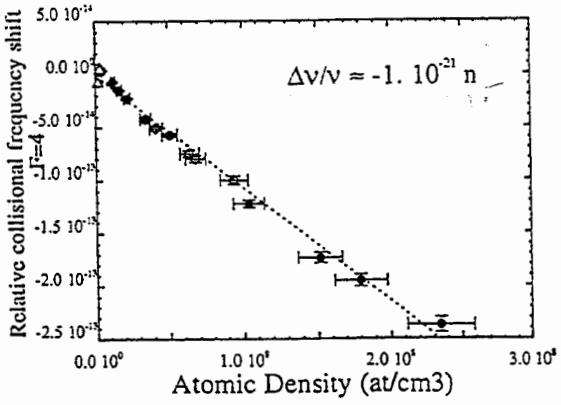


Fig.9:The collisional frequency shift for atoms prepared in F=4.Loading in a MOT: ● temperature 6 μK, ◇ 3 μK. Loading in an optical molasses: Δ 8 μK

Other measurements have been done with atoms prepared in F = 3 with a pulsed optical pumping laser beam tuned below the transition F=4--->F'=4 by about 4MHz. The 4-4 beam has been superimposed to the down going vertical beam. To perform a complete optical pumping, we needed a $\sim$ 8ms pulse for a weak power of $\sim$ 0.1mW. The atoms remaining in F = 4 are pushed down by radiation pressure with the z cooling beam switched on for few ms. The optical pumping heats the atoms of around 3μK. We measured a relative collisional shift of $\sim 5.9\ 10^{-22}\ n_{ave}$ (fig.10).

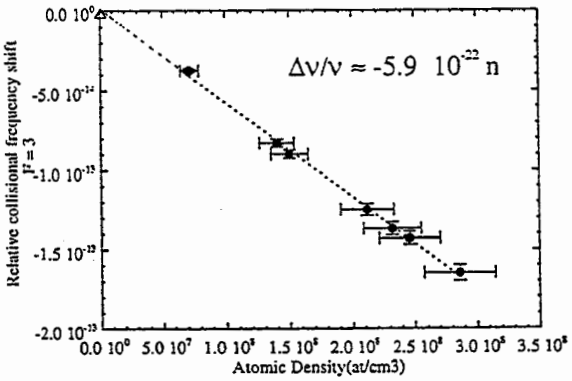


Fig.10: The collisional frequency shift for atoms prepared in F=3.Loading in a MOT: ● temperature 8 μK. Loading in an optical molasses: Δ 11 μK.

We have symetrized the zeeman populations in F=3 by adjusting the 4--4 pumping beam's polarization.

We have also done preliminary measurements for most of atoms prepared in m_F=0. The corresponding collisional shift is larger than the previous ones.These measurements will be repeated in the new fountain in which a microwave cavity will be

added for such a microwave selection. The uncertainty in the absolute density determination is at least about 50% due to uncertainties on the solid angle collection and the determination of the cloud's size. The systematic errors, being the same for both measurements, are eliminated when compa-ring the results. The collisional shift for atoms prepared in F = 3 is of the order of about 60% of that for atoms distributed in F = 4. The results of the Stanford group agree with ours[12].

5-CONCLUSION AND PROSPECTS:

At present, when using optical molasses, collisions constitute a limitation of the accuracy at the 10^{-15} level. The spin exchange shift which seems to be the ultimate limitation of the accuracy of any laser cooled atomic standard can be reduced to 10^{-16} by ejecting from the fountain the atoms which contribute to collisions but are not detected. These atoms are those which are in the wings of the transverse velocity distribution. This can be realized with 2D velocity selective Raman transi-tions [15] , [16].

A further improvement of the frequency stability of a factor of two can be achieved by using the best available quartz oscillator. Presently, the microwave leakage is not negligible in our device.

The induced shift is estimated to be $\sim 10^{-14}$. A new cavity will replace the present one and we expect that the leakage will be considerably reduced. With a quality factor of around 10^4, the cavity pulling shift will be completely negligible.

Before the accuracy of our device can be assessed, we have to investigate fully all the magnetic field's effects: Majorana transitions, Rabi and Ramsey shifts...and the spectral impurities. Nevertheless, an accuracy of 10^{-15} can be achieved reasonably in the near future.

A better evaluation of our device requires a comparison with an other standard with the similar performances. A second fountain using a 2 D Raman selection is being built.

REFERENCES:

[1]:N.F.Ramsey, Phys.Rev., 78(1950), 695.

[2]:J.R.Zacharias, unpublished(1953).

[3]:A. Clairon et al, IEEE Trans. Instrum. Meas.,April 1995.

[4]:C. Monroe, W. Swann, H. Robinson and C. Wieman, Phys.Rev.Lett., vol. 65, pp.1571-1574,Sept1990.

[5]:M.Kasevich, E.Riis, S.Chu and R. DeVoe, Phys.Rev.Lett., vol.63, pp.612-615,1989.

[6]:J.L.Hall and M.Zhu, J.Opt.Soc. Am.B, 6(1989)2194.

[7]:A.Clairon, C.Salomon, S.Guellati and W.D.Phillips, Europhys.Lett., vol.16, pp.165-170, Sept1991.

[8]:J.Dick, Proc. of PTTI, Redendo Beach, 1987, pp.133-147.

[9]:J.Vanier and C.Audoin, The quantum physics of atomic frequency standards,vol.2, pp.785

[10]: J.Viennet, C.Audoin and M. Desainfuscien, IEEE Trans.Intrum.Meas., vol.21, pp.204-209, August1972.

[11]: G.Vecchi and A. De Marchi, IEEE Trans.Instrum.Meas., vol42, pp.434-438, April1993.

[12]:K.Gibble and S.Chu, Phys.Rev. Lett., vol.70, pp.1771-1774, March1993.

[13]:L.Cutler, A. De Marchi, C.Flory and R.Giffard, J.Appl.Phys., vol.69, pp.2780-2792, March1991.

[14]:E.Tiesinga, B.J.Verhaar, H.C.T. Stoof and D. Van Bragt, Phys.Rev.A, vol.45, pp.R2671-R2673, March1992.

[15]:M.Kasevich, D.SWeiss, E.Riis, K.Moler, S.Kasapi and S.Chu, Phys.Rev. Lett., vol.66, pp.2297-2300, May1991.

[16]:S.Lea et al, Physica Scripta , vol.51, 1994.

PROPOSAL FOR THE GENERATION OF A CONTINUOUS BEAM OF COLD ATOMS

N. DIMARCQ, E. AUCOUTURIER, P. PETIT and P. CEREZ.

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

ABSTRACT.

This article is divided in three parts. The first one concerns the generation of a continuous flux of cold atoms which may be more intense and more reproducible that the conventional pulsed ones. In the second part, we describe a new 3 dimensional cooling configuration in a laser isotropic light. This configuration appears to be very efficient and insensitive to various misalignments of the cooling laser beams. In the third part, we show that new extended cavity SC laser structures, including self-alignments concepts, can correctly operate during a very long time without any further setting.

I - INTRODUCTION.

It is now admitted that the operation of atomic clocks with cooled atoms must improve both frequency stability and accuracy of these devices [1]. The key point of such clocks is the generation of a cold atoms beam. At the present time, all the cold atoms sources, which have a pulsed operation, generate fluxes whose intensities are low and not reproducible. This is detrimental for the clock S/N ratio and for the clock frequency stability. In order to solve this problem, we propose an original lauching configuration which should generate a continuous beam of cold atoms. The flux of this continuous beam is expected to be more reproducible and more intense than a pulsed one. This configuration is described in Section II.

The flux of cold atoms contained in a continuous or a pulsed beam depends on the number of atoms which are initially cooled before launching. In a conventional molasse, this number closely depends both on the laser beams diameter and on the capture velocity. In basic experimental set-ups, the diameters of the laser beams are limited due to the low optical power delivered by laser diodes. As a consequence, the number and the volume of cold atoms are reduced. We have experimentally tested a new 3 dimensional cooling configuration of cesium atoms contained in a cell with an isotropic monochromatic laser light. This configuration, which has been revealed to be very efficient, is described in Section III.

All the cooling experiments require high spectral purity laser sources. Extended-cavity techniques are the most efficient to improve the spectral purity of commercial laser diodes. Nevertheless, the long term operation is often disturbed by detrimental mechanical drifts. We have designed new laser structures including self-alignment concepts whose principle is detailed in Section IV. We have demonstrated that such a laser is insensitive to misalignments of the cavity components.

II - Generation of a continuous beam of cold atoms.

II-1 Conventional pulsed cold atoms sources.

At the present time, all the cold atoms sources have a pulsed operation whose typical cycle is the following. In a first time, a ball of atoms is cooled in an three dimensional (3D) optical molasse or a magneto-optical trap (MOT) [2]. Then, the cooling laser beams are shut and the ball of cold atoms is launched using a technique which is described in Section II-2. In such a pulsed operation, the rate of launching is rather low : the cycle period is about one second. Moreover, the number of cold atoms launched at each period is reduced ($\approx 10^6$ atoms) and its fluctuations are far larger than the usual Poissonnian atomic shot noise. The main consequence of this enhanced shot noise is a degradation of the clock S/N ratio. The influence of this important atomic shot noise can be reduced by using complicated normalization techniques for the detection of the clock signal. An other problem due to the pulsed operation of a cold atoms clock is related to the dead time between two successive cycles. As a matter of fact, as the microwave interrogation is not continuously performed, we have no information about the clock frequency between two successive interrogation processes. The duty cycle (interrogation duration compared to cycle duration) is rather low.

The simplest way to solve these two problems is the use of a continuous beam of cold atoms. In such a beam, it is expected that the cold atoms flux will be more intense and that the related fluctuations will be reduced compared to the case of a pulsed beam. Moreover, with a continuous beam, the microwave interrogation can be performed almost continuously and the corresponding duty cycle is close to unity.

II-2 Generation of a continuous flux of cold atoms.

It is necessary to perform simultaneously the cooling of the thermal atoms and the launching of the cold atoms .

a) Transverse cooling of the atoms.

The good collimation of the cold atoms beam is ensured by the cooling and the trapping of the atoms in the plan perpendicular to the launching direction. This is realized by a two dimensional (2D) magneto-optical trap (MOT). The principle of such a 2D-MOT is depicted in Figure 1. Two orthogonal stationnary laser waves directed along the x and y axis perform the cooling processes. Each stationnary wave has progressive components which are respectively circularly polarized σ^+ and σ^-. An inhomogeneous static magnetic field is generated by two pairs of coils supplied with suitable current polarities. The variations of its transverse component B_x (resp. B_y) versus x (resp. y) is shown in Figure 2 for a current intensity I equal to 0,5 A. The B_z component is equal to zero along the z axis.

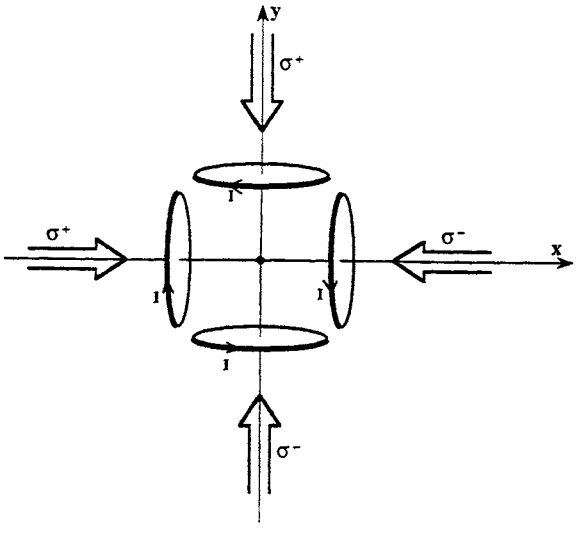


Fig.1 2D Magneto-optical trap.

The combined effects of the polarized laser light and of this inhomogeneous magnetic field induce a transverse cooling (along the x and y axis) and a trapping of the atoms towards the z axis.

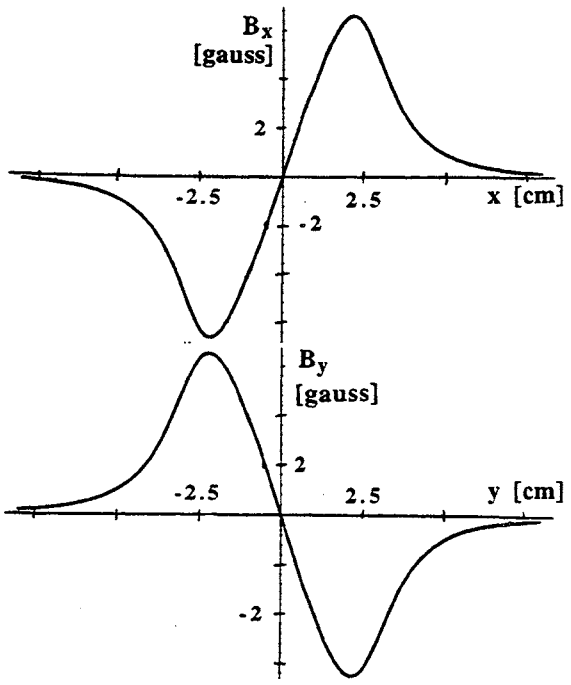


Fig.2 Transverse components of the magnetic field.

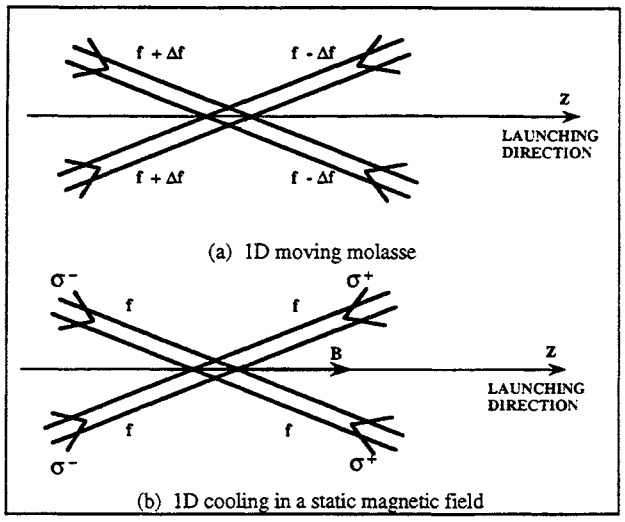


Fig.3 Launching laser standing waves

b) Lauching of the cold atoms

The launching of the atoms require special one dimensional (1D) cooling configurations.

1D moving molasse

The atoms are submitted to a stationnary laser wave made of two progressive waves whose frequencies are symmetrically shifted around the cooling laser frequency. The atoms are thus cooled in a frame which is moving at the speed $v_{launching}$ approximatively given by :

$$v_{launching} = \lambda \Delta f \qquad (1)$$

where λ is the laser wavelength and Δf is the frequency shift.

1D cooling in a static magnetic field

The atoms are cooled in a stationnary laser wave whose progressive components are respectively circularly polarized σ^+ and σ^-. A static magnetic field is directed along the laser beams. The cold atoms are then launched at a speed given by [3] :

$$v_{launching} = \frac{\Delta\omega_{Larmor}}{k} \qquad (2)$$

where $k = 2\pi/\lambda$. The difference $\Delta\omega_{Larmor}$ between the Larmor frequencies of the fundamental and the excited states is proportional to the static magnetic field magnitude.

The second technique is easier to operate but the first one is more efficient for the launching at a speed larger than 1 $m.s^{-1}$. After they have been launched, the cold atoms cross the various interaction regions (atomic state preparation, microwave interrogation, optical detection of the clock signal) of the resonator. In the case of a continuous operation, spurious light in the resonator disturbs the clock operation because it creates light shifts and it may modify atomic trajectories due to radiation pressure effects. Two types of spurious light must be distinguished :

Spurious laser light.

No laser beam must be along the atoms trajectories in the resonator. As a consequence, the launching is not performed using a single standing laser wave but two laser standing waves which intersect in the interaction zone (Figure 3). The atoms are launched along the bisecting line of the laser beams. This configuration is valid for the 1D moving molasse (Fig. 3.a) and for the 1D cooling in a static magnetic field (Fig. 3.b). As a result there is no laser light along the atomic paths.

Spurious fluorescence light.

The fluorescence light emitted by the cold atoms source induces a frequency shift which degrades the clock accuracy. Two solutions can be considered. The first one consists in precisely evaluating the frequency shift due to fluorescence light effects. This solution, which may be available for miniature clocks using cold atoms, cannot be obviously applied in the case of primary standards. The second solution consists in chopping the cold atoms beam : in this case, the microwave interrogation occurs when the spurious fluorescence light is stopped by the chopper. The chopping rate must be adjusted to the microwave interrogation duration. With this solution, the atomic beam is quasi continuous.

In all the launching techniques (pulsed or continuous ones), the atomic density in the beam is directly related to the number of atoms initally cooled before launching. This number directly depends on the capture velocity which is a function of various parameters such as the optical intensity and the diameter of the cooling laser beam. Increasing this number requires higher optical power which cannot be reached with laser diodes. We have developed an original 3D laser cooling configuration where the cesium atoms contained in a cell are efficiently cooled with an isotropic laser light.

III - 3D cooling of cesium atoms with isotropic light.

III-1 Description of the technique.

The experimental set-up is described in Figure 4. The pyrex cell (diameter = 4 cm) containing the atoms is enclosed in an integrating sphere which creates the isotropic light. This sphere (diameter = 8 cm) is made with SPECTRALON (Labsphere) which reflects more than 99% of the incident light. The laser light is injected inside the sphere with 14 multimode fibers symmetrically disposed around the sphere. The surface of the pyrex cell is frosted in order to prevent specular reflexions. The cooling light comes from a SDL laser diode which delivers about 100 mW and which is spectrally purified with an injection locking technique. The master laser is a SC extended cavity laser which is described in Section IV. Its frequency f_L is tuned below the resonance frequency f_0 of the $4 \rightarrow 5$ transition of the cesium D_2 line. The pumping laser, tuned to the $3 \rightarrow 4$ transition, avoids the trapping of the atoms in the F=3 fundamental state.

III-2 Advantages of the cooling with isotropic light.

In conventional cooling configuration using collimated laser beams, the atoms are slowed down if they belong to the intersection region of the laser beams. The optical molasse volume is then directly related to the beam diameter. In the presented configuration, the whole cell is irradiated by isotropic light. As a consequence, all the atoms contained in the cell are submitted to the radiation force induced by successive {absorption + spontaneous emission} cycles. We can expected that the molasse volume will be almost equal to the cell one.

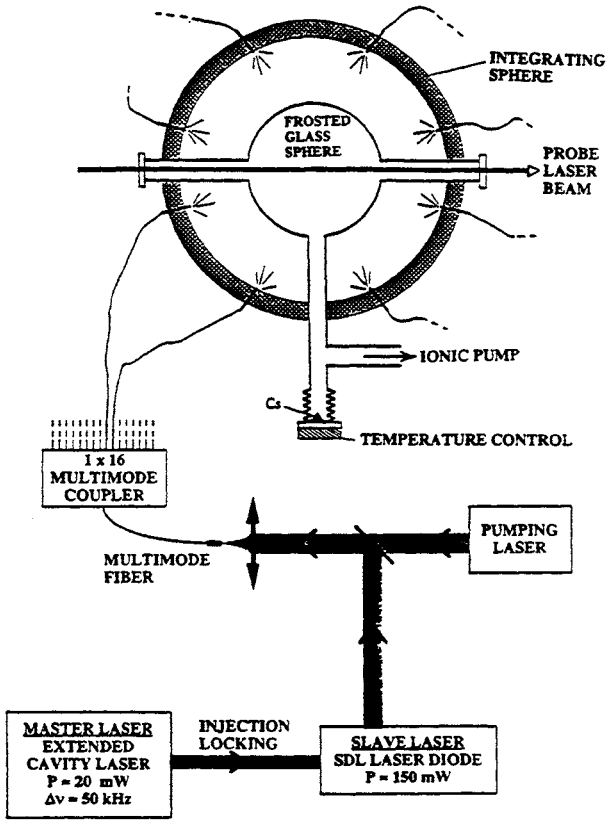


Fig.4 Experimental set-up

Due to the Doppler effect, the frequency of the light in the moving atomic frame is different from the one in the laboratory frame. The cooling processes are efficient if the laser frequency is resonant in the atomic frame. This implies that :

$$f_0 = f_L + \frac{v \cos\theta}{\lambda} \qquad (3)$$

where f_0 is the resonance frequency, f_L is the laser frequency, v is the atomic speed and θ is the angle between the atomic velocity and the laser beam direction. In conventional cooling schemes with collimated laser beams, this relation cannot be always fulfiled because the speed v changes when the atom is slowed down whereas f_0, f_L and λ are fixed and θ is random. In the cooling configuration with isotropic light, there is always a light cone (centered on the atomic velocity) which is resonant in the atomic frame (Figure 5). The angle θ of such a cone is fixed by the resonance condition (3). During the slowing down of the atom, this condition is always fulfilled because the variation of the speed v is balanced by the variation of the angle θ of the resonant light cone (Figure 5). This self-adaptation effect lead to a noticeable increase of the cooling processes efficiency. This has been verified with Monte-Carlo simulations and tested experimentally.

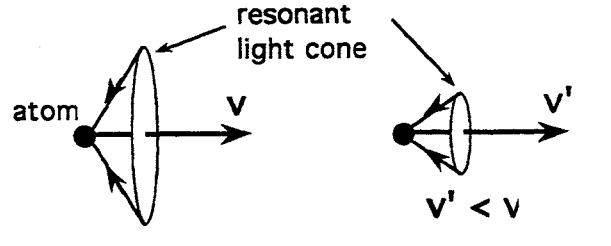


Fig.5 Self-adaptation of the resonance condition

In conventional cooling configurations, a low magnitude magnetic field is required for the good operation of the cooling processes. Moreover, the exact alignment of the different progressive waves is necessary : any misalignment breaks the symmetry between the radiation pressures exerced by the laser beams. In the cooling configuration in isotropic light, the cold atoms have been observed without any compensation of spurious magnetic fields (Earth or instrumental magnetic fields). The only critical alignment is the injection of the laser beam in the optical fiber. The equilibrium between the optical powers emitted by each fiber in the integrating sphere is not important because the integration sphere drastically filters the light inhomogeneities. This configuration should be able to operate with a lower number of optical fibers which could be reduced to 2 fibers only.

III-3 Experimental results.

In order to point out the cooling phenomenon, we observe the linear absorption of a probe laser beam which crosses the cell (Figure 6). When there is no cooling laser, the observed signal is Doppler broadened. The width at half maximum is about 500 MHz. When the cooling and pumping lasers are in operation, we point out the narrow absorption peak of the cold atoms. There are three peaks corresponding to the three authorized optical transitions.

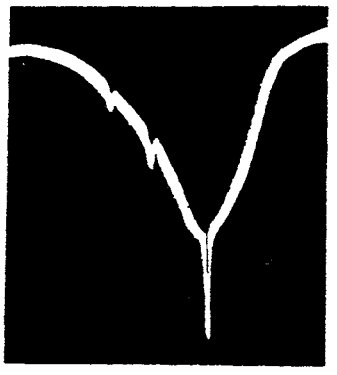


Fig.6 Linear absorption signal of a probe laser beam

It is impossible to deduce the cold atoms temperature from the absorption lines width because the observed width is the radiative linewidth (≈ 5 MHz) which correspond to a temperature about 40 mK. The expected temperature depends on the cooling process : it is about 120 μK in the Doppler limit but it can be lower (a few μK) if light and polarizations gradients induce Sisyphus effect [4].

With the presented preliminary experimental set-up, the typical number of cold atoms is about 10^9 for an injected optical power equal to 30 mW.

IV - Self-aligned extended cavity SC lasers.

IV-1 Settings of an extended cavity laser.

The schematic principle of an extended cavity laser (ECL) is depicted in Figure 7. The cavity is closed by a grating in a Littrow configuration. Due to the wavelength dispersion of the grating, the emission wavelength of the ECL is fixed by the wavelength of the light exactly fed-back in the laser diode. The good operation of the ECL thus requires three settings of the collimation lens along the three axis :

- along the z axis (focusing) : a 10 μm misadjustment is acceptable
- along the horizontal y axis : there is no requirement on the adjustment because this setting corresponds to a wavelength tuning of the ECL
- along the vertical x axis : any spurious misalignment is the main origin of the mechanical instabilities.

In order to reduce the influence of misalignments, we have developed self-alignment techniques. With such techniques, the laser diode is properly fed-back, even if an optical component of the extended cavity is misaligned.

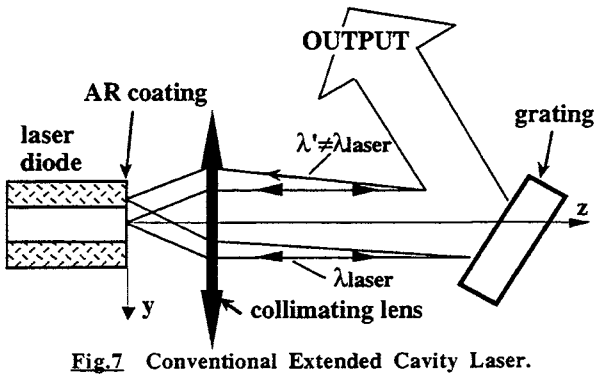
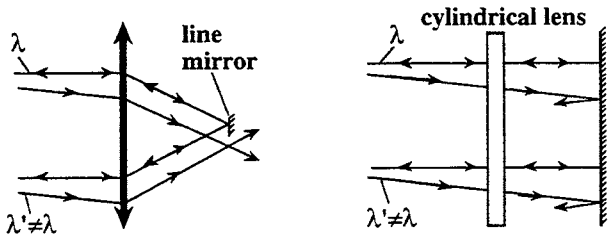


Fig.7 Conventional Extended Cavity Laser.

IV-2 Self-alignment (S.A.) techniques.

The optical component which ensure a S.A. function is a cat's eye. A cat's eye is made of a lens and a plan mirror located in the focal plan of the lens. This cat's eye plays the role of a mirror which reflects a beam parallel to the incident beam for any incidence angle of the incidence beam. Nevertheless, this S.A. must not operate along the y direction which is the direction of the wavelength selectivity realized by the grating. As a consequence, we have designed one dimensional (1D) cat's eye which ensure a S.A. function along the x axis but which has no function along the y axis. Two kinds of 1D cat's eye are shown in Figure 8. In Fig. 8-a, the mirror of the cat's eye is a very thin line mirror (width ≈ a few μm) which is parallel to the x axis. The spatial selectivity of the reflexion ensures the wavelength selectivity. In Fig. 8-b, The lens of the cat's eye is a cylindrical lens whose axis is parallel to the y axis.



(a) (b)

Fig.8 1D cat's eye configurations

One of the tested self-aligned ECL is depicted in Figure 9. In this case, the grating is not in a Littrow configuration. Its first diffraction order is directed towards the 1D cat's eye. The laser optical output is the zero order of the grating.

In order to test the efficiency of S.A. techniques, we have observed the variations of the threshold current as a function of a misadjustment of the collimation lens along the x axis. These variations are shown in Fig.10. In Fig. 10-a, there is no S.A. configuration beacuse the cylindrical lens has been removed. The threshold current rapidly increases as soon as the misadjustment is equal to the dimension (along x) of the active zone of the laser diode (≈ 1 μm for SDL laser diodes). In Fig. 10-b, the S.A. configuration is in operation and leads to a dependance on misalignements reduced by a factor 50.

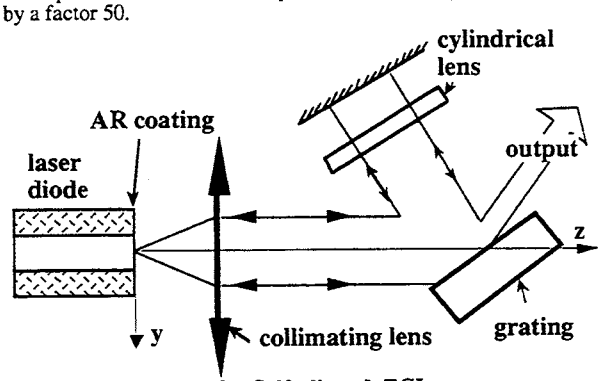
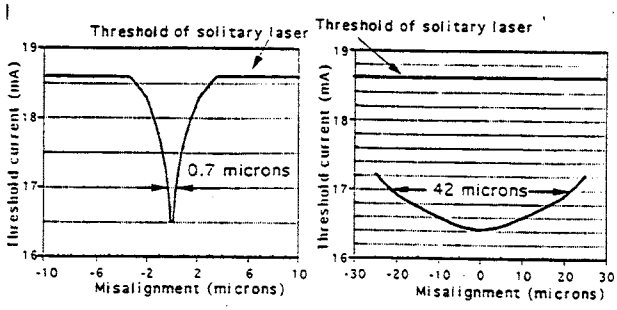


Fig.9 Self-aligned ECL.



(a) without S.A. configuration

(b) with S.A. configuration

Fig.10 Variations of the threshold current as a function of the vertical misadjustment.

We have tested an other self-aligned ECL structure where the grating plays the role of the mirror in the 1D cat's eye (Figure 11). In this case the grating is in the Littrow configuration. The cylindrical lens is used both for the S.A. in the 1D cat's eye and for the collimation of the output laser beam. The reliable operation of this simplest self-aligned configuration has been experimentally demonstrated.

Thanks to S.A. techniques, it is thus not necessary to control the vertical setting. We have designed a compact ECL structure where the positions of all the optical components are initially adjusted and fixed by sticking. The only remaining setting is the wavelength selection realized by combined translation and rotation movements of the grating. We have demonstrated that such a ECL has remained correctly aligned during 18 months without any further adjustment. Morover, such a structure has been successfully tested in microgravity during parabolic flights.

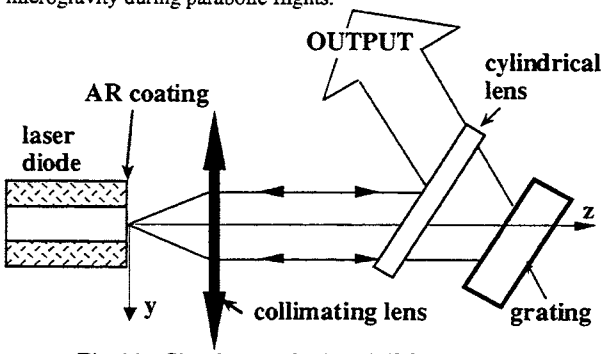


Fig.11 Simplest self-aligned ECL structure

V - Conclusion.

We must demonstrate experimentally the faisability of the generation of a continuous beam of cold atoms. The created beam will be characterized : atomic density, collimation of the atomic paths, reachable launching velocities, etc...

We will also determine the main characteristics of the cooling configuration with an isotropic laser light : atomic density and temperature, localisation of the atoms, ... It will not be possible to measure the temperature with a time of flight method. Other methods, such as Raman stimulated processes, will be tested.

We have demonstrated that self-aligned extended cavity lasers are very reliable. The optimal structure, which increases the continuous tuning range of such lasers, must be now designed.

Acknowledgments

The authors want to thank the Bureau National de Metrologie (Paris) for its financial support.

References

[1] A. Clairon and al., IEEE Trans. on IM , to be published
[2] C. Monroe and al., Phys. Rev. Lett. 65, 1571 (1990)
[3] C. Valentin and al., Europhys. Lett. 17, 133 (1992)
[4] J. Dalibard and al., J.O.S.A. B6, 2023 (1989)

Atomic Trapping and Cooling in an Anisotropic Trap

N. Sagna, G. Dudle, P. Thomann

Observatoire Cantonal Neuchâtel, 58, rue de l'Observatoire, CH-2000 Neuchâtel

Abstract

We report on investigations of the magneto-optic trapping of Cs atoms using other geometries than the usual, quasi isotropic anti-Helmholtz coil pair. We have trapped large samples of cold atoms in a configuration where 6 laser beams provide a tridimensional viscous force for cooling, while a two-dimensional magnetic gradient confines atoms along the third, neutral axis.

We compare the trapping efficiency of such a structure with that of an isotropic trap and with predictions of a model that includes the effects of the magnetic field. The potential of anisotropic traps for producing beams of slow atoms for atomic frequency standards is also discussed.

1 Introduction

Frequency standards using laser cooled Cesium atoms are about to be operational [1, 2]. All proposals and works done in this field up to now are based on a pulsed scheme: the atoms are first trapped in a magneto-optical trap (MOT), in a second step cooled in an optical molasses and then vertically launched on a ballistic flight where the microwave interrogation takes place. This procedure has undeniable advantages. First, the atoms pass twice through the same cavity on their ballistic flight. The end to end phase shift is therefore completly cancelled. Second, for a given interaction distance, this method maximises the interaction time since the vertical velocity goes through zero. Third, since all lasers are off during the RF-interrogation there are no light shift problems related to irradiation from the source into the cavity.

However, there are different drawbacks to this scheme. The first is related to its low duty cycle. Indeed, the fact that one cycle must be finished before the next one can start puts severe limits to the average signal. To achieve a good S/N ratio one needs high atomic densities in each shot which means also sizeable collision effects [3, 4]. A second point is the stability transfer from the atomic resonator to the local oscillator. It has been shown [5] that when the loop gain in a passive standard is time-dependent, as is the case in a pulsed beam, the local oscillator noise limits the short-term stability of the standard to a level significantly higher than that expected from the resonator Q and S/N alone.

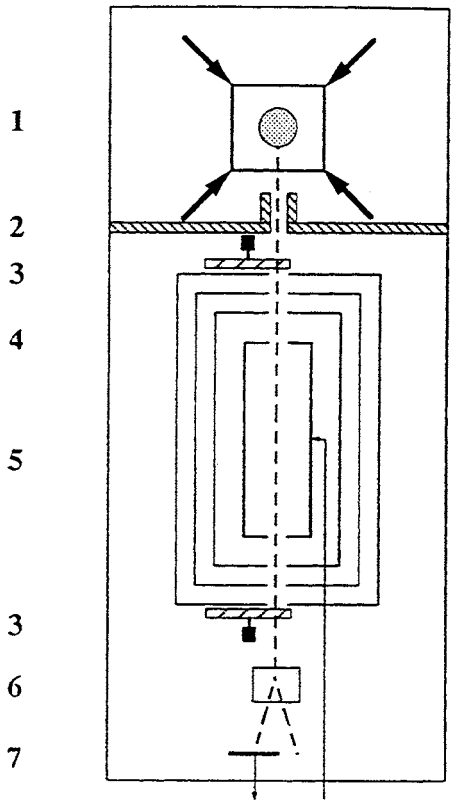


Figure 1: Scheme of the cold cesium beam resonator being built.

For these reasons a frequency standard based on a continuous beam of laser cooled atoms is an interesting alternative to a pulsed version in spite of its apparent disadvantages.

Figure 1 shows a scheme of the cold Cesium beam resonator that is being built at the Observatoire de Neuchâtel. The cold Cesium source (1) is located in the upper part of the apparatus. This part is separated from the rest of the vacuum system by a graphite getter (2). From the source the cold atoms fall through a rotating light trap into the interrogation region. Such a light trap (3) is necessary here to avoid light shift

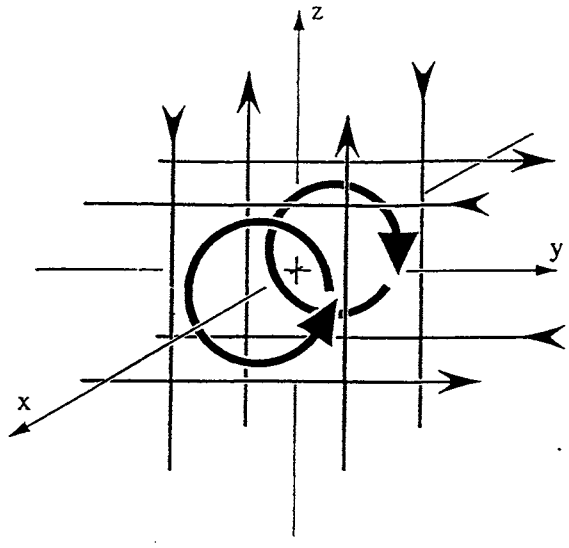


Figure 2: Structure for the magnetic field gradient. A set of 4 straight wire conductors is parallel to the z-axis another independent is parallel to the y-axis.

problems due to irradiation from the source into the microwave cavity. The trap is based on velocity selectivity and lets atoms in while keeping light out of the microwave cavity. A number of magnetic shields (4) surround the microwave cavity (5). At the exit of the cavity again a light trap, followed by the state selector (6) and finally the detector (7).

We work currently on the continuous source of cold atoms. We present here the experimental set up (section 2), a systematical study of the capture efficency *vs.* the important cooling parameters and a comparison between the behaviour of a 3-D and a 2-D trap (section 3). In section 4 we discuss the potential of producing a continuous beams from a 2-D trap.

2 Experimental Set up

A well known way of producing a magnetic gradient is the anti-Helmholtz coil pair: two coaxial loops are supplied with currents of opposite sign. The zero divergence of the magnetic field and the rotational symmetry imply $dB_y/dy = dB_z/dz = -0.5\ dB_x/dx$ if x is the common axis of the coils.

In this work we use a different geometry for the magnetic field gradient(see Fig. 2): four straight wire conductors parallel to the vertical z-axis are disposed on a square centered at the origin. If adjacent conductors are supplied with currents of opposite sign this array produces a two dimensional quadrupole field in the horizontal Oxy plane. A second, independent set of conductors parallel to the y axis adds a similar quadrupole field in the Oxz plane. The combination of both structures yields a gradient almost equivalent to that of the anti-Helmholtz coils. This fact is not surprising since near the center the eight conductors form together two rectangular quasi loops parallel to the Oyz plane with opposite current in the two loops. While the x-axis is a rotational symmetry axis in the case of anti-Helmholtz coils, it is a four fold axis here. Each four-wire structure can produce a maximum gradient dB_x/dx of 7.5 G cm^{-1}. When both structures are operated at maximum current, the x-component reaches 15 G cm^{-1}. But due to the independence of the current segments, this assembly can be used not only to study three dimensional magneto-optical traps but also to investigate anisotropic situations. One example is the 2-D trap where only one of the two structures is supplied and produces a two-dimensional gradient in the Oxy plane or the Oxz plane respectively. In both cases there is no gradient component in the third direction. This situation leads to a trapping in the two directions with magnetic gradient and an optical molasses in the third direction.

Six independent laser beams provide a 3-D viscous force for cooling. Four of them are used to produce two standing waves in a vertical Oyz plane, two propagate along the horizontal x axis, which is also the common axis of the rectangular loops. The beams in the Oyz plane are at an angle of 45° to the horizontal plane, such that no beam propagates parallel to the vertical atomic beam path.

While the beams at 45° in the Oyz plane contribute to the trap in the y direction and to the molasses in the z direction, the two x-axis beams act only in a trapping direction. The independence of the four beams in the vertical plane allows to introduce frequency or intensity differences between beams of opposite z-components.

The light of an additional laser is added to the x beams. This repumping laser is locked to the $F = 3 \rightarrow F' = 4$ transition.

3 Results and Discussion

For each one of the two magnetic field configurations, we have measured the number of trapped atoms as a function of the essential trapping parameters such as laser intensity, laser detuning and magnetic field gradient. While analytical treatments [3, 6] of the capture process and the number of trapped atoms neglect the effect of the magnetic field, there is some experimental evidence [1, 7] of the critical role the magnetic field gradient plays in the capture process.

Let us briefly present simple ways to describe the capture process in the two situations. As discussed in Ref. [8] under the assumption of its isotropy, a three dimensional magneto-optical trap can be described by an entirely 1-D model that takes into account the magnetic field gradient. One indeed assumes that when an atom enters the trapping volume it sees a pair of co- and counterpropagating beams. Although most of the atoms enter the trap off-axis, it is shown [8] that if typical 3-D features – such as the power broadening due

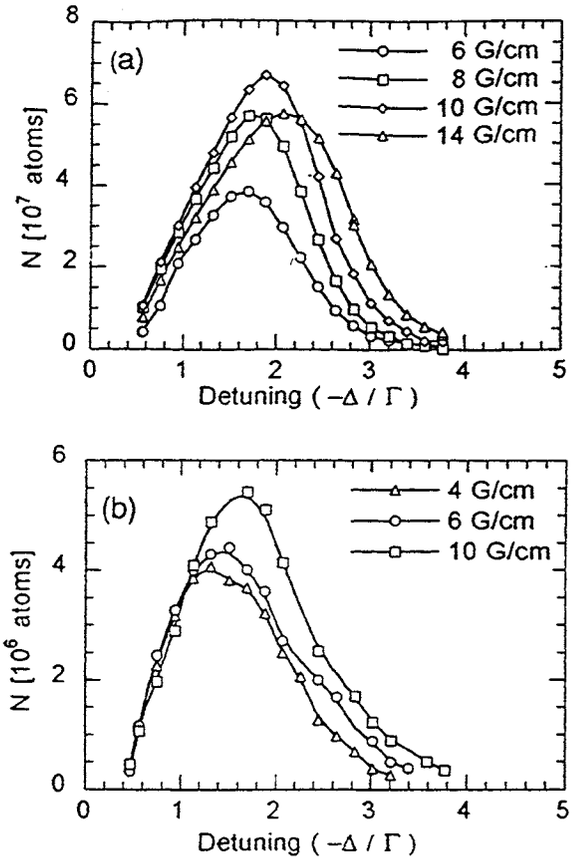


Figure 3: Number of trapped atoms *versus* laser detuning for different values of the magnetic gradient. (a) Case of a 3-D trap: the laser intensity is kept constant ($I = 1.8 I_{sat}$). (b) Case of a 2-D trap: the laser intensity is kept constant ($I = 1.4 I_{sat}$).

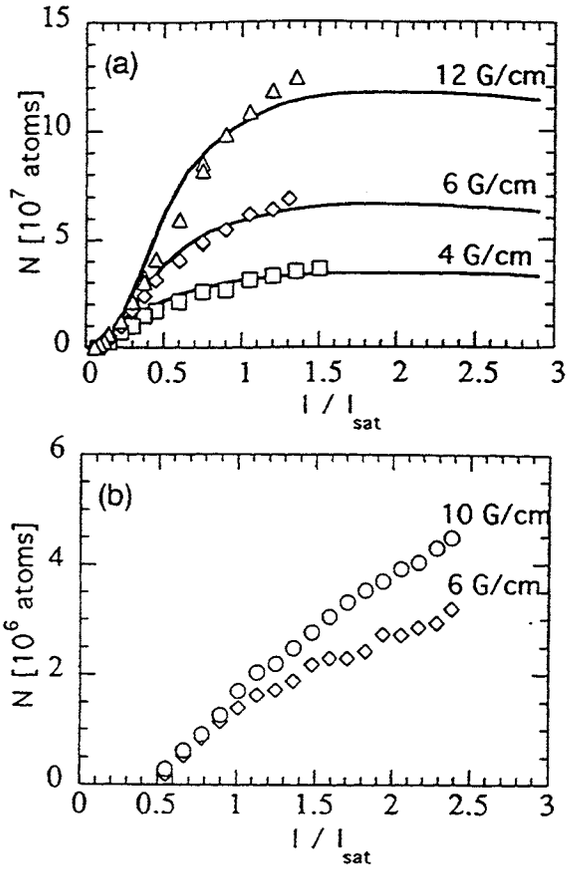


Figure 4: Number of trapped atoms *versus* cooling beams intensity for different values of the magnetic gradient. (a) Case of a 3-D trap: the symbols represent the experimental data whereas the lines are the predictions of a 1-D model multiplied by a normalization factor. For $dB_x/dx = 4, 6, 12$ G.cm^{-1}, the laser detuning is kept constant at the respective values $\Delta = -1.3, -1.7, -2\Gamma$ and the normalization factor was respectively 1.35, 2.4 and 3.81. (b) Case of a 2-D trap: the detuning is kept constant at the value $\Delta = -1.5\Gamma$.

to the transverse beams and the mixing of the population of the Zeeman sublevels by the changing magnetic field – are taken into account, the capture process and the influence of the magnetic field on it may be described by considering the on-axis case. The efficiency of the capture process is given by the capture velocity that is determined by numerically solving the Newton equation of motion, using a Doppler expression of the magneto-optical force.

To make a theoretical comparison between the number of trapped atoms in the two configurations, we study the capture velocity in the *Oxy* plane using a 2-D model. In this plane both the isotropy and the anisotropy of the magnetic field can be easily modelled.

Figure 3 shows the dependence of the number of trapped atoms on the detuning for different values of the magnetic gradient. In both magnetic configurations, the optimum detuning increases with the magnetic gradient. In the 3-D (respectively 2-D) case all the optimum detuning observed experimentally lie between -1.6 and - 2.2 Γ (resp. -1.3 and -1.7 Γ). This difference of range in the capture detuning is only related to the fact that the diameter of the cooling beams is smaller in the 2-D case. Additional measurements have shown that for identical conditions of irradiation the optimum detuning is the same. As one can see, the dependence of the number of trapped atoms on the detuning of the cooling beams is not modified when the magnetic field becomes anisotropic.

The same remark prevails for the influence of the laser intensity on the capture process that is presented in Fig. 4 for different values of the magnetic gradient. The values on the abscissa correspond to the average intensity over the $1/e^2$ diameter of each of the six cooling beams and are normalized to the saturation intensity of the $|F = 4, m_g = 4\rangle \rightarrow |F' = 5, m_e = 5\rangle$ transition ($I_{sat} = 2.2$mW cm^{-2}). In addition we represent in Fig. 4a, the predictions of the 1-D model (solid lines) that describes the 3-D magneto-optical trap. It is interesting to notice that this simple 1-D model reproduces accurately how the capture process depends on laser intensity for different values of the magnetic

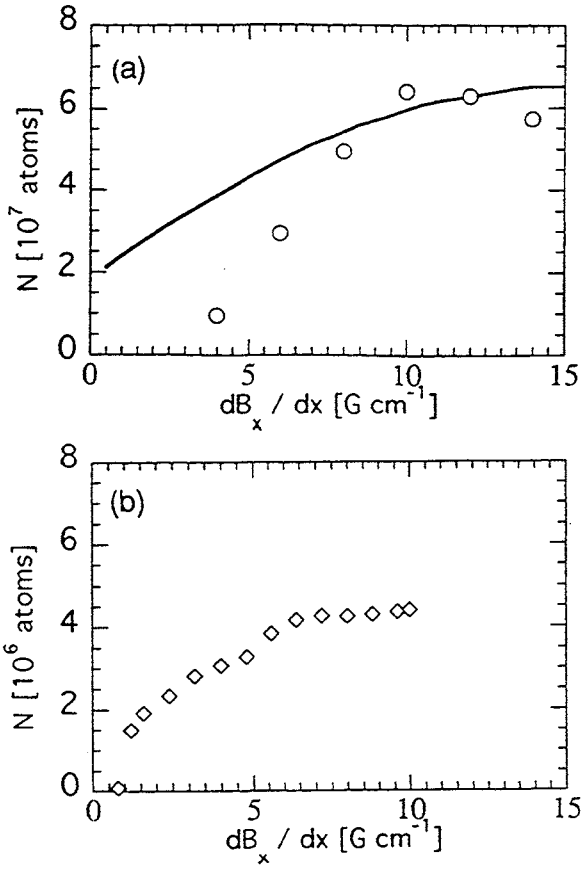


Figure 5: Number of trapped atoms versus the x-component of the magnetic field gradient. (a) Case of a 3-D trap:the laser detuning and intensity are kept constant ($\Delta = -2\Gamma, I = 1.8I_{sat}$). The symbols (o) are experimental values, the line represents the theoretical prediction. (b) case of a 2-D trap: the laser detuning and intensity are kept constant ($\Delta = -1.7\Gamma, I = 1.4I_{sat}$). The symbols (◇) are experimental values.

gradient; both the experimental results and the theoretical predictions represented in Fig. 4a shows that the number of trapped atoms saturates with the intensity.

An important difference between the capture process in the two magnetic configurations resides in the efficiency of the capture. For identical irradiation conditions and for the same value of dB_y/dy in the two configurations we found that the number of trapped atoms in the 3-D case was eight time larger than in the 2-D case. Such a factor was not expected and can not only be due to the influence that the magnetic field has by bending the atomic trajectories. As mentioned before, we compare the capture velocity of the atoms moving in the Oxz plane for the two configurations of magnetic field: we certainly found that the capture velocity decreases by a factor 1.05 when going from an isotropic 3-D to an anisotropic 2-D MOT but this can not explain such a difference in the number of trapped atoms. We believe that this difference could be due to the fact that in the 2-D trap other losses than those due to collsions between hot and cold atoms appear. Since the z-axis is free of trapping, there are in this direction phenomena such as spatial diffusion and atomic drift due to imperfections in the alignments or in the intensity balance. Meanwhile, these loss phenomena do not affect the loss rate of already trapped atoms: several measurements have shown that for identical conditions of irradiation the loading time of the two types of trap was very close (typically, 1 s for the 3-D and 0.8 s for the 2-D trap). These loss phenomena seem to expel atoms that are in the process of beeing cooled: therefore one observes a difference in the steady state number of trapped atoms, but not in the loading time. This explanation is in line with the remark one could make about Fig. 5a: the comparison between the predictions of the 1-D model and the experimental results shows that at low magnetic field gradient - when the time constant of the capture becomes comparable in magnitude with the average time between two expelling collisions - the model overestimates the capture efficiency. This difference can be partly explained by the fact that our 1-D model disregards both exclusion of atoms during the positioning time and loss phenomena other than collisions between hot and cold Cesium atoms. The great sensitivy of $\sigma^+ - \sigma^-$ molasses (or low magnetic gradient trap) to imperfections could be explained [9] by the fact that a 3-D $\sigma^+ - \sigma^-$ molasses exihibits vortices network with several leakage directions for the atoms. An imperfection in the cooling conditions leads to important atomic drift along some preferential directions. Besides this difference in the absolute number of trapped atoms, the experimental results presented in this part show that the way the capture efficiency depends on the trapping parameters is not modified by the change from a 3-D to a 2-D trap.

4 Continuous beam of cold atoms: problems and perspectives

The experimental apparatus presented in Section 2 is designed to obtain a vertical beam of cold Cesium atoms. The extraction of the atoms from a 2-D trap can be realized using three techniques. The first one was used in the first atomic fontain [2] and consists in creating an intensity imbalance between the beams propagating towards $+\vec{e_z}$ and those propagating towards $-\vec{e_z}$ that leads to a net radiation force directed along the z-axis. A second technique is to apply a static magnetic field in the z-direction: theoretical and experimental studies [10] have shown that the effect of such a magnetic field on an 1-D $\sigma^+ - \sigma^-$ molasses is to induce an atomic drift of velocity 30 cm s^{-1} G^{-1} for cold Cesium atoms. The atom can also be extracted from the 2-D trap by introducing a frequency difference $\Delta\nu$ between the beams going towards $+\vec{e_z}$ and

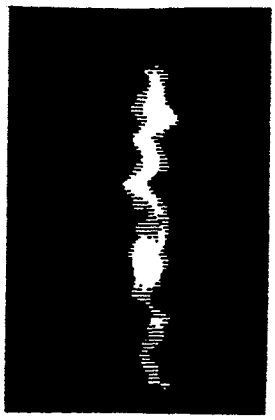


Figure 6: Image of a typical column in a two-dimensional trap.

$-\vec{e_z}$: the atoms are then cooled in a moving frame of velocity $v_z = \lambda\Delta\nu$. We used the first two techniques on the 2-D trap but the application of both an intensity imbalance up to 50% and a static magnetic of magnitude up to 1.5 G along $+\vec{e_z}$ did not allow us to extract atoms from the 2-D trap and detect them on the hot wire detector. We strongly believe that the production of a vertical beam of cold atoms was thwarted by the existence of a spurious structure in the 1-D cloud (fig.6). Indeed, the images of the CCD camera show that the atoms are arranged in a "furrow" that oscillates spatially in a irregular way around the z-axis: the amplitude of these oscillations is of the order of one millimeter. We notice that the features of the irregularities depend on the imperfections of the irradiation system such as the laser beams misalignment or unsmooth spatial profile; however, none of these imperfections is entirely responsible of these irregularities. Arguments, such as collective effects or laser beam disalignments that have been given by other groups [9, 11, 12] to explain special structures in a 3-D traps are not suitable for our experimental observations.

This problem of structure alters considerably the directivity of the cold atom beam and hence represents one of the major obstacles to the use of this beam in a an atomic frquency standard. One of the solutions actually under study is the dark channel: in the center of the trap one creates a region where no repumping light is present [13]. This channel is obtained by projecting a shadow in the repumping beam. As long as an atom is repumped, it is submitted to a cooling force in all directions and is driven to the z-axis by the 2-D trapping force in the Oxy plane. If the distance from the vertical axis becomes smaller than the radius of the dark channel, the atoms is no longer repumped, it becomes transparent to the cooling and trapping light and begins to fall under the action of gravity. Small horizontal components could bring the atom back into a repumped region. Here it would be cooled and trapped again. The dark channel could therefore be seen as a collimator with a directivity determined by its geometrical parameters.

5 Conclusion

We propose a new design for a frequency standard based on a continuous beam of cold atoms. The atomic source is a anisotropic two-dimensional magneto-optic trap with a geometry for the magnetic gradient which is different than the usual anti-Helmholtz configurattion. With this trap we have succeeded to trap vertical columns of 10^6 to 10^7 cold atoms in loading times of approximately 1s. This gives us good hope that it will be possible to extract a sufficient amount of cold atoms for a continuous beam. Inhomogeneities in the cold column clouds seem to be one of the main problems to be overcome. A dark channel in the repumping beam might be a solution to this obstacle. The extraction is the aim of our current work.

References

[1] A. Clairon *et al.*, Proceedings of the 6th EFTF, Noordwijk, The Netherlands, (1992)

[2] M.A. Kasevich, E. Riis, S. Chu, Phys. Rev. Lett. 63(6), 612 (1989)

[3] K. Gibble, S. Kasapi, S. Chu, Opt. Lett. **17**, 526 (1992)

[4] S. Ghezali *et al.*, these proceedings

[5] G.J. Dick *et al.*, Proceedings of the 22nd PTTI, Vienna, Virginia, 487 - 508 (1990)

[6] C. Monroe *et al.*, Phys. Rev. Lett. **65**, 1571 (1990)

[7] K. Lindquist *et al.*, Phys. Rev. A **46**, 4082 (1992)

[8] G. Dudle *et. al.*, Proceedings of the 8th EFTF, Munich, Germany (1994)
N. Sagna *et al.*, to be published

[9] A. Hemmerich *et al.*, Europhys. Lett. **21**(4), 445 (1993)

[10] C. Valentin *et al.*, Europhys. Lett. **17**(2), 133 (1993)

[11] D. W. Sesko, T. G. Walker, C. E. Wieman., J Opt Soc Am B **8** (5), 946 - 958 (1991)

[12] I. Guedes *et al.*, J Opt Soc Am B **11** (10), 1935 - 1940 (1991)

[13] W. Ketterle *et al.* Phys. Rev. Lett. 70(15), 2253 (1993)

Acknowledgement

This work was supported by the Swiss National Found for Scientific Research, the Swiss Federal Office of Metrology and the Swiss Federal Office of Education and Science.

SESSION 7

RESONATORS II

Chairman : J. ZELENKA

FINITE ELEMENT ANALYSIS USING TRIGONOMETRIC INTERPOLATING FUNCTIONS FOR HIGH OVERTONES THICKNESS MODES RESONATORS

Fanjaharisoa RAOELIJAONA and Bernard DULMET

Laboratoire de Chronométrie Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
Chemin de l'Epitaphe, 25030 Besançon Cedex - FRANCE

ABSTRACT

A finite elements program is developped at LCEP to model quasi-thickness modes in Piezoelectric Crystals. To efficiently render the pattern of eigenmodes, this program uses trigonometric interpolating functions along the thickness of the plate and parabolic functions along the in-plane axes, in such a way that only one layer of elements is required along the thickness. Significal advances have been made since prior publication at IEEE 1994 Symposium on Ultrasonics : the choice of shape functions for fundamental thickness modes has been improved, and we have designed new elements, with a proper set of interpolation functions along the thickness to model third overtones modes with only fours nodes along the thickness. Obtained numerical results for fundamental and 3rd overtone modes exhibit a close agreement in compare with analytical results from Tiersten's model.
First results concerning the frequency temperature behaviour have been obtained and a standard hexaedric element has been designed to validate our program comparing its results for flexural and torsional modes with results from a finite element commercial software. A simple visualization tool with help of gnuplot has been made.

I - INTRODUCTION

Finite Element Method has been proven as a powerful numerical technique for obtaining approximate solutions of differential equations subject to known boundary and initial conditions whenever available analytical methods reach their limits.
The analytical model proposed by Stevens and Tiersten [1] gives satisfactory results in compare with experiments, moreover it provides valuable information on helping us to choose our trigonometric interpolating function. Nevertheless, slight changes in the crystal shape or electrodes geometry result into serious complexity [2] to obtain accurate solutions. In this case FEM seems a flexible tool to predict resonant modes and to have an insight of temperature influence in modal analysis.
In this paper, we have improved the set of trigonometric interpolating functions presented in a previous paper [3] and aimed to the modelling of fundamental modes. For comparison we have also designed a more classical triangular prismatic element using cubic interpolations.
Moreover, we have designed new prismatic elements specifically aimed for high overtones modelling. Using them leads to neatly smaller matrices in compare with classical elements using polynomial shape functions, along every direction (increasing the degree of polynoms would reduce the number of elements layers and increase accuracy, but would not decrease the size of the problem).

All dynamic fields are mapped onto the material coordinates of a fixed, known, reference state. Thus we take advantage of this feature which facilitates the computation of the static thermal sensitivity of obtained eigenmodes, since using so-called Lagrangian formalism is efficient, and in fact mandatory, to avoid the additionnal work that would be necessary to compute the actual shape of the piezoelement along temperature variations.
Curves showing the influence of temperature on the frequencies (fundamental, third overtone) are compared with results from the analytical models and modal decomposition [4].

II - BASIC EQUATIONS AND MATRIX FORM OF VARIATIONAL FORMULATION

Every finite element program is built upon an integral formulation. We found useful and even necessary to base our program consistently with respect to the general theory for thermo-electro-elasticity as exposed by Tiersten [5], although the present development state of our program is only aimed to the solution of linear problems. Thus, we found that the following variational formulation is pertinent for the class of problems considered:

$$\int_{t_0}^{t} dt\Big[\int_{V_o}(K_{L\alpha}\delta U_{\alpha,L} + D_L\delta\phi_{,L} + \rho_o\omega^2 U_\alpha\delta U_\alpha)dV\Big] - \int_{t_0}^{t} dt\Big[\int_{S_o}(F_\alpha\delta U_\alpha + \sigma\delta\phi)dS\Big] = 0 \qquad (1)$$

where all geometric characteristics are considered at some fixed, stress free, reference state, that we may arbitrarily define as the natural state of the crystal at 25°C, $\boldsymbol{K}$ is the Piola Kirchhoff stress tensor and $\boldsymbol{D}$ the material electric displacement respectively. ρ_O, V_O, S_O denote the mass density, volume and outer surface of the considered crystal in the natural state. $\boldsymbol{F}$ is some applied static force and σ the net charge on some part of the outer surface. Because finite elements method is known to require much computational power to get eigenmodes of volumic structures and it is of prime interest to be able to predict at least the static thermal behaviour of those modes, the use of lagrangian formulation referred to a fixed natural state is in fact mandatory since it is the only means to avoid the need of computing the actual shape of the piezoelement along the temperature variations. Then we use the linearized following behaviour laws:

$$\begin{cases} K_{L\alpha} = G_{L\gamma M\varepsilon} U_{\varepsilon,M} + R_{A\alpha L}\phi_{,A} \\ D_A = -R_{A\alpha L} U_{\alpha,L} + N_{AL}\phi_{,L} \end{cases} \quad (2)$$

where G, R, N respectively denote the sets of so-called effective material elastic piezoelectric and dielectric constants[3].

Since at this moment we only consider eigenmodes in non-dissipative media, we shall not compute the current output of the vibrating piezoelement submitted to time-harmonic driving forces and voltages. So, the surface integral vanishes in Eq(1). Also we did not account for the electrostatic equations in the surrounding medium. Hence Eq(1) implicitly incorporates the following simple boundary conditions:

$$\left.\begin{matrix} n_L K_{L\gamma} = 0 \\ n_L D_L = 0 \end{matrix}\right\} \text{ on the surface} \quad (3)$$

except at the nodes where either U or ϕ is imposed to zero.

In order to turn our problem into a matrix form suitable for FEM we use the compressed matrix notation instead of the extended tensorial notation. As usual in formalisms using Piola-Kirchhoff stress tensor and material electric displacement, index compression turns the elastic coefficients and piezoelectrics coefficients matrices into 9 by 9 and 3 by 9 matrices, respectively (see Ref [6]) In such analysis the independent field variables to consider are the potential and mechanical displacement gradients with respect to the material coordinates of the body in a well defined, known, reference state

$$\begin{aligned} \langle \phi' \rangle &= \langle \phi_{,1} | \phi_{,2} | \phi_{,3} \rangle \\ \langle U' \rangle &= \langle U_{1,1} | U_{2,2} | U_{3,3} | U_{2,3} | U_{3,1} | U_{1,2} \\ &\qquad U_{3,2} | U_{1,3} | U_{2,1} \rangle \\ \langle U \rangle &= \langle U_1 | U_2 | U_3 \rangle \end{aligned} \quad (4)$$

Thus, we obtain the matrix form of the variational formulation suitable for the considered class of problems:

$$\begin{aligned} &\int_V \left(\langle \delta U' \rangle [G] \{U'\} + \langle \delta U' \rangle [R]^t \{\phi'\} \right) dV \\ &+ \int_V \left(\langle \delta \phi' \rangle [R] \{U'\} - \langle \delta \phi' \rangle [N] \{\phi'\} \right) dV \\ &\qquad + \int_V \rho_o \omega^2 \langle \delta U \rangle \{U\} dV = 0 \end{aligned} \quad (5)$$

III - DISCRETISATION OF THE PROBLEM AND SHAPE FUNCTION CONSIDERATION

In this section we describe the next step of finite elements analysis, "the discretisation process", which allows us to reduce the infinite numbers of degrees of freedom of a continuum to a finite value. In this purpose every continuously-varying unknown field is interpolated in each element from its values at the nodes, most often located on the boundary of the elements. Shared nodes at the common interface of two elements connect the field representation from a given element to the next one, and so on.

For instance, we mesh the resonator with 24-nodes finite elements consisting of four layers of nodes along the thickness upon a six-nodes triangular basis. This kind of element is shown in Fig 1:

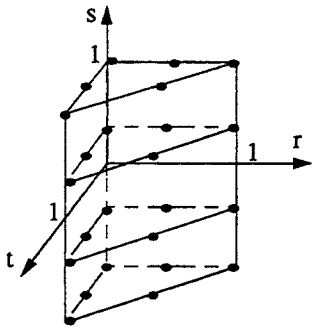


Fig 1: 24 nodes-prismatic element

The mechanical displacement and electrical potential may be calculated by nodal interpolation defined in a given element:

$$\begin{cases} U_k = \sum_{j=1}^{n} H_j(r,s,t)\, U_k^j \quad (k = 1,2,3) \\ \phi = \sum_{j=1}^{n} H_j(r,s,t)\, \phi^j \end{cases} \quad (6)$$

n denotes the number of nodes within an element, H_j denotes the shape function, U_j and φ_j respectively denote the mechanical displacement and electrical potential at the nodes.

Concerning the shape functions, because standard elements available in commercial packages take low degree polynomial interpolations, many layers of such elements are required to take into account the fast trigonometric variations of mechanical displacement along the thickness. This results into a loss of efficiency and very large matrices requiring much computational power. For instance it is not possible to simulate thickness modes with an overtone number higher than three on our reasonably sized hardware (Unix workstation with 112Mb RAM). To address this problem, we found necessary to adapt element and shape functions according to sought modes. And by definition, the thickness vibrations have an oscillating behaviour along the thickness. Thus to correctly fit this behaviour, we propose prismatic element and trigonometric interpolations along the thickness in such a way that only one layer of elements is required for each sought mode.

First, we consider the case of the fundamental modes. We have found that the interpolating functions basis proposed in a previous paper [3] does not provide with a correct ordering of modes. We presently obtain far better results with the following basis along the thickness:

$$\langle P_s \rangle = \langle 1 \ \sin(\pi s/2) \ \sin(\pi s) \ \cos(\pi s) \rangle \quad (7)$$

Of course, the number of terms of P should match the

number of nodes along the thickness. To simulate third overtone modes we have considered the following basis :

$$\langle P_s \rangle = \langle 1 \ \sin(3\pi s/2) \ \sin(\pi s) \ \cos(\pi s) \rangle \quad (8)$$

Along the in-plane axes, we have taken the usual Lagrange polynomial functions [7].
Consider that the basis along the plane *(ot,or)* is $<P_{t,r}>$. Thus we can compute the overall basis P :

$$\langle P \rangle = \langle P_{t,r} \rangle \langle P_s \rangle \quad (9)$$

This allows us to build the shape functions according to the well known formula[7] :

$$\langle H \rangle = \langle P \rangle [P_n]^{-1} \quad (10)$$

where P_n is the matrix of the basis values at every nodes of reference element.
We should remark that our formulation is not isoparametric since we use different (polynomial) shape functions to define the geometry of the element. Thus we have the following formula :

$$\langle x_k \rangle = \sum_{j=1}^{n} \bar{H}_j(r,s,t) x_k^j \qquad k = 1,2,3 \quad (11)$$

All considered shape functions are expressed in term of dimensionless parametric coordinates *r, s, t* which vary from [0,1], [-1,1], [0,1]. In fact this system of coordinates allows a straightforward numerical integration to evaluate any element integrals which are required during the computation of stiffness and mass matrices.

To elegantly include the static temperature influence in the modal analysis, the components of the gradient of mechanical displacement and electrical potential must be listed in the following vectors

$$\langle U' \rangle = \langle U_{1,1} \, U_{2,2} \, U_{3,3} \, U_{2,3} \, U_{3,1} \, U_{1,2} \, U_{3,2} \, U_{1,3} \, U_{2,1} \rangle$$
$$\langle \phi' \rangle = \langle \phi_{,1} \, \phi_{,2} \, \phi_{,3} \rangle$$

The gradient of mechanical displacements (resp. electrical potential)-displacement (resp. potential) relationships in accordance with the equations (9) can be written in the following form :

$$\begin{aligned} \{U'\} &= [B]_{9X96} \{U^n\} \\ \{\phi'\} &= [D]_{3X96} \{U^n\} \\ \{U\} &= [T]_{3X96} \{U^n\} \end{aligned} \quad (12)$$

where

$$\{U^n\} = \{U_1^1 \, U_2^1 \, U_3^1 \, \phi^1 \ldots U_1^{24} \, U_2^{24} \, U_3^{24} \, \phi^{24}\}$$

$$[B] = \begin{bmatrix} H_{1,1}^1 & 0 & 0 & 0 & \ldots \\ 0 & H_{2,2}^1 & 0 & 0 & \ldots \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & H_{1,2}^1 & 0 & 0 & \ldots \end{bmatrix}_{9X96_{(4X24)}}$$

$$[D] = \begin{bmatrix} 0 & 0 & 0 & H_{4,1}^1 & .. \\ 0 & 0 & 0 & H_{4,2}^1 & .. \\ 0 & 0 & 0 & H_{4,3}^1 & .. \end{bmatrix}_{3X96_{(4X24)}}$$

$$[T] = \begin{bmatrix} H_1^1 & 0 & 0 & 0 & .. \\ 0 & H_2^1 & 0 & 0 & .. \\ 0 & 0 & H_3^1 & 0 & .. \end{bmatrix}_{3X96_{(4X24)}}$$

The whole treatment from Eq(6) yields the following eigenvalue matrix equation :

$$KU = \lambda MU \quad (13)$$

where K and M are the overall stiffness and mass matrices obtained by assembling all elementary stiffness and mass matrices :

$$\begin{aligned} [K^e] &= \int_{V^e} \big([B]^t [G][B] + [B]^t [R]^t [D] \\ &\qquad + [D]^t [R][B] - [D]^t [N][D] \big) dV \\ [M^e] &= \int_{V^e} \rho_o [T]^t [T] dV \end{aligned} \quad (14)$$

In order to evaluate such integrals, we adopt the well-known Gauss-Legendre numerical integration procedure. This method consists of approximating the definite integral of a function *f(x)* by a properly weighted sum of particular values of *f(aj)* suitably distributed within the limits of integration[8] :

$$\begin{aligned} \int_V f(x) dV &= \sum_{j=1}^{n} \omega_j \, f(a^j) \\ x(x_1, x_2, x_3) & \\ a^j(a_1^j, a_2^j, a_3^j) & \end{aligned} \quad (15)$$

n denotes the number of Gauss' points and a^j denotes Gauss points.
It should be mentionned that the special features of our shape functions require to pay some careful attention to the numerical integration. At the difference of the case of polynomial interpolations, the derivatives involved by the gradients of displacement appearing in B (hence in K) do not improve the integration accuracy of K in compare with M.
In our case we found that taking twelve Gauss' points along the thickness and four along the plane *(ot,or)*, allows us to have a sufficient approximation of our integrals.

IV - OVERALL STRUCTURE OF THE FINITE ELEMENTS PROGRAM

Our program has the same structure as an usual finite elements program. All routines are written in C language. We can summarize the main steps of this program as follows :

- generates automatically the mesh of the resonator, it provides the set of nodal coordinates and the elements constituting the mesh
- reads data and evaluates stiffness and mass matrices, writes them on the output file
- renumbers nodes in order to minimise the stiffness and mass matrices-bandwidth
- assembles elementary stiffness and mass matrices into K and M
- solves the eigenproblem $KU=\lambda MU$ with subspace iteration method
- visualizes the results

In order to predict the interest modes, a post-treatement is designed to visualize the results. It first generates data from the finite elements calculation. Those data are thereafter visualized with help of a common graphical software.
It should be noted that the stiffness and mass matrices are symmetric, therefore it is only necessary to store the upper triangular part of them. In this way the memory requirements for storage of the global stiffness equations is approximately halved. We could even more reduce the memory requirements upon adopting a skyline storage: it consists of only storing in memory the part of each matrix column which is located between the diagonal and the farthest non zero element. Thus, for the cases that we considered in practice, the memory requirements are roughly reduced to one tenth of the size of the initial full matrix.

V - NUMERICAL RESULTS

We considered two classes of structures : fundamental and third overtones AT resonators. Also, we have designed an isoparametric hexaedric element with 20 nodes, suitable for parabolic interpolation along any direction. It was built to simulate flexural and torsional vibrations of a piezoelectric bar. Putting all piezoelectric coefficients to zero, we have obtained exactly the same frequencies and eigenmodes as provided by SYSTUS a commercial software. This appears as a good checking of the overall functionning of our program. We will not further discuss the results obtained from this hexaedric element, since they are out of the scope of the present paper.

A - Fundamental mode

We have investigated an AT-cut resonator with the following geometry:

AT 35°25'
Thickness 0.6878mm
Radius of curvature 106cm
disk diameter 20mm

To treat fundamental thickness modes problem, our program proposes two kinds of interpolation functions along the thickness: functions obtained from the trigonometric basis (8) and from the simple cubic basis :

$$\langle P_s \rangle = \langle 1 \; s \; s^2 \; s^3 \rangle \qquad (16)$$

Both kind of shape functions are used in 24-nodes prismatic elements (Fig 1) previously described. Table1 shows comparative frequencies obtained from both sets of interpolations, analytical model from Stevens and Tiersten and modal decomposition model from Dulmet.

	frequencies - 25°C (kHz)			
	FEM 1	FEM 2	MDM	AM
1 0 0	2450.7	2452.5	2446.6	2441.4
1 0 1	2481.0	2481.3	2475.5	2470.8
1 1 0	2487.0	2486.0	2479.1	2478.8
1 0 2	2510.7	2499.6	2504.0	2500.1
1 2 0	2525.2	2538.6	2511.2	2516.9

Table 1

FEM 1 : trigonometric interpolations
FEM 2 : polynomial interpolations
AM : analytical methods
MDM : Modal decomposition method

All models propose the same mode ordering. As an example, we give on Fig 2 the U1 pattern in lower surface of the plate. The mesh appearing in this figure is only made for visualisation purpose. The finite element mesh that we have used for all FEM analyses presented in this paper actually consists of one layer of triangular prismatic elements regurlarly distributed inside four concentric circular regions, as shown in Fig3. So we get only 868 nodes and 96 elements, which represents a very reasonably sized analysis.
The considered geometry corresponds to a contoured resonator with shallow contour. We have not yet designed two dimensionnal elements that could be used to combine the modelling of trapping effects due to contouring and mass loading, as was analytically done at the reference [2]. In this respect, the considered geometry is not perfectly appropriated to validate our program. On figure 2, it is interesting to observe that the residual amplitude at the disk edge is not really zero.

B - Third overtone mode

In the case of third overtone, we considered the design of a 10Mhz AT resonator with the following geometry:

AT 35°27'
Thickness 0.5mm
Radius of curvature 200mm
disk diameter 15mm

This geometry corresponds to a very stronger trapping than for the considered fundamental mode geometry. According to the serious size of the analysis that would have been required piling up three layers of prismatic elements with cubic interpolations, we could only test

trigonometric interpolation defined by (9), and we propose comparative results at table2.

	Frequencies 25°C (kHz)			
n m p	32 points	48 points	MDM	AM
3 0 0	10024.4	10036.4	10049.3	10046.8
3 0 1	10126.4	10137.9	10128.6	10126.8
3 1 0	10128.5	10146.1	10131.4	10130.8
3 0 2	10228.7	10240.8	10206.9	10206.1
3 2 0	10234.1	10248.3	10212.8	10214.1

Table 2

It is worth remarking that the number of Gauss integration points has a sensible influence on the frequency, as discussed at the end of the section III. Corresponding mode patterns for harmonic modes are presented on Fig4 and Fig5.
These results have been obtained for an electroded plate.

	Mode (3 0 0)	
	short-circuit	open-circuit
25°C	10036425	10052841.2
0°C	10037157.9	10053566.4
50°C	10035661.5	10052071.9
100°C	10034468.3	10051239.7

Table 3

If we set to zero the potential in a central region, hence simulating plate with short-circuit, zero thickness, electrodes, we found a frequency lowering roughly amounting to 16 Khz, which is too large a value. In the nearest future, we intend to use different interpolations along the thickness for the potential and the mechanical displacement to address this problem.
For the same geometry but 35°17' orientation angle an analysis was done to predict the turnover points. The results of this analysis are presented at the table 4 and the Fig 5b.

	Mode(3 0 0)
	open-circuit
25°C	10050661.6
0°C	10051164.8
50°C	10050110.6
100°C	10049495.4

Table 4

Fig5a and 5b show the frequency-temperature curves we predict for (3 0 0) C-mode. Those curves are obtained by a polynomial regression from a set of four frequencies at 0°, 25°, 50° and 100°. At the given θ angles, we obtain too high turnover points. To obtain more realistic value of the turnover it is at the moment necessary to reduce the θ angle by roughly 10-15 minutes. We plan to address this problem which might have various reasons, in further work.

VI - CONCLUSION

The overall functioning of our program is now well established. It is already able to investigate third overtones. The inner design is made to facilitate a fast implementation of a large range of specific elements. The main interest in short term will be the possibility of modelling the influence of arbitrarily shaped electrodes, as well as defects in their positioning.
Further work still remains necessary to check the validity of this program for the modelling of doubly rotated cuts and the improvement of temperature-frequency sensitivity.

REFERENCES

[1] D.S. Stevens and H.F. Tiersten
"An analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation" J.AC.Soc of AM, V79 n°6 pp 1811-1826 (1986).

[2] H.F. Tiersten, B.J. Lwo and B.Dulmet
"An analysis of transversely varying thichkness modes in trapped energy resonators with shallow contours" Rensselear Polytechnic Institute, Troy, Ny 1280 - 3590.

[3] F. Raoelijaona and B. Dulmet
"Finite Elements analysis using Trigonometric interpolations for quasi-thickness piezoeletric resonators" IEEE 1994 Ultrasonics Symposium.

[4] B. Dulmet
"Contribution à l'étude des vibrations à énergie piégée des résonateurs piézoélectriques à ondes de volume" Thesis Besançon n°243 - 1991

[5] H.F. Tiersten
"On the non-linear equations of thermo-electro-elasticity" Int.J.Eng.Sci 9, 587, 1971.

[6] B. Dulmet and R. Bourquin
"Influence of the definition of material constants in study of crystal plates thermal behaviour" Proc 3rd EFTF pp 305 - 312, 1989

[7] G. Dhatt, G. Touzot
"Une présentation de la méthode des éléments finis" Maloine S.A Editeur Paris - 1984

[8] Z. Kopal
"Numerical analysis - Mechanical quadratures" Chapman & Hall Ltd - 1961

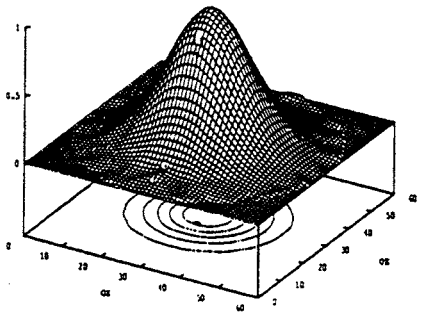


Fig 2 : Mode (1 0 0)

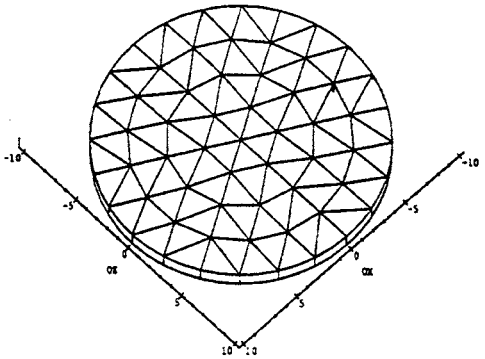


Fig 3 : Mesh of the resonator

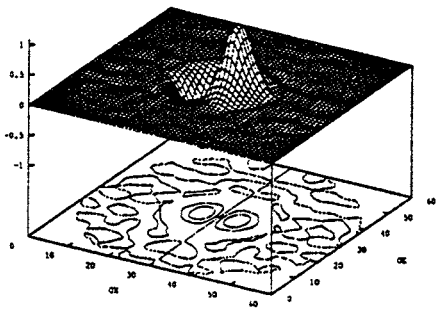


Fig 4 :Mode (3 1 0)

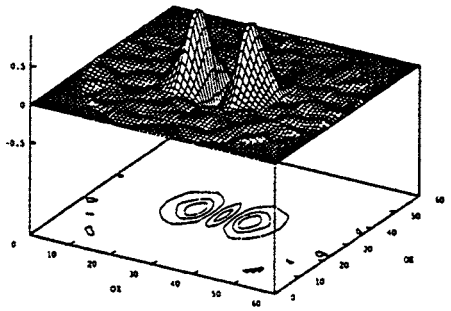


Fig 4 :Mode (3 2 0)

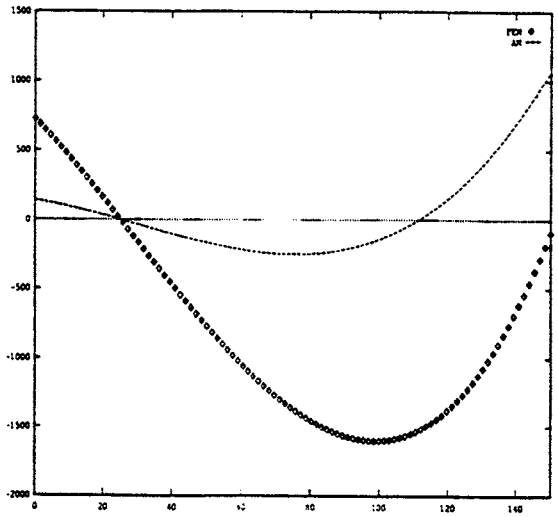


Fig 5a: Fequence Temperature 35°27' C-mode (3 0 0)

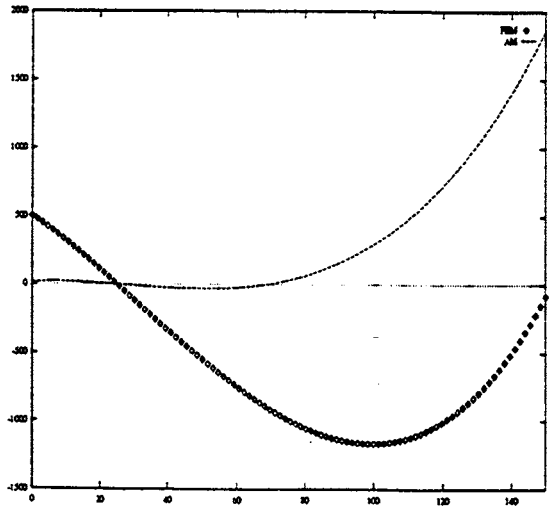


Fig 5b : Fequence Temperature 35°17' C-mode (3 0 0)

TORSIONAL DOUBLE-ENDED TUNING FORK QUARTZ CRYSTAL FORCE SENSOR

P. TRUCHOT AND R. BOURQUIN

Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
25030 Besançon Cedex, France

SUMMARY :

The sensitivity of the resonant frequency of a rectangular bar vibrating in torsion mode is analytically calculated by using the Rayleigh method in the case of a clamped-clamped bar. A study by finite element method enables to take into account of the influence of the base shape of a single or double bar resonators. Experimental results are gived for resonators of small sizes having a cut-angle (single rotation) of 32-34° about the X axis and made by chemical etching.

I. INTRODUCTION

Bar vibrating in torsional mode can be used to make a force to frequency transducer. The change of resonant frequency, when a force (traction or compression) is exerted along the length of the bar, is due to an apparent variation of the constant of torsion of the bar. This variation is coming from a "geometric effect" : the increase of the stress in the resonator when the fibers, stretched by the vibration, are submitted to an additional static tension. (The same phenomena cause the variation of frequency in a bar vibrating in flexure mode).

Whole resonant device is not restricted to the bar itself , but also include the fixations of it. Often, practical realisations use double-ended tuning fork. Analytical calculation of the force sensitivity of the sensor can be made only in the case of simple assumptions about the boundaries conditions. More realistic predictions can be obtained by using finite elements method. By this way it is possible to investigate several boundaries conditions and to take fully into account of the anisotropy of the material.

To make a sensor of practical use, it is necessary to minimize the temperature dependance of the resonant frequency of the resonator. In the case of the torsional resonator, this may be achieved by correctly choosing the orientation of the bar and the width to thickness ratio. This last parameter allows to adjust the frequency versus temperature curve of the resonator.

II FORCE SENSITIVITY OF A TORSIONAL RESONATOR

In a general way, the frequency f of a resonator submitted to a static stress $T_{ij}^{(0)}$ can be computed by the Rayleigh method [1][2]:

$$f^2 = \frac{1}{4\pi^2}\frac{U+U_T}{K} \qquad (1)$$

U is the "internal" potential energy due to the vibration when no static stress is exerted, and :

$$U_T = \int_v T_{ij}^{(0)} S_{ij}^{(2)} dV \qquad (2)$$

is the amount of energy resulting from the static stress. $S_{ij}^{(2)}$ is the non linear (quadratic) part of the deformation arising from the vibration. If u_i are the dynamics displacements between the static constrained state and the final state resulting from the superimposition of a vibration to the preceding state, we have

$$S_{ij}^{(2)} = \frac{1}{2}\frac{\partial u_m}{\partial x_i}\frac{\partial u_m}{\partial x_j} \qquad (3)$$

K is the kinetic energy.

The here considered resonator is shown by the Figure 1.

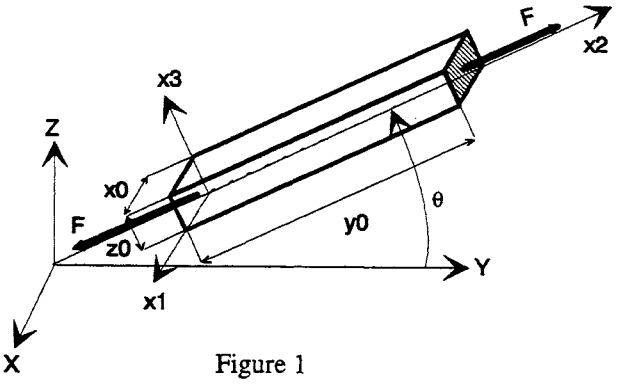


Figure 1

The mechanical displacements corresponding to the torsion are given by :

$$u_1 = -\phi(x_2,t).x_3$$
$$u_2 = \frac{\partial\phi(x_2,t)}{\partial x_2}.\psi(x_1,x_3) \qquad (4)$$
$$u_3 = \phi(x_2,t).x_1$$

where $\phi(x_2,t)$ is the rotation angle of the beam section and $\psi(x_1,x_3)$ is the St Venant's torsion function, $[x_1, x_2\ x_3]$ is the rotated frame corresponding to a particular orientation of the resonator. If in this frame, the elastic coefficients C_{46} may be neglected, the ψ function take the form [3] :

$$\psi(x_1,x_3) = -x_1.x_3 + \frac{8{z_0}^2}{\pi^3 K_s}.$$

$$\sum_{i=0}^{\infty} \frac{(-1)^n}{(2i+1)^3} \frac{\sinh((2i+1)\pi K_s x_1 / z_0)}{\cosh((2i+1)\pi K_s x_0 / 2z_0)} \sin((2i+1)\pi x_3 / z_0) \tag{5}$$

with $$K_s = (s_{66} / s_{44})^{1/2} \tag{6}$$

in the case of the rectangular beam depicted in Figure 1. If we assume that both ends of the beam are rigidly clamped, we have :

$$\phi(x_2,t) = A \sin(\frac{n\pi x_2}{y_0}) e^{j\omega t} \quad , \text{n = odd} \tag{7}$$

as solution of the equation governing the behaviour of the resonator when no static force F is exerted :

$$\rho I_0 \frac{\partial^2 \phi(x_2,t)}{\partial t^2} = C_t \frac{\partial^2 \phi(x_2,t)}{\partial x_2^2} \tag{8}$$

where $$I_0 = \frac{x_0 z_0 (x_0^2 + z_0^2)}{12} \tag{9}$$

is the moment of inertia of a bar section, and

$$C_t = \frac{z_0^4}{3 s_{66} r}\left[1 - \frac{192 r}{\pi^5 K_s} \sum_{i=0}^{\infty} \frac{1}{(2i+1)^5} \tanh\frac{(2i+1)\pi K_s}{2r}\right] \tag{10}$$

the constant of torsion with $r = \frac{z_0}{x_0}$.

Thus, for the torsional resonator, the internal energy U is:

$$U = \frac{1}{2}\int_0^{y_0} C_t \left(\frac{\partial \phi}{\partial x_2}\right)^2 dx_2 = \frac{C_t}{4}\frac{n^2\pi^2 A^2}{y_0} \tag{11}$$

and if we consider that the static stress is uniform in the whole bar, only $T_{22}^{(0)}$ is non zero. As $S_{22}^{(2)}$ we have :

$$2S_{22}^{(2)} = (x_1^2 + x_3^2)\left(\frac{\partial \phi}{\partial x_2}\right)^2 + \psi^2 \left(\frac{\partial^2 \phi}{\partial x_2^2}\right)^2 \tag{12}$$

The last term in the right side contains terms of fourth degree in x1 or x3. Its contribution in the integral (2) is negligible. So U_T can be writed with a good approximation as :

$$U_T = \int_{-x_0/2}^{x_0/2}\int_0^{y_0}\int_{-z_0/2}^{z_0/2} \frac{F}{x_0 z_0}\frac{x_1^2 + x_3^2}{2}\frac{n^2\pi^2 A^2}{y_0^2}\cos^2(\frac{n\pi x_2}{y_0})dV$$

$$= \frac{1}{48}\frac{n^2\pi^2 A^2}{y_0}(x_0^2 + z_0^2)F \tag{13}$$

For the kinetic energy we have :

$$K = \frac{1}{2}\rho I_0 \int_0^{y_0}\left(\frac{\partial \phi}{\partial t}\right)^2 dx_2 = \rho y_0 I_0 \frac{\omega^2 A^2}{4} \tag{14}$$

By substituting in (1), we obtain

$$f^2 = \frac{n^2 C_t}{4\rho I_0 y_0^2}\left(1 + \frac{x_0^2 + z_0^2}{12 C_t} F\right) \tag{15}$$

whence the relative sensitivity K_f of the torsional resonator

$$K_f = \frac{1}{F}\frac{\Delta f}{f_0} = \frac{z_0^2}{24 C_t}\frac{1 + r^2}{r^2} \tag{16}$$

The Figure 2 shows the dependance of $K_f . z_0^2$ as a function of the width to thickness ratio r in the case of a resonator having a cut angle of 34°.

The accuracy of (16) is restricted by the validity of the expression (12) of the deformation of the resonator. If they are additional deformations coming from other excited modes in the bar, the force sensitivity may be significantly modified.

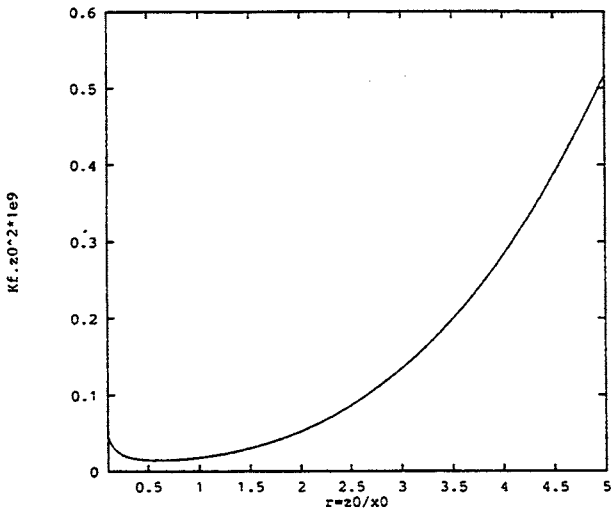


Figure 2: $K_f . z_0^2$ vs $r = \frac{z_0}{x_0}$

III. FINITE ELEMENTS STUDIES

The main purpose of the finite elements studies is to find the force to frequency sensitivity of the whole sensors when they are submitted to a traction force with realistic shape. In particular, they allow to investigate several shapes of attachments. In addition, they give informations about the modes of vibration.

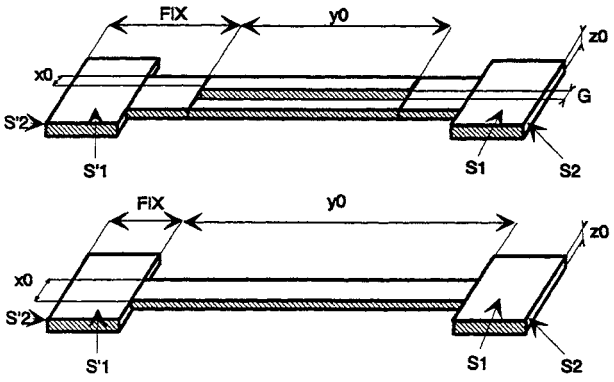


Figure 3

Double beam

The drawing of a double ended fork quartz sensor, with its attachments, is shown on the Figure 3. The analysis by finite elements method was performed by using Systus code. Linear elements have been used. The Figure 4 shows the deformed state of such a device when it vibrates in the fundamental mode of torsion. The boundaries conditions are such that the displacements at the ends S2 and S'2 of the bases are zero, excepted the u_2 displacement at the end where the static traction force is exerted. The moments exerted by the two beams on the bases introduce a deformation of them. So the ends of the beams are not perfectly clamped and the mode of vibration of the resonator is not a pure torsion mode. The displacements along the width at the end of a beam (at $x_2 = x_3 = 0$) are shown on the Figure 5a and those along the midline of the upperface (at $x_1 = 0$, $x_3 = z_0/2$) on the Figure 5b. The u_3 displacement on this last figure correspond to a flexure mode of high order.

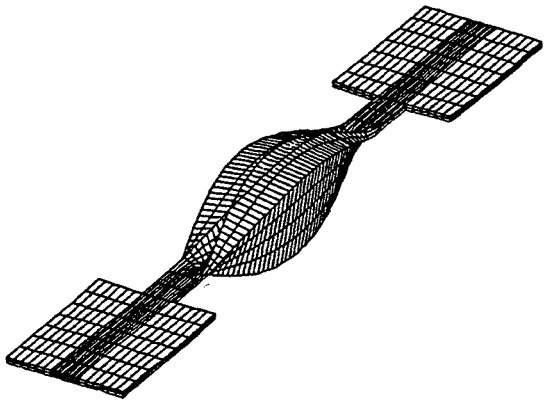


Figure 4: fundamental mode of torsion for a double beam

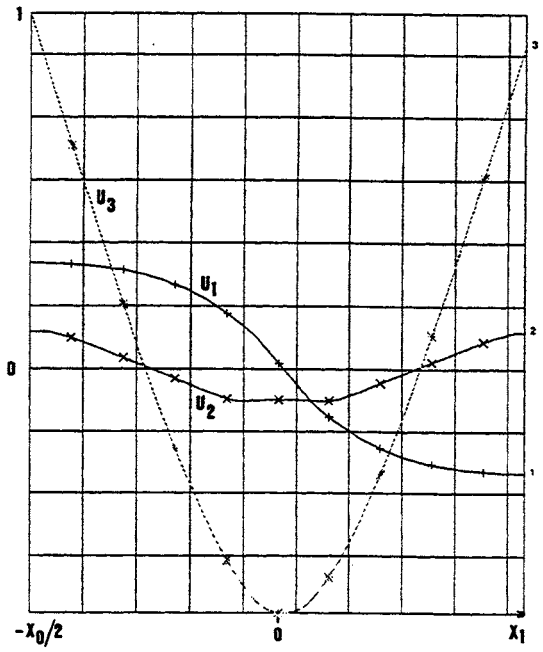


figure 5a: displacements u1, u2, u3 at the end of the beam for a width to thickness ratio r=0.3

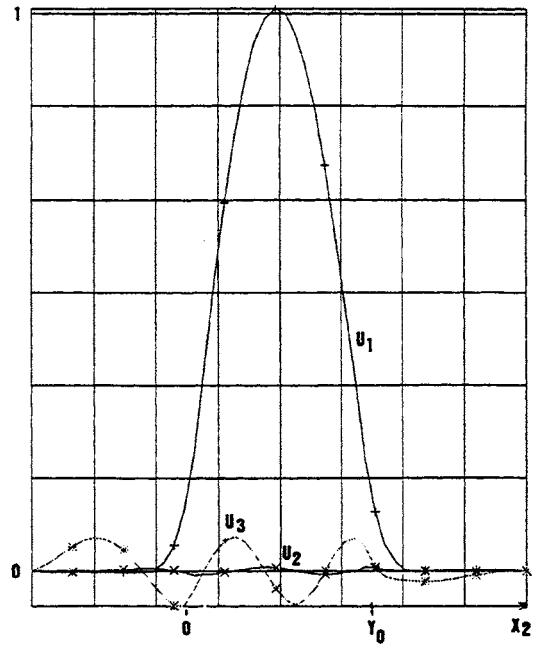


Figure 5b: displacements u1, u2, u3 along the midline of the uppersurface (r=0.3)

Single beam

In order to make comparisons with the analytical model, we have also studied a single beam sensor. The Figure 6 shows its deformed state. In this case, the bases exhibit only a small deformation (rotation) when the same boundaries conditions than above are used (Figure 7a) and the displacements along the midline of the upperface of the beam (Figure 7b) correspond to a pure torsion mode.
The fact that the boundaries conditions are not exactly the same at both ends implies a dissymmetry of the vibration along the length of the beam.
One can note that the deformations of the bases are smaller in the double beam, because the two beams vibrate in opposite phase, than in the single beam. For this reason, the double beam is preferred for practical realisations so as to have a higher Q factor.

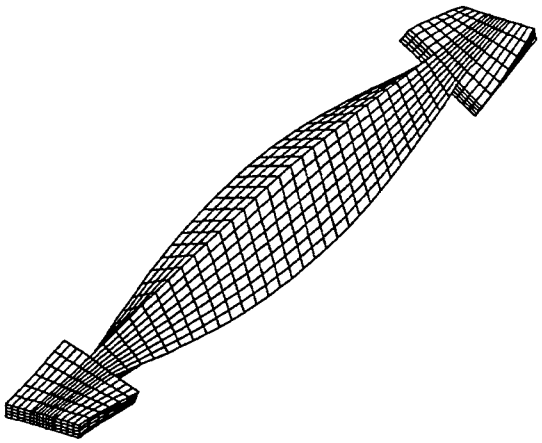


Figure 6: fundamental mode of torsion for a single beam

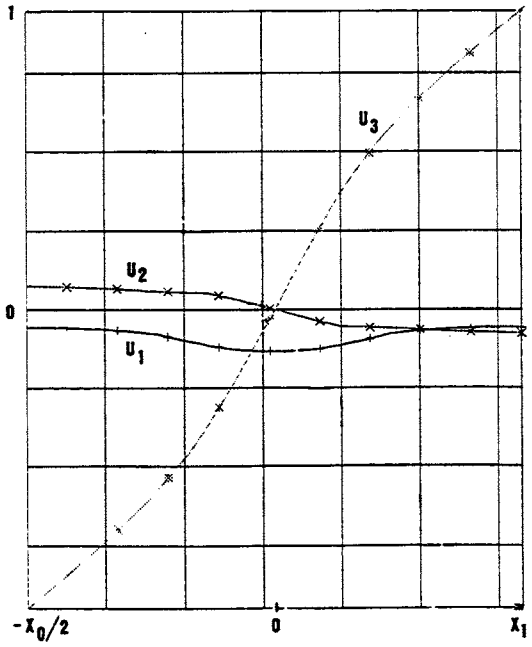


Figure 7a : displacements u1, u2, u3 at the end of the beam (r=0.3)

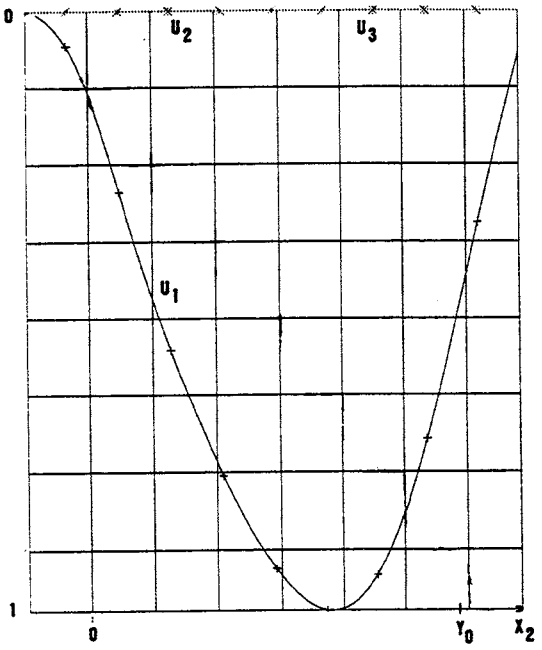


Figure 7b: displacements u1, u2, u3 along the midline of the uppersurface (r=0.3)

IV. EXPERIMENTAL RESULTS

Tests were performed on double ended fork devices having the following characteristics :

- cut angle : 32° or 34°
- thickness : z_0 = 127 microns
- width : x_0 = 250, 315 or 420 microns
- distance between beam : G = 110 microns
- length : y_0 = 6 mm
- length of fixation : FIX = 5 mm

These devices have been made by chemical etching. Their force sensitivities K_f where measured by using an electromagnetic force generator to apply a traction in the range 0 - 0.5N.

The table 1 gives the measured values of Kf, the resonant frequencies of the double ended tuning fork and the corresponding ones obtained by finite elements method. We have also included results for a single beam device. Analytical values are obtained from (16) and (15). In finite elements studies, the two values showed correspond to two different boundaries conditions : a/ a null displacement is imposed on the main surfaces (S1 and S'1) of the bases (excepted the u2 displacement on the base where the static force is exerted), b/ the fixed surfaces (S2 and S'2) are those at the extremities of the resonator (same condition than for the double beam device).
Some remarks may be done :

- the two factor between the Kf of double and single beam,

		Single beam			Double ended tuning fork		
			Finite elements				
	X_o (micron)	Analytical	a/	b/	Finite elements	Experimental	Cut-angle
K_f	250	$9.27.10^{-4}$	$8.89.10^{-4}$	$10.31.10^{-4}$	$4.29.10^{-4}$	$4.28.10^{-4}$	32°
	315	$9.99.10^{-4}$	$9.50.10^{-4}$	$11.56.10^{-4}$	$4.57.10^{-4}$	$5.37.10^{-4}$	34°
	420	$11.6.10^{-4}$	$10.83.10^{-4}$	14.10^{-4}	$4.95.10^{-4}$	/	32°
Frequency	250	211425	216598	188304	204100	204400	32°
	315	181083	186852	160186	170640	173300	34°
	420	144124	152393	126490	137050	/	32°

Table 1

- the experimental error in the Kf measurement is about 10%, due to the relative imprecision of the knowledge of the applied force and some variations in the width of the devices. No experimental values are available for x_0=420 microns because a bad driving of the torsion with the used shape of electrodes. Agreement between predicted and experimental values is fairly good,

- in the case of single beam sensor, conditions of the a/ case are quasi identical than those of the analytical model, numerical values are quite equal. In the b/ case, the effective length of the resonator is higher than in the a/ case, so the frequencies are lower and Kf is higher since the frequency variation coming from the applied force is the same in both cases.

ACKNOWLEDGEMENTS

The authors would like to thank Dr B. DULMET for his help in performing FE analysis and P.BLIND from "Institut des Microtechniques de Franche-Comté" for the chemical etching of sensors.

REFERENCES

[1] **GERARDIN, RIXEN :** Théorie des vibrations. Ed. Masson, Paris 1993.

[2] **William C. ALBERT** : Force sensing using quartz crystal flexure resonators, 38th AFCS,p.233-239, Philadelphia 1984

[3] **P. TRUCHOT, R. BOURQUIN** : Torsional quartz resonators, 8th EFTF, p 862-871, Munich 1994.

9TH EUROPEAN FREQUENCY AND TIME FORUM - BESANCON - 8 - 9 - 10 MARCH 1994
Session 7 Number 3

ON MINIMISING QUARTZ CRYSTAL RESONATOR AGEING BEHAVIOUR AND COMPREHENSIVE AGEING STATISTICS FOR THIRD AND FIFTH OVERTONE QUARTZ CRYSTAL RESONATORS IN THE FREQUENCY RANGE 5 TO 16.384 MHz

Kanak K. Tuladhar and Gérald Jenni
OSCILLOQUARTZ SA, CH - 2002 NEUCHATEL

ABSTRACT

This paper is an extension to earlier paper (EFTF 1993) by the present authors. In this paper three new methods of minimising ageing behaviour will be outlined. Ageing could be negative or positive (occasionally ageing reversal effect has also been observed). It will be shown analytically that in general positive ageing resonators are superior to negative ageing resonators in terms of overall stability of the resonators/oscillators.

Comprehensive ageing data compiled over many years on some 2000 resonators/oscillators operating in the frequency range, 5 to 16.384MHz at third and fifth overtone modes will also be given and wherever possible statistically analysis will be carried out. Based on these wealth of production data, a number of interesting results and relationships between resonator designs, processing techniques, overtone modes, operating frequency range will be outlined. It will also be shown that conventional third and fifth overtone mode AT-cut-crystal resonators operating at 5MHz could have ageing better than 2E-10/day and 5E-11/day respectively by utilising optimum designs and processing techniques.

Finally difficulties involved in modelling ageing behaviour will be highlighted and wherever possible the relationship between daily, monthly, ageing will also be given.

1. INTRODUCTION

Oscillator ageing has been defined as the systematic change in frequency with time due to internal changes in the oscillator, when factors external to the oscillator are kept constant (refs. 1, 2). The ageing phenomena is still not well understood although it has been one of the most researched topics in the history of quartz crystal resonators and oscillators (refs. 3, 4).

In this paper three methods of minimising ageing will be given. Ideal, positive, negative ageing resonators / oscillators will be compared and their advantages and disadvantages will be highlighted.

Comprehensive ageing data compiled over many years on some 2000 resonators / oscillators operating in the frequency range 5 to 16.384MHz at third and fifth overtone modes will be given. Ageing data analysis techniques will be described and data analysis will be carried out. Relevant discussion will be given to highlight the relationship between the ageing behaviour and crystal resonator parameters, overtone mode & frequency.

Finally difficulties involved and associated problem in modelling ageing behaviour will be described.

2. AGEING MINIMISATION TECHNIQUES

The effect of C0/C1 (i.e the ratio between static and motional capacitances) on oscillator stability and ageing is well known and has been described elsewhere (refs. 3 and 4). However more general approach to minimise ageing behaviour is to maximise the ratio, [C0/C1] / F, where F is the frequency. This clearly indicates to minimise ageing one has to use largest possible electrode size and must operate at highest possible overtone mode. Furthermore quartz resonator blank must be contoured to highest possible dioptre value and must operate at lowest possible frequency.

Note that above parameters are not mutually exclusive and further constraints will be imposed by the available package size at a given operating frequency and therefore in general trade-off has to be made to achieve optimum practical resonator.

Note also that in general, the ratio, [C0/C1] / F for the third-overtone SC-Cut resonator is greater than the corresponding ratio for the third-overtone AT-Cut resonator, and lower with respect to fifth-overtone AT-Cut resonator. As such if optimum processing is used fifth-overtone AT-Cut Crystal could have lower ageing than third-overtone SC-Cut Crystal.

Recently a great deal of work has been carried out to minimise the acceleration sensitivity of the AT- and SC-Cut resonators. The work carried out by H Tiersten (ref. 5) and his associates has resulted in defining successfully a best possible resonator configuration and associated mounts, number of mounts, etc. But unfortunately to date no one has carried out similar sort of comprehensive work either analytically or experimentally with a view to minimise ageing. As such

currently, relative to quartz blank, shape, contour, size and frequency, optimum resonator design parameters, electrode shape, size, thickness, electrode tab width & positions, and number of mounts & their positions are not known.

Based on work carried out some twenty years ago (refs. 6 & 7) at least for the plano-plano resonator, by considering decay constants,α_x and α_z (see fig 1 of ref. 6) it can be shown that the optimum resonator blank shape is rectangular (square) for square (rectangular) electrode shape, for the optimum resonator Q (i.e to some extents for ageing). The ratio between the sides of rectangular blank or electrode will depend on the thickness of the electrode. An extension of this work to cover precision resonators will be a great help to minimise resonator ageing.

To date Crystal oscillators are designed using hierarchical approach, i.e to say oscillator designers will be asking improved resonators, and similarly resonator designers will be asking superior quartz material for minimising ageing. Although at each stage TQM, quality circle might be utilised, for the state-of-the-art oscillator design harmonious design technique, Fig. 1 is really preferred. This approach involves making each designer aware of the problems faced by other designers and material suppliers, as such wherever possible compromise could be made jointly to improve the overall oscillator performance. Finally user must be aware of the limitations and should not specify incompatible specifications, i.e low-ageing and large frequency pulling.

3. IDEAL, NEGATIVE AND POSITIVE AGEING RESONATORS / OSCILLATORS

Generally ageing could be negative or positive (occasionally ageing reversal effect has also been observed). An ideal resonator (i.e resonator, which does not age) has been proposed by J. Vig (ref. 8) at the 48th AFCS in Boston. However during discussion session, H. Tiersten (ref. 9) pointed out that in real life ideal resonator is not possible to realise because even at moderate drive level, crystal resonator will experience ageing. To expand the argument further, real resonator, Fig. 2 can be seen as combination of ideal resonator together with an ageing part of the resonator, say Cx. Ageing effects on Cx for the negative, ideal and positive ageing resonators are detailed in Table 1. This shows, firstly, for the ideal resonator the value of Cx must be infinity at all time, this in fact would not be possible. Secondly the value of Cx goes to negative (positive) for the negative (positive) ageing resonators. This indicates in practice positive ageing resonators are preferred, as for the negative ageing resonator, the ageing part of the resonator, Cx is in fact inductive.

Now consider two identical oscillators and suppose after initial ageing period, one is ageing linearly at + 5E-10 / day and other at - 5E-10 / day. Noise on the ageing curve due to varactor noise, instability within the mode selector, and instability within the load termination can be lumped into physical load capacitor instability. Note also the load capacitance will be increased (decreased) for the positive (negative) ageing oscillators during periodic alignment to compensate for ageing. Instabilities within the ageing curve due to load capacitor instability of +/- 5ppm have been computed for a typical third-overtone resonator equipped oscillator. The results are tabulated in Table 2 for the lapsed time of 3, 7 and 10 years. These results clearly indicate, although initially both positive and negative ageing oscillators have the same noise level, as the lapsed time increases noise level increased (decreased) for the negative (positive) ageing resonators. Therefore again positive ageing resonators/oscillators are preferred. Note also that the effect described above will be 5 to 10 times worse for the fundamental mode resonator based oscillator and similarly will be better by 5 to 10 times for the fifth-overtone-mode resonator based oscillator.

The loaded Q for the positive (negative) ageing oscillator will increase (decrease) as the lapsed time increases. This factor is well known and will not be elaborated.

4. DATA ANALYSIS TECHNIQUES

Data analysis techniques will follow the same pattern as one described earlier (ref. 4). Firstly time taken to meet a prescribed ageing specification will be analysed, then for the statistically significant batches detail ageing analysis will be carried out.

5. DISCUSSION ON MEASURED AGEING RESULTS

Out of many thousands of oscillators manufactured, ageing data has been compiled for some 2000 oscillators equipped with third and fifth overtone mode cold-weld encapsulated AT-Cut resonators operating in the frequency range, 5 to 16.384MHz. These include, (403, 424, 298) and (360, 412, 84) units of third and fifth overtone modes operating at (16.384MHz, 10 MHz, 5MHz) and (10MHz, 8.192MHz, 5MHz) respectively. And corresponding values of [C0 / C1], [C0 / C1] / F and prescribed ageing specifications after four weeks of pre-ageing are given in Table 3.

Table 3. Useful Design Information and Prescribed Ageing Specification for various AT-Cut Resonators.

Frequency [F]	[C0 / C1]	[C0 / C1] / F	Ageing Specification after 4 weeks
THIRD - OVERTONE - MODE			
16.384 MHz	5800	350 E - 06	5.0 E - 10 / day
10.000 MHz	3800	380 E - 06	5.0 E - 10 / day
5.000 MHz	6900	1380 E - 06	2.0 E - 10 / day
FIFTH - OVERTONE - MODE			
10.000 MHz	20,000	2000 E - 06	2.0 E - 10 / day
8.192 MHz	33,000	4000 E -06	2.0 E - 10 / day
5.000 MHz	22,000	4400 E - 06	1.5 E - 10 / day

5.1 Third-Overtone Mode Resonators / Oscillators

Detailed ageing analysis carried out for the third overtone-mode 16.384MHz resonators / oscillators has already been reported (see ref. 4).

Now consider the distribution, Fig. 3a, the time taken to meet the ageing of 5E-10/day for some 424, 10.000MHz resonators/oscillators. This clearly demonstrates that some 84% and 95% met the specified ageing within 22 and 29 days respectively. It can also be seen that around 44% and 29% met the ageing during the periods, 10 to 15 and 16 to 22 days respectively. Note also that some 10% met the prescribed ageing within 10 days and remaining 5% took between 31 to 45 days.

Distributions, Figs. 3b and 3c are the detailed ageing statistics for the two most significant batches meeting the prescribed ageing during the period 10 to 15 and 16 to 22 days respectively. These results, Figs. 3a, 3b, 3c and Figs. 2, 3a, 3b of ref. 4 clearly demonstrate that, although 10MHz and 16.384MHz have about the same [C0 / C1] / F values, 10MHz resonators / oscillators are better in terms of ageing behaviour.

Similarly consider the distribution, Fig. 4a, the time taken to meet the ageing of 2E-10 / day (a tough specification for the third-overtone AT-Cut resonator !) for some 298, 5.000 MHz resonators / oscillators. This demonstrate that some 77% met the specified tough specification within 30 days and remaining 23% took 31 to 55 days.

Distributions, Figs. 4b and 4c are the detailed ageing statistics for the two cumulative batches meeting the prescribed ageing during first 30 days and during the period 19 to 30 days respectively. Figs. 4b and 4c demonstrate that there is no significant differences in terms of ageing distribution for these two cumulative batches. Therefore the tough specification of 2E-10/day after 4 weeks of pre-ageing could be considered as the limit for the third-overtone AT-Cut resonators.

5.2 Fifth-Overtone Mode Resonators / Oscillators

Now consider the distribution, Fig. 5a, the time taken to meet the ageing of 2E-10/day for some 390, 10.000MHz resonators/oscillators. This clearly demonstrate that some 74% and 94% met the prescribed ageing within 19 days and 29 days respectively. Remaining 6% took between 30 and 60 days. Distribution, Fig. 5b is the detailed ageing statistics for the combined two most significant batches, meeting the ageing during the period 10 to 29 days.

Similarly consider the distribution, Fig. 6a, the time taken to meet the ageing of 2E-10/day for some 412, 8.192MHz resonators/oscillators. This demonstrate that some 76% and 87% met the specified ageing within 21 days and 29 days respectively. Some 14% met the ageing within first 11 days and remaining 13% took between 30 to 60 days. Distribution, Fig. 6b is the detailed ageing statistics for the combined two most significant batches meeting the ageing during the period, 12 to 21 days.

These results, Figs. 5b and 6b clearly demonstrate that 8.192MHz resonators / oscillators have about 10% lower three-sigma ageing value relative to 10MHz.

Finally consider the distribution, Fig. 7a, the time taken to meet the ageing of 1.5E-10/day for the limited 84 pieces of 5.000MHz resonators / oscillators. This clearly demonstrates that 62%, 92% and full 100% met the prescribed ageing within 20, 29 and 39 days respectively. Distribution, 7b is the detailed ageing statistics for the complete batch of 84 units. This shows some 66% have ageing less that 1E-10/day. Improved ageing less than 1E-10/day can be achieved by pre-ageing resonators / oscillators for an extended period and will be described elsewhere in this paper.

5.3 Ageing Models and Approximations

Extended ageing measurement carried out for two 16.384MHz crystal resonators/oscillators and one 5.000MHz fifth overtone crystal resonator/oscillator are given in Figs. 8a and 9a. Relevant discussion for 16.384MHz third-overtone resonator/oscillator has already been reported (ref. 4). However note that for the fifth-overtone mode 5.MHz typical ageing of 1.5 E-10/day has been achieved after some 16 days. Further improvements have been achieved by extending the pre-ageing period. Note that for the same resonator ageing of 1E-10/day and 5E-11/day have been achieved after some 25 and 75 days respectively.

The ageing curves given in Figs. 8a and 9a can be approximated by,

$$y = C1 \cdot \ln (C2 \cdot t + 1) + C3 \cdot t$$

where C1, C2, C3 are constants,
t is the lapsed time period
and y is the change in frequency from the nominal value.

These curves clearly demonstrate that initially, i.e. during first few days after oscillator switch-on, the ageing rate is quite high. This phenomenon has been demonstrated in Figs. 8b and 9b by computing the slope of the curves given in Figs. 8a and 9a for every successive 8 day - period. The ageing contributions from the logarithmic part and linear part are equal after about 50 days and 40 days respectively for the third-overtone 16.384MHz and fifth-overtone 5.000MHz. And similarly contribution from the logarithmic part is about ten-times lower after 550 and 400 days for 16.384MHz third-overtone and 5MHz fifth-overtone resonators /oscillators respectively. Computed and measured monthly ageing rates are given in table 4.

Table 4 : Monthly Ageing

Period	P3, 16.384MHz, AT		P5, 5.000MHz, AT	
	Computed	Measured	Computed	Measured
First	2.34E-08	2.35E-08	5.61E-09	5.67E-09
Second	1.30 E-08	1.30E-08	1.99E-09	2.01E-09
Third	1.05E-08	1.00E-08	1.60E-09	1.68E-09
Fourth	9.40 E-09	9.20E-09	1.45E-09	1.36E-09

It should be noted that ageing modelling cannot be generalised and must be localised to a specific manufacturer and their resonator design and processing techniques together with raw-material used at a given frequency. And validity must be verified by extended ageing measurement over medium/long-term period.

Although the final ageing after extended ageing period could be nearly the same for all resonators from a given batch (see C3 value for the two curves given in Fig. 8a), initial ageing during first few days, when the ageing contribution from the logarithmic part is predominant, could be different (see C2 value for the two curves given in Fig. 8a). Determining the value of C2 is the most difficult part in modelling ageing behaviour.

5.4 Further Discussions

Results presented in Figs. 3b, 3c, 4b, 4c, 5b, 6b, 7b and Figs. 3a and 3b of reference 4 clearly showed 5 to 10% of resonators/oscillators have negative ageing. This could have been caused due to a number of reasons, currently are under study and will be reported in due course.

6. CONCLUSIONS

It has been shown in this paper that to minimise ageing behaviour of the resonator, the term [CO/C1] / F has to be maximised. And conventional hierarchical oscillator / resonator design techniques with severe contraint on quartz material have been extended to a new approach called harmonious design technique.

It has been pointed out that for the conventional resonator, circular quartz blank having circular electrode might not be an optimum for the minimum ageing. Based on work carried out some twenty years ago for the plano-plano resonator, an extension of this work has been proposed to derive optimum electrode shape, size and thickness together with appropriate electrode tab width, number of mounts and their positions, for a given quartz blank shape, contour and size with a view to minimise ageing.

In real life ideal resonator (i.e resonator, which does not age) is not possible, however ageing can be minimised for any given resonator / oscillator. Positive and negative ageing resonators have been compared and demonstrated that oscillators having positive ageing are superior to negative ageing ones.

Based on some 2000 resonators / oscillators operating in the frequency range 5 to 16.384MHz at third and fifth overtone modes, ageing of 5E-10 / day and 2E-10 / day have been routinely achieved at 16.384MHz and 5MHz respectively for the third-overtone mode resonators. And similarly for the fifth-overtone mode resonators ageing of 2E-10 / day and 0.5 to 1.5 E-10 / day at 10MHz and 5MHz respectively have been routinely achieved in an industrial production environment.

Finally some preliminary ageing modelling carried out suggested that ageing model can not be generalised, it must be localised to a specific manufacture and their resonator design, processing and packaging techniques together with quartz material used.

7. ACKNOWLEDGEMENTS

The authors wish to thank Dr. E.P. Graf for his helpful comments and suggestions. Thanks also to Mr. Y Schwab of Oscilloquartz for carrying out most of the ageing measurements.

8. REFERENCES

1. CCIR Recommendation No. 686, "Glossary", CCIR 17th Plenary Assembly, Vol VII, "Standard Frequency and Time Signals (Study Group)", 1990.
2. "IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology", IEEE, Std. 1139-1988, IEEE NY, 10017.
3. Vig, J.R, and Meeker, T.R, "The Ageing of Bulk Acoustic Wave Resonators, Filters and Oscillators", Proc. 45th AFCS, May 1991, pp. 77-101.
4. Tuladhar, K. K, and Jenni, G, "Low Ageing HC-36/U Encapsulated Third-overtone AT-Cut 16.384MHz Crystal Resonators and Oscillators, Proc. 7th EFTF, March 1993, pp. 297-304
5. Lwo, B.J, and Tiersten, H.F, "Calculation of the Optimal Clip Dimensioning to Minimise the Influence of Fabrication Imperfection on the Acceleration Sensitivity of SC-Cut Quartz Resonators with Stiffened Rectangular Support System", Proc. 48th AFCS, June 1994, pp. 165-171.
6. Werner, J.F, and Dyer, A.J, "The Relationship between Plateback, Mass Loading and Electrode Dimensions for AT-Cut Quartz Crystals having Rectangular Resonators Operating at Fundamental and Overtone Modes", Proc. 30th AFCS, June 1976, pp. 42-53.
7. N.H.C. Reilly and M. Redwood, "Wave Propagation Analysis of the Monolithic Crystal Filter", Proc. IEE, Vol. 116, No. 5, May 1969, pp. 653-660.
8. Vig, J.R, and Walls, F.L, "Fundamental Limits on the Frequency Instability of Quartz Crystal Oscillators", Proc. 48th AFCS, June 1994, pp. 506-523.
9. Tiersten, H.F, Comments to above ref. No. 8 during discussion, at 48th AFCS, Boston, June 1994.

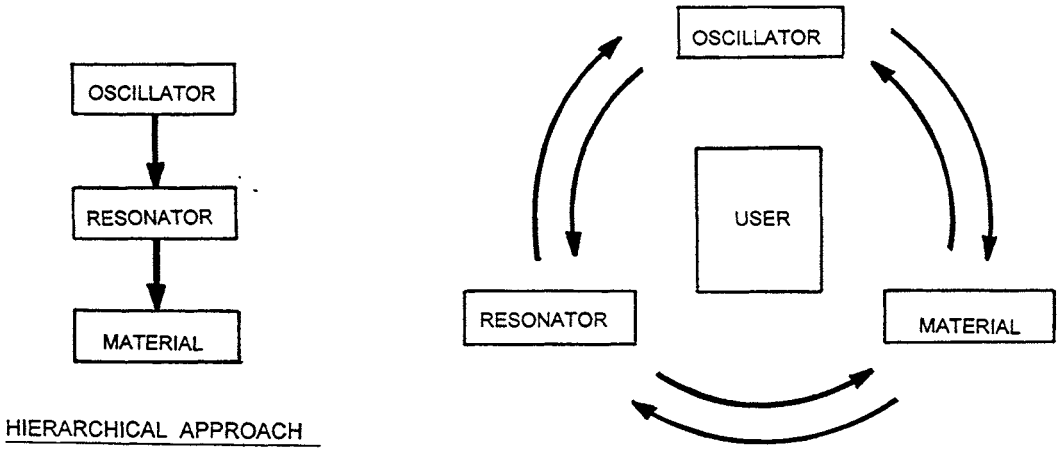


Fig. 1. HIERARCHICAL AND HARMONIOUS DESIGN TECHNIQUES

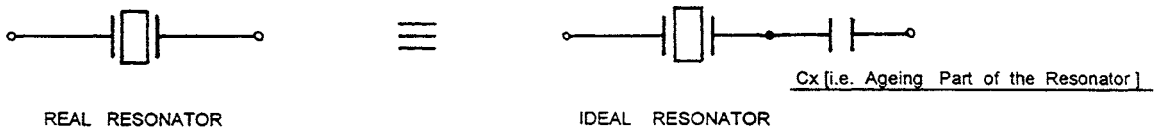


Fig. 2. REAL and IDEAL RESONATORS

TABLE 1 EFFECT ON Cx due to RESONATOR AGEING BEHAVIOUR

	at t = 0	t →	t1 to t2
NEGATIVE AGEING	Cx = ∞	Cx := Large and Negative value	Cx(t2) > Cx(t1)
IDEAL [i.e. No Ageing]	Cx = ∞	Cx = ∞ for all t	Cx(t2) = Cx(t1) = ∞
POSITIVE AGEING	Cx = ∞	Cx := Large and Positive Value	Cx(t2) < Cx(t1)

TABLE 2 NOISE DUE TO LOAD CAPACITOR INSTABILITY(+/-5ppm) FOR THE RESONATOR HAVING C1=0.8fF,Co=4.0pF and CL=15pF

TIME[YEARS]	AGEING RATE	OSCILLATORS HAVING		TOTAL AGEING
		POSITIVE AGEING	NEGATIVE AGEING	
0	5.0E-10 /DAY	+/- 8.31E-11	+/- 8.31E-11	0
3	5.0E-10 /DAY	+/- 8.15E-11	+/- 8.47E-11	5.40E-07
7	5.0E-10 /DAY	+/- 7.94E-11	+/- 8.67E-11	1.26E-06
10	5.0E-10 /DAY	+/- 7.77E-11	+/- 8.82E-11	1.80E-06

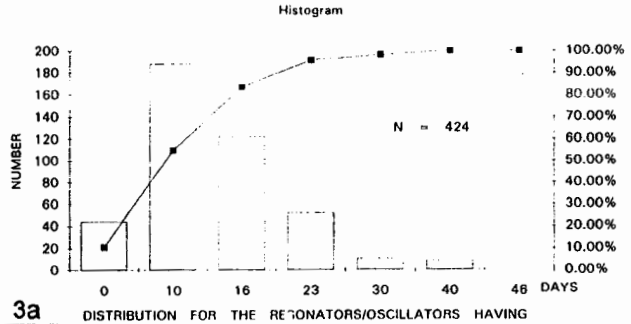


Fig. 3a DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 5E-10/DAYS

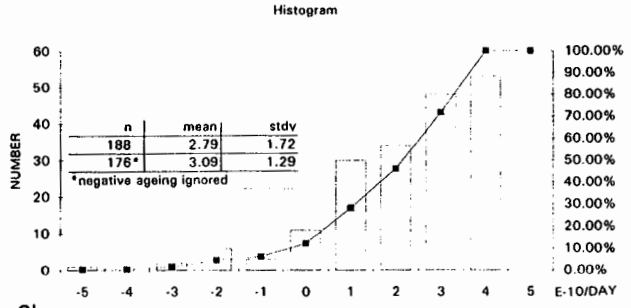


Fig. 3b DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 5E-10/DAY DURING 10 TO 16 DAYS

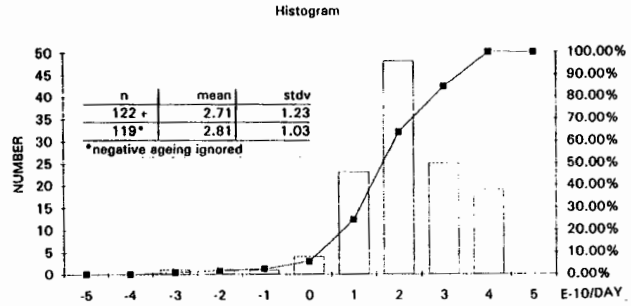


Fig. 3c DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 5E-10/DAY DURING 17 TO 23 DAYS

Figs. 3a to 3c THIRD OVERTONE 10.000MHz AT

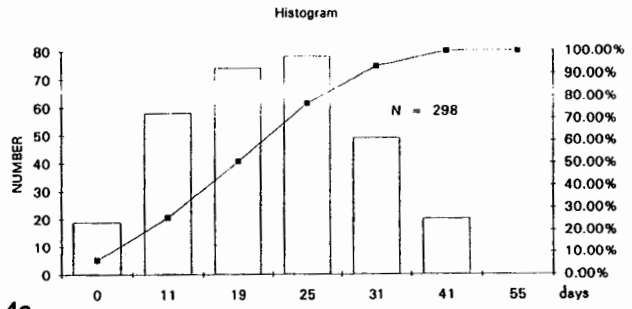


Fig. 4a DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 2E-10/DAY

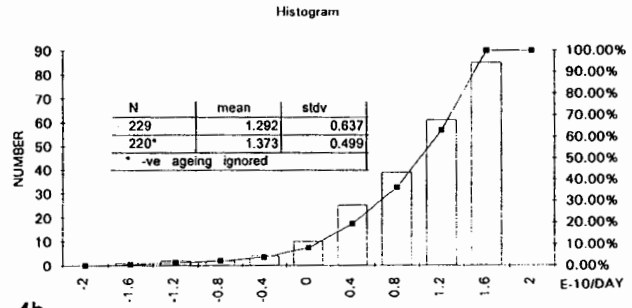


Fig. 4b DISTRIBUTION FOR THE RESONATEORS/OSCILLATORS HAVING AGEING LESS THAN 2E-10/DAY DURING FIRST 30 DAYS

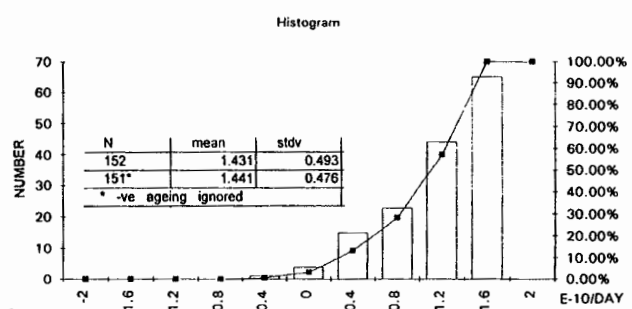


Fig. 4c DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 2E-10/DAY DURING 19 to 30 DAYS

Figs . 4a to 4c THIRD OVERTONE 5.000MHz AT

AT P5 10 MHz

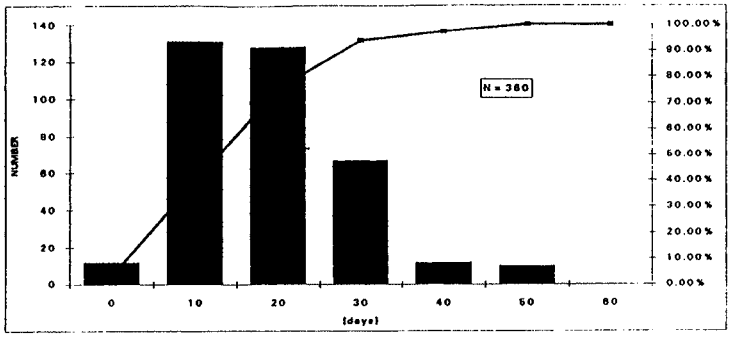


Fig. 5a : DISTRIBUTION FOR RESONATORS / OSCILLATORS HAVING AGEING LESS THAN 2 E-10/day

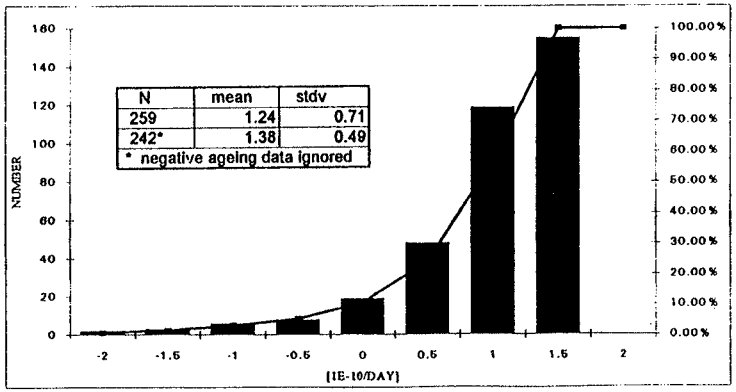


Fig. 5b : DISTRIBUTION FOR RESONATORS / OSCILLATORS HAVING AGEING LESS THAN 2E-10 DURING 10 to 29 DAYS

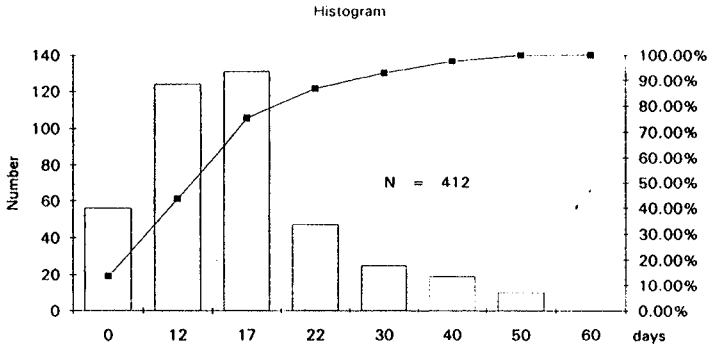


Fig. 6a DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 2E-10/DAY

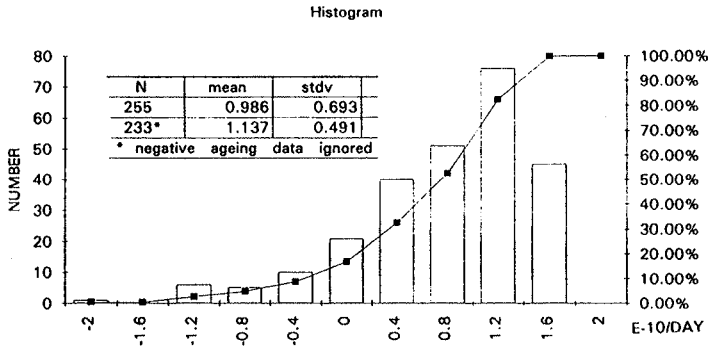


Fig. 6b DISTRIBUTION FOR THE RESONATORS/OSCILLARORS HAVING AGEING LESS THAN 2E-10/DAY DURING 12 to 22 DAYS

Figs . 6a to 6b FIFTH OVERTONE 8.192MHz AT

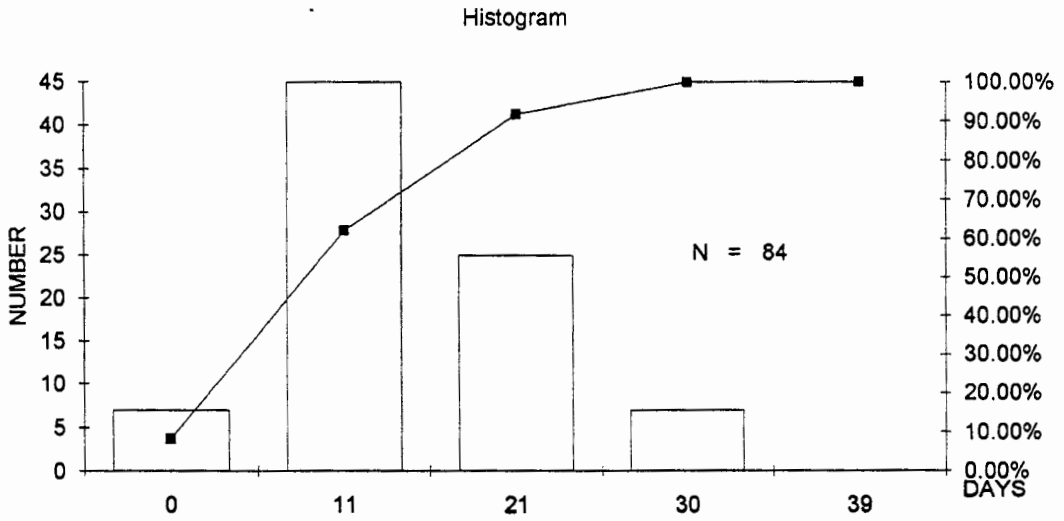


Fig. 7a DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 1.5E-10/DAY

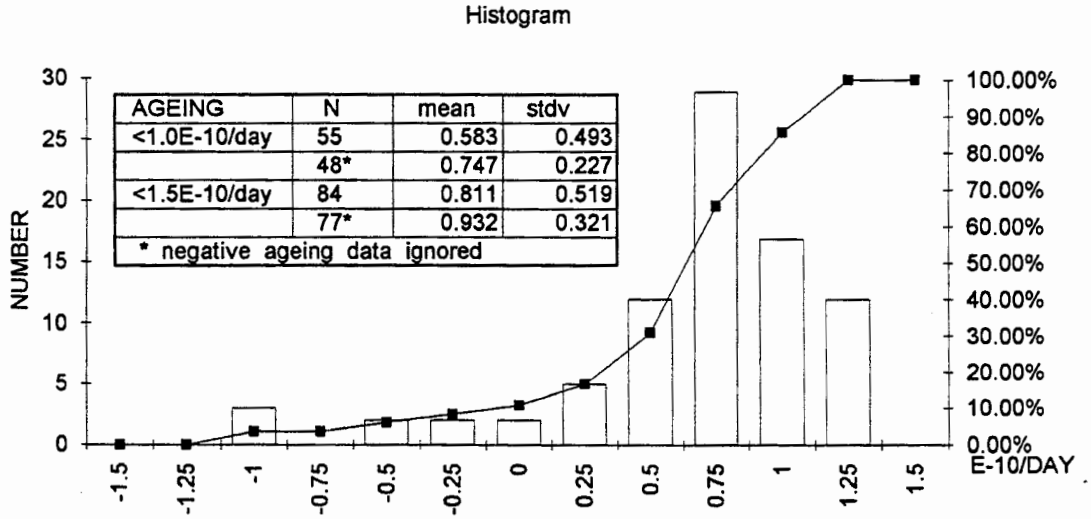


Fig. 7b DISTRIBUTION FOR THE RESONATORS/OSCILLATORS HAVING AGEING LESS THAN 1.5E-10/DAY DURING FIRST 39 DAYS

Figs. 7a to 7b FIFTH OVERTONE 5.000MHz AT

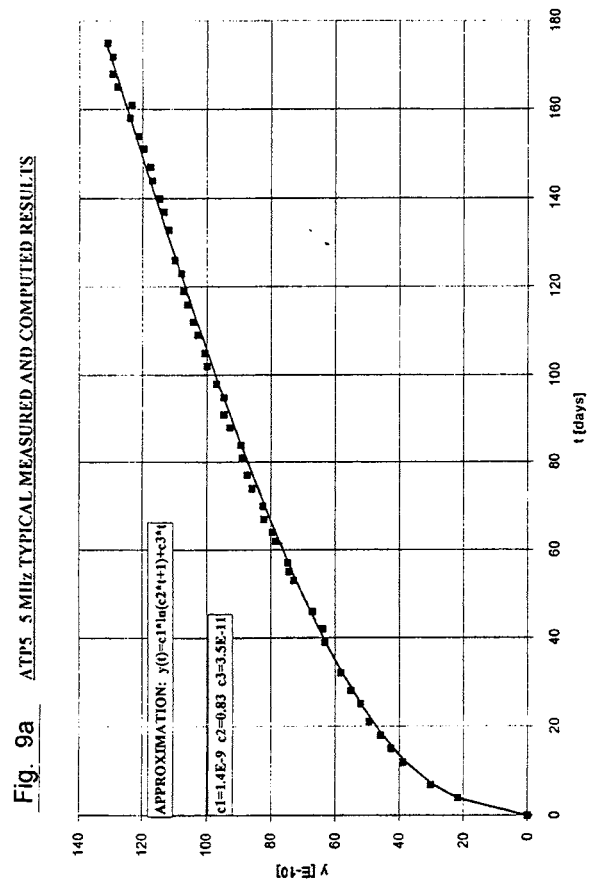


Fig. 9a ATP5 5 MHz TYPICAL MEASURED AND COMPUTED RESULTS
APPROXIMATION: y(t)=c1*ln(c2*t+1)+c3*t
c1=1.4E-9 c2=0.83 c3=3.5E-11
y [E-10]
t [days]

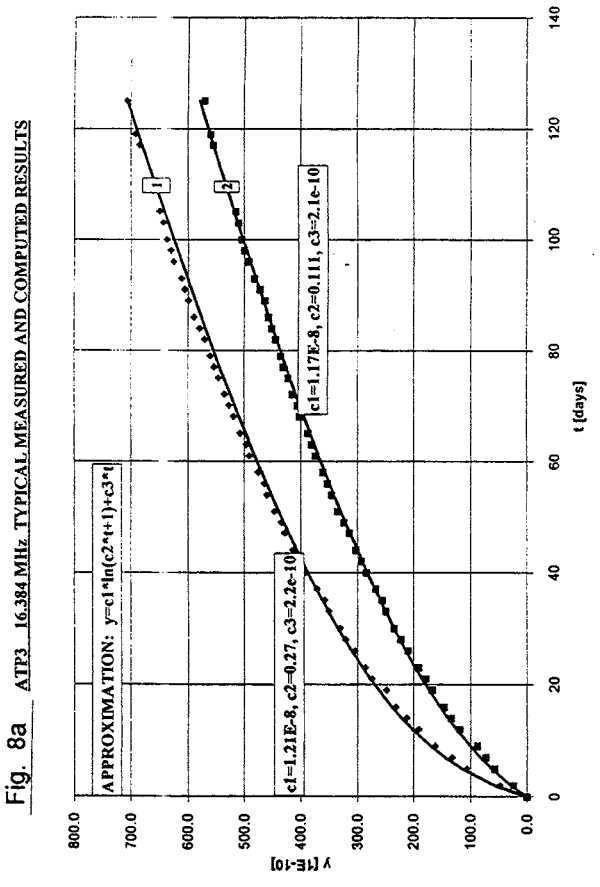


Fig. 8a ATP3 16.384 MHz TYPICAL MEASURED AND COMPUTED RESULTS
APPROXIMATION: y=c1*ln(c2*t+1)+c3*t
c1=1.21E-8, c2=0.27, c3=2.2e-10
c1=1.17E-8, c2=0.111, c3=2.1e-10
y [1E-10]
t [days]

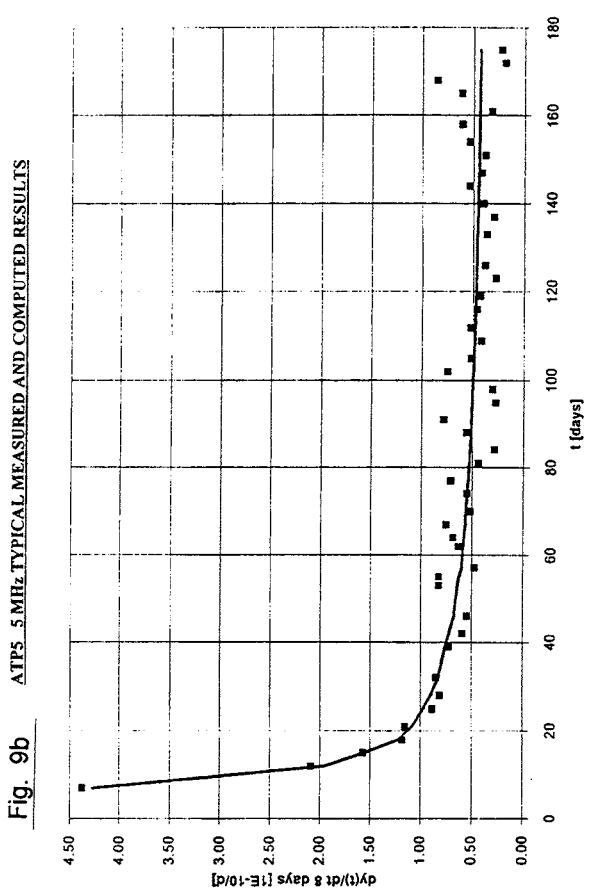


Fig. 9b ATP5 5 MHz TYPICAL MEASURED AND COMPUTED RESULTS
dy(t)/dt 8 days [1E-10/d]
t [days]

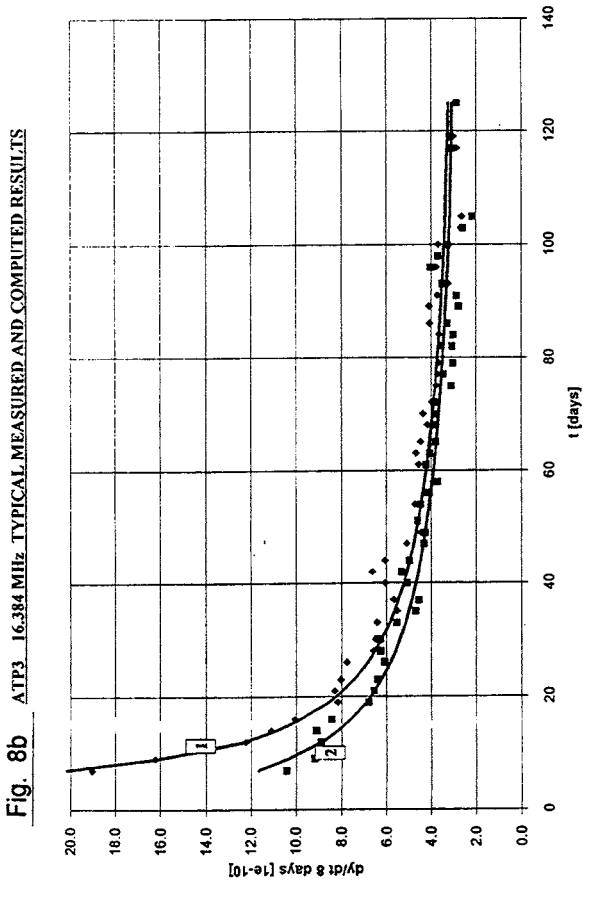


Fig. 8b ATP3 16.384 MHz TYPICAL MEASURED AND COMPUTED RESULTS
dy/dt 8 days [1e-10]
t [days]

AN ELECTRODELESS BT CUT, 5th OVERTONE QUARTZ CRYSTAL RESONATOR AT 50 MHz

Ch. Longet, G. Robichon

bva INDUSTRIE SA
1 boulevard Fleming - BP 1985 - 25020 BESANÇON CEDEX - FRANCE

Abstract

A 50 MHz, BT cut, 5th Overtone Quartz Crystal has been developed by bva INDUSTRIE SA using the electrodeless technology approach.
The target was : Maximum Q value, keeping Q x frequency F = 30 E12 which is classical for the BT cut.
Experimental results are as good as expected :

Q	=	600 000
R_1	=	100 Ohms
TOP	=	70-75° C (temperature turnover point)

An ultra low noise oscillator has then been designed for an application with ensuing multiplication to almost 10 GHz. The minimum of $S_y(f)$ had to be at about 150-200 Hz from the carrier frequency. Two iterations were necessary for improved phase noise results. Two output frequencies were requested :

50 MHz	N > 10 dBm
100 MHz	N > 10 dBm

Two oscillators were made with a particular requirement concerning the output level between the two units :

N1 - N2 < 0.3 dBm

In order to keep the minimum position for $S_y(f)$ during all the life of the product which is phase locked to a frequency reference, aging has to be controlled.

The target values were :

ΔF/F	< 5 E-10/day
ΔF/F	< 1.5 E-8/month
ΔF/F	< 1.5 E-7/year
ΔF/F	< 4 E-7/3 years

I - THE RESONATOR

1.1 - Introduction

After several years of development and experiences principally with AT and SC cuts and at frequencies of 5 and 10 MHz, the electrodeless BVA technology resonators have proved their excellent capabilities [1, 2].

The LHA's atomic clock application at the origin of this paper needs a low phase noise oscillator with a main frequency output of 50 MHz. One way to get there is to take a well-known 10 MHz resonator and multiply its frequency by 5. In this case, we keep the advantages of the known 10 MHz quartz crystal, but we also lose 14 dB on the phase noise level. Therefore, we decided to make a 50 MHz resonator in BT cut which exhibits a better Q factor, and a plate thickness easier to manufacture in a classical way. The main characteristics of the electrodeless technology is that the electrodes are not deposited directly on the active part, i.e, the vibrator, but on two pieces, like "condensers", placed on each side with a gap of few microns [3]. Moreover, the active part is decoupled from the supporting part by oriented quartz bridges.

These three parts are placed in a special suspension system [4] and encapsulated in a standard HC-40/U cold-weld can (Fig. 1).

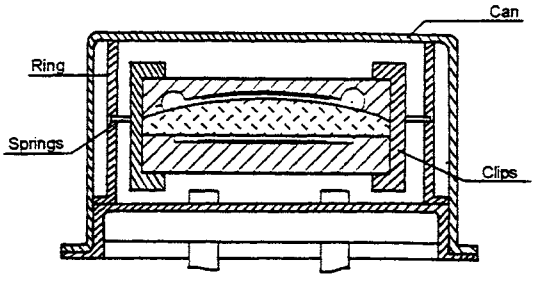


Figure 1

1.2 - Theoretical approach of the resonator design

Because this type of product has never been made before, we had to perform some calculations to determine the most appropriate design.
We then decided to manufacture units with the following specifications :

Frequency : 50 MHz ± 10^{-6}
Overtone : 5t h
Turn over point : 75° C ± 5° C
Quality factor : $> 600.10^3$
Serial resistance : $< 100\ \Omega$
Aging : $< 1.5\ 10^{-8}$ /month
First spurious mode : $f_0 + 60\ kHz$
Encapsulated in HC-40 can

For these specifications, the following plano-convex crystal design parameters are adequate :

$\varphi = 0$ $\quad R_c = 700\ mm$
$\theta = 46°\ 00'\ 45"$ $\quad \varnothing_{el} = 3.5\ mm$

where R_c is the radius of curvature.
The frequency temperature characteristic considered in the form

$$\frac{\Delta f}{f} = a(T - T_0) + b(T - T_0)^2 + c(T - T_0)^3$$

is represented in Fig. 2.

1.3 - Results

Two batches were manufactured. The first to confirm the simulation results and the second to make the resonators to be used in the oscillators.
The first tests performed were the frequency-temperature behaviour and the motional parameters measurement, thanks to a specific software using the CEI444 method. The results are presented in table 1 and an example of frequency-temperature curve in Fig. 3. The turn over point temperature is calculated by a 3rd order fitting of the measured data. The cubic coefficients

$a = 3.8\ 10^{-6}$
$b = 2.36\ 10^{-8}$
$c = -1.9\ 10^{-10}$

give a frequency difference at T_{top} between the real points and the fitted curve $< 10^{-8}$.

N°	f_{req} (Hz)	T_{top} (°C)	Q (10^6)	R_r (Ohm)	L_1 (H)
1	50 046 031	69	0.62	101	0.2
2	50 058 396	71.4	0.552	105	0.185
3	50 063 910	73.2	0.651	100	0.207
4	50 064 914	74.2	0.635	98	0.198
7	49 999 946	76.1	0.34	145	0.184
8	49 999 995	73.4	0.705	91	0.203
9	50 000 019	73.9	0.67	95.7	0.204
10	50 000 007	74.8	0.633	99.5	0.2
11	50 000 031	75.9	0.655	95	0.2

Table 1

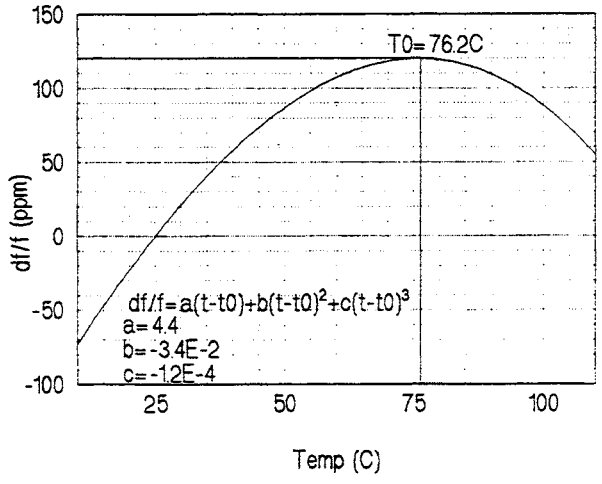


Figure 2

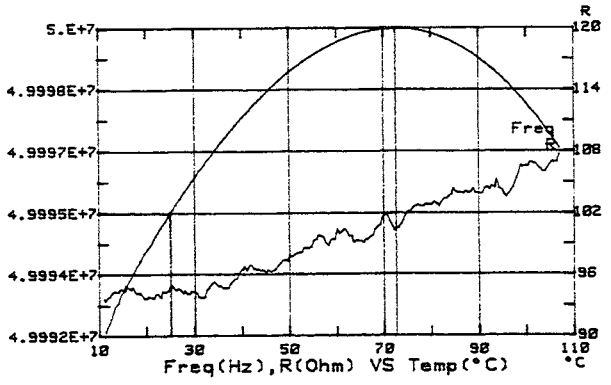


Figure 3

The pieces N° 1, 2, 3 and 4 made from the first batch, were adjusted by taking the fundamental mode as reference, because we didn't have the equipment to measure frequency over 20 MHz. This explains the frequency shift.
For the second batch, an appropriate equipment was available and this permitted to reach a $-10^{-6}/+6.10^{-7}$ frequency adjustment.

The most important result is the Q factor. Except N° 7 (bad etching process), the average value is about $6.4\,10^5$ (7.10^5 for one piece), this is better than we expected.

The dispersion of the T_{top} values is explained by the small thickness of the plates which makes the angular adjustment more difficult, but the values stay in a satisfactory range.

Fig. 4 shows the location of unharmonic modes on one sample.

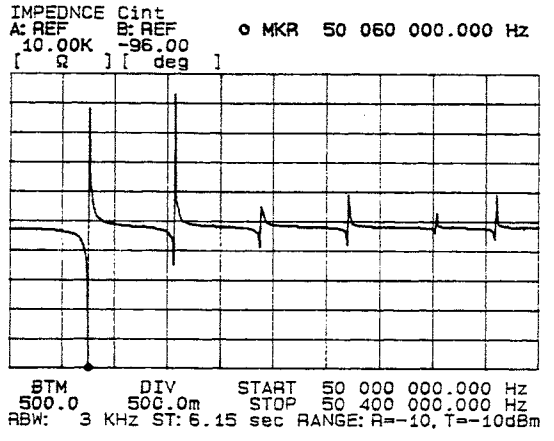


Figure 4

II - THE OSCILLATOR

2.1 - Context

The oscillator was ordered with an ambitious phase noise specification at 50 MHz and 100 MHz. The hardest target was at 100 Hz from the carrier.

For 50 MHz L(100 Hz) ≤ -140 dBc/Hz
For 100 MHz L(100 Hz) ≤ -134 dBc/Hz

The typical value at 10 MHz is L(100Hz) ≈ -145 dBc/Hz

So, after multiplication of a classical 10 MHz, we get :

At 50 MHz L(100 Hz) ≈ -131 dBc/Hz
At 100 MHz L(100 Hz) ≈ -125 dBc/Hz

In order to meet the phase noise specification and to keep these performance for more than ten years, it was decided to use a 50 MHz BT cut electrodeless type with a maximum of Q value and a minimum of 1/f noise for the first transistor.

2.2 - Development

The loaded Q value for an oscillator with high Q is obtained with the cross point position between f^{-3} and f^{-1} shape of phase noise which gives f_1 with :

$$f_1 = \frac{F_0}{2\,Q_L}$$

where Q_L is the loaded Q and F_0 the carrier (F_0= 50 MHz).

• It was reasonable to assume that Q_L should be about 300 000 if Q was of about 600 000.

• The motional resistance of about 100 Ohms allowed to design an electronics with only one transistor in the loop.

The options above can be considered the best dispositions for the phase noise performance of the oscillator at 50 MHz.

Remember that :

Q.F ≈ 30.10^{12} for BT cut
Q.F. ≈ 16.10^{12} for SC cut

It also supposes that thermal effects due to the more important slope of BT cuts compared to SC cuts can be rejected.

2.3 - Problems

It appears that the fundamental (R = 12 Ohms) and 3rd overtone modes were present in the output signal whereas they should have been rejected.

The modulation of the 5th overtone mode (R = 100 Ohms) by the fundamental mode limits the power put on the crystal so that the phase noise floor is restricted to -158 dBc/Hz at 50 MHz, and -152 dBc/Hz at 100 Hz obtained by multiplication.

However, the results close to the carrier were better than expected (see Fig. 5).

L(1 Hz) = -80 dBc/Hz at 50 MHz
L(1 Hz) = -74 dBc/Hz at 100 MHz

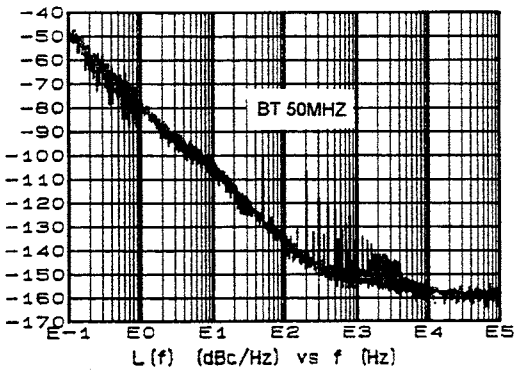


Figure 5

Short term stability was (see fig. 6)

σ(1 sec) = 3.5 E-12
σ(0.1 sec) = 4.8 E-12

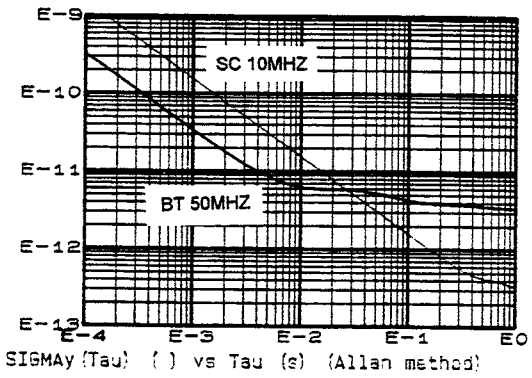


Figure 6

The specification for the crystal shall be for future series :

BT cut 5th overtone electrodeless
Q maximum
R(5th) < 100 Ohms
R(5th)/R (fundamental) < 4

This new spec. for the resonator will allow to make an oscillator with a better phase noise floor.

2.4 - Results and comparison with 10 MHz

We see Fig. 7 Sy(f) obtained with :

10 MHz SC cut 3rd overtone and
50 MHz BT cut 5th overtone

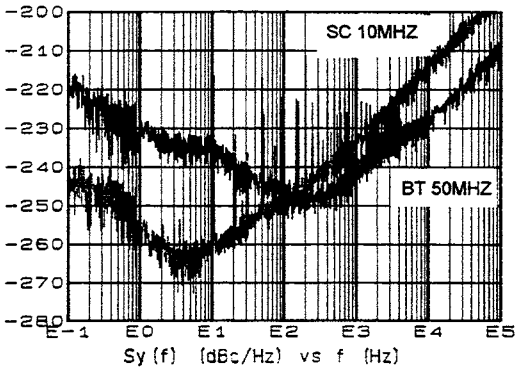


Figure 7

Sy(f) is the relative frequency fluctuations spectrum sometimes noticed, $S\left[\frac{\Delta\gamma}{\gamma_0}\right](f)$, used for comparing oscillators at different frequencies.

$$S_y(f) = \frac{f^2}{2\,\gamma_0^2} \quad L(f)$$

where

γ_0 = carrier frequency (Hz),
f = offset from the carrier (Hz).

III - CONCLUSIONS

Most objectives in the framework of our study have been reached. The product is available and bva INDUSTRIE is willing to supply the

resonator alone or the complete oscillator.

Future developments are planned with the objectives of direct resonator frequency at 100 MHz and of oscillator size reduction.

IV - ACKNOWLEDGMENTS

The authors wish to thank Prof. R.J. BESSON and Mr B. DULMET of the ENSMM-LCEP, Besançon, for their help in the calculation and theoretical phase of this project.

We also wish to thank Mr R. BARILLET of LHA who initiated this study.

BIBLIOGRAPHY

[1] J.NORTON, R.J. BESSON
"Tactical BVA quartz resonator performance"
Proc. 1993, IEEE Frequency Control Symp. pp 609-613.

[2] J. NORTON
"BVA type oscillators for spacecraft"
Proc. of the 45th Annual Symposium on Frequency Control, May 1991, pp 426-430.

[3] R.J. BESSON
"A new electrodeless resonator design"
Proc. of the 31st Annual Symposium on Frequency Control, June 1977, pp 147-152.

[4] R. DELAITE
"Special miniaturized SC-cut quartz resonator for severe environment applications"
Proc. of the 3rd EFTF, March 1989, pp 296-300.

SESSION 8

ATOMIC FREQUENCY STANDARDS : SIGNAL GENERATION AND ANALYSIS

Chairman : C. AUDOIN

Reducing the Effect of Interrogation Phase Noise on the Frequency Stability in Passive Frequency Standards : Experimental Results

R. Barillet , D. Venot and C. Audoin

LABORATOIRE DE L'HORLOGE ATOMIQUE

Unité Propre de Recherche du CNRS associée à l'Université Paris-Sud - Bât. 221 - UPS - 91405 ORSAY - Cedex - France

Abstract

The local oscillator phase noise limits the frequency stability in passive frequency standards [1] [2]. When the probing signal is permanently applied to the atomic resonator, this limitation is due to an intermodulation effect which takes place in the modulation-demodulation process. We have theoretically established that, with a square wave frequency modulation, a proper signal processing using a notch filter in the frequency synthesis [3] and a proper demodulation waveform [4] [5] can significantly reduce the spurious effect. Experimental test confirms the optimisation of the performance by the theoretical demodulation waveform and the previous State of the Art is improved by a factor of 4.5.

I. Introduction

Optical pumping and laser cooling make it possible to significantly improve the ultimate performance of passive frequency standards. Nevertheless a limitation related to the phase/frequency noise of the interrogation signal had been foreseen [1], and the expected improvements are presently limited by this spurious limitation in several high performance experiments [3][6]. For passive frequency standards where the probing signal is permanently applied to the atomic resonator, this effect is due to the modulation-demodulation process [2] which is necessary to precisely determine the top of the atomic resonance : when the frequency of the microwave interrogation signal is modulated at frequency f_M, an intermodulation effect translates fluctuations from all the even harmonics of f_M (in the Fourier domain) into the low frequency component in the error signal of the VCXO servo.

The main spurious contribution is related to the frequency noise existing near the second harmonics of the modulation frequency, in the Fourier domain. From this remark, a group at NIST realized the first reduction of the limiting effect : it consists in rejecting the phase noise at $\pm 2f_M$ from the carrier by means of a proper dual notch filter [3]. For sine modulation and demodulation waveforms, the possible improvement depends on the depth of the frequency modulation around the carrier : in the experimental conditions used by Szekely et al [3], the dual notch filter made it possible to reduce the spurious effect by one order of magnitude.

This first approach is interesting because it improves the ultimate frequency stability that may be obtained, but in fact it decreases the main noise contribution in the probe signal itself and does not reduce the intermodulation effect described in [2]. A remark by Kramer [1] forecasts that, with square modulation and demodulation waveforms, the limiting effect disappears if the output signal of the resonator is not filtered. This remark led us to investigate the field of Signal Processing - in low frequency - in order to reduce the intermodulation effect itself.

II. Theoretical results

In the case where the modulation frequency is small compared to the linewidth of the atomic resonator, the quasi-static assumption makes it possible to analytically express the frequency stability limit in the standard operating conditions [2], which correspond to a selective filtering of the resonator response. With the quasi-static assumption, a matrix calculation shows that in the general case, there is no possibility to fully suppress the spurious limiting effect when there is any truncation of the spectrum of the resonator signal : the only possible improvement consists in optimizing the behavior, probably by a proper signal processing.

In the general case where the resonator response is not selectively filtered, without any assumption concerning the phase noise in the probe signal, one obtains expressions based on infinite series which describe how the frequency stability is limited, depending on modulation and demodulation waveforms [7]. For example, when the frequency of the probe signal is square wave modulated, the frequency stability limit is proportional to the square root of the coefficient ξ^c given by

$$\xi^c = \frac{\displaystyle\sum_{n=1}^{+\infty} \lambda_{2n} \cdot \left(\sum_{k=0}^{+\infty} C_{2k+1} \cdot \frac{(4k+2)\,n}{(2k+1)^2 - 4n^2} \right)^2}{4 \cdot \left(\displaystyle\sum_{k=0}^{+\infty} \frac{C_{2k+1}}{2k+1} \right)^2} \qquad (1)$$

where λ_{2n} is a coefficient describing the shape of the phase noise spectrum, or a power gain applied to the Fourier component $2nf_M$ in the probe signal - for example an attenuation by a proper notch- and C_{2k+1} is the rate of the $(2k+1)f_M$ component in the demodulation signal. From relation (1), we

then determine how the spurious limiting effect can be reduced by the dual choice of a demodulation waveform and a convenient rejection.

The first case we have considered is a local oscillator showing a white phase noise in the (Fourier) frequency domain of interest, without any rejection in the probe signal : there is an optimum demodulation function -given by C_1=1, C_3=1/27, C_5=1/125...- but the improvement is only 1to 2%, which is not interesting in practice.

After that, we have considered a local oscillator where the phase noise component at $2f_M$ is more or less rejected in the Fourier domain : one finds interesting optimized waveforms described by their various spectral components. The following table shows the improvement ζ that can be expected for the frequency stability limit expressed in Allan standard deviation. Value 1 (no optim., no rejection in the table) corresponds to the usual case where the output of the resonator is selectively filtered and the $2f_M$ component is not rejected. We have observed that the optimum values do not change significantly in the various optimized configurations : for example C_3 remains in the vicinity of -0.2.

We notice that with a square wave frequency modulation, the expected improvement related to the only rejection of the $2f_M$ component is less favorable than in the sine wave frequency modulation mentioned in [3]. At last the expected improvement is obtained with harmonics coefficients quite different from those of a square wave (C_3=1/3, C_5=1/5...)

(Square wave freq. modulation) (C_1 = 1 in any case) SYSTEM ↓	Improvement factor ζ		
	No rejection	Full rejection	Rejection 30 dB
No optim.	(A) 1	(B) 1.89	1.89
Optim. with C_3 (≈-0.193)	1.00615	5.67	5.54
Optim. with C_3, C_5	1.00695	(C) 10.75	9.89
Optim. with C_3, C_5, C_7	1.00994	17.21	14.17
Optim. with C_3,C_5,... C_{11}	1.01018	32.3599	19.7307
Optim. with C_3, C_5... C_{21}	1.01037	83.653	23.8181

Table I - Theoretical results

Finally, we have simulated the behavior of the frequency standard [5][7] in three different configurations mentioned in the previous table : (A) without a rejection, (B) with a 30dB rejection and no otimization, (C) with a 30dB rejection and (C_3,C_5) optimization. The results of the simulation were similar to that of the previous table, but always a little lower : the obtained improvement is 1.7 (instead of 1.89) with a full rejection at $2f_M$ and approximately 6 instead of 10 with the (C_3,C_5) optimization.

III. Experimental apparatus

The investigation of new signal processing possibilities requires the synthesis of various waveforms that will be used to demodulate the resonator response. The required waveforms will be generated by a simple digital apparatus followed by a DAC. For the demodulation, commercial synchronous detections realize a square wave demodulation, whilst we need any demodulation waveform : for that reason, the demodulator will be an analog multiplier operated in satisfactory conditions of "linearity" -that is properly realize the product of the two input signals.

III.1. Digital synthesis of the demodulation waveform

III.1.a. Operating principle

The waveform generator must be able to

- synthesize any waveform in low frequency up to 100Hz, with 21 harmonics showing adjustable rates ; the fundamental frequency is considered as the first harmonics
- adjust the phase within 1° inside the period

A series of numbers describing the desired waveform is loaded into a memory ; the analog demodulation signal is obtained by permanently reading the memory and converting the digital data.

The corresponding functional scheme is presented in figure 1 : it looks like that of a DDS, but does not show so many possibilities. A ramp generator gives the initial address, coding wheels make it possible to adjust the phase offset - in fact the address offset - and an adder gives the final address in the memory. The digital signal delivered by the memory is converted into an analog signal by the DAC, then filtered. The sampling frequency F_e is generated from the clock frequency F_k which has the same time base as the modulation frequency F_M, so the whole apparatus is synchronous.

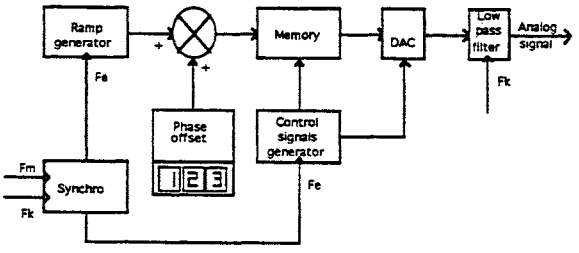


Fig.1 : Demodulation waveform synthesis

III.1.b. Realization of the apparatus

In order to adjust the phase within 1° anywhere inside the period, we need more than 360 points, that is 512 points - 9bits of address- and a sampling frequency $F_e = 512F_M$. Taking into account the number of synthesized harmonics (21), we see that Shannon condition is quite fulfilled. For the amplitude of the signal, quantization noise led us to choose 12bits for data.

An EPROM memory is used to enable us to easily modify the demodulation waveform. In fact this memory is realized by two EPROMs (NMC27C16, 8bits each). From the coefficients of the various harmonics to be synthesized, a program calculates the value of the function for the 512 required points ; after levelling the function so that its maximum is $FFF and its minimum $000 -respectively +5V and 0V at the output of the DAC- the program loads the EPROM memories.

The upper modulation frequency is related to the conversion time in the DAC, which theoretically gives a maximum sampling frequency F_e = 125kHz. Practically the board shows a satisfactory operation up to F_e = 107kHz ; the corresponding modulation frequency is F_M = 209Hz, larger than our practical requirements.

The operation of this apparatus has been carefully checked. It has been specially verified that the obtained waveform and the corresponding spectrum where very precisely identical to the expected wave and harmonics coefficients.

III.2. Analog demodulator

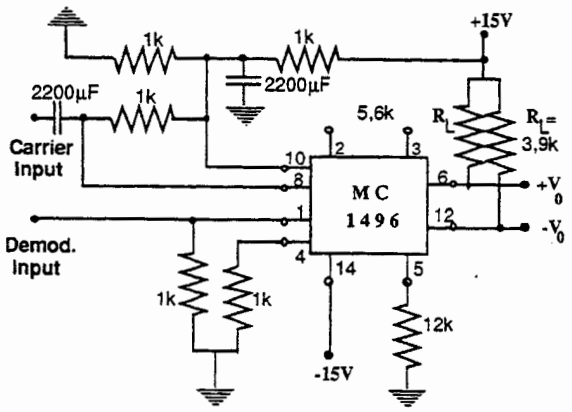


Fig.2 : Analog demodulator

The MC1496 circuit is a classical modulator-demodulator usually operated in the radio frequency domain. We have modified the values of the input components according to the scheme shown in figure 2, in order to use this circuit in very low frequency (35Hz and harmonic frequencies). The "Carrier Input" receives the signal delivered by the atomic resonator, while the "Demod Input" receives the demodulation waveform. To be operated in "linear" conditions, this circuit must fulfill two requirements, indicated in the data sheet : less than $25mV_{p-p}$ on the "Carrier Input" and less than $6V_{p-p}$ on the "Demod Input". The first condition strongly limits the noise performance, specially because we do not filter the output signal of the atomic resonator. The differential output $+V_0$, $-V_0$ is followed by a differential amplifier showing a gain of 20dB and an offset adjustment.

The proper noise of the demodulator in "linear" conditions has been characterized. When the "Demod Input" receives the demodulation waveform and the "Carrier Input" is loaded by 50Ω, one may observe that a slow drift happens during the thermal stabilization and the corresponding spectrum of the output signal shows a $\approx f^{-2}$ dependence in very low frequency. When stable thermal conditions have been reached, the spectrum of the output signal shows the classical f^{-1} dependence in very low frequency. In the fluctuations of the locked VCXO, this noise characteristics induces a "flicker frequency noise" which gives the well known "flicker floor" of the Allan variance. In our experimental conditions, the calculation by means of the classical formula [8] gives the ultimate limit at 3.10^{-13} expressed in Allan standard deviation. This value is very bad for a frequency standard but quite sufficient for our measurements : taking into account the usual performance of the Cs3 resonator used for the experimental verification, it only limits the sampling time to τ<45s, or else the measurement is not significant.

III.3. Overall experimental apparatus

The experimental apparatus is shown in figure 3. It enables us to simultaneously determine the spectrum of the interrogation signal at 9192MHz and the corresponding frequency stability of the locked VCXO. The detail of the operation has been described in [9]. Briefly, DRO_1 and DRO_2 are identical low noise microwave generators, where a 9192MHz signal delivered by a Dielectric Resonator Oscillator (DRO) is phase-locked to a 10MHz signal (VCXO). The VCXO signal will be frequency-locked to the atomic transition frequency, with a time constant of 1s. The functional scheme of $DRO_{1,2}$ is shown in figure 4.

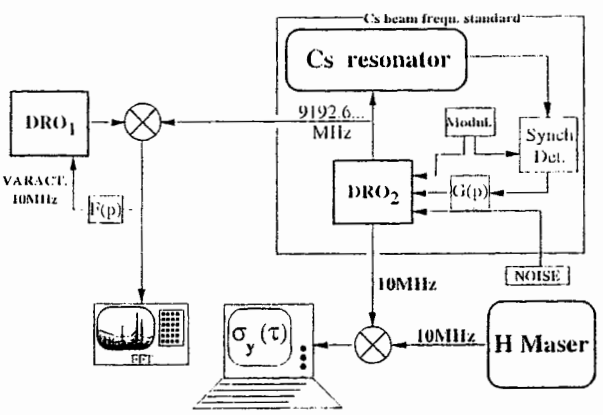


Figure 3: Overall experimental apparatus

Additional noise may be injected into the PLL of the DRO, which gives a controlled degradation of the phase noise in the probe signal. In order to easily adjust the corresponding Fourier spectrum, we do not use a real dual notch filter : the effect of the notch is present -or not- in the additional noise generated by an FFT analyzer Spectral Dynamics 380.With this system, the initial spectrum of the probe signal can be degraded by as much as 35dB in the Fourier frequency range of interest. In our operating conditions, without additional phase noise, the "natural" performance of the cesium resonator is in the range 2 to $2.5\ 10^{-12}\ \tau^{-1/2}$, depending on ambient conditions.

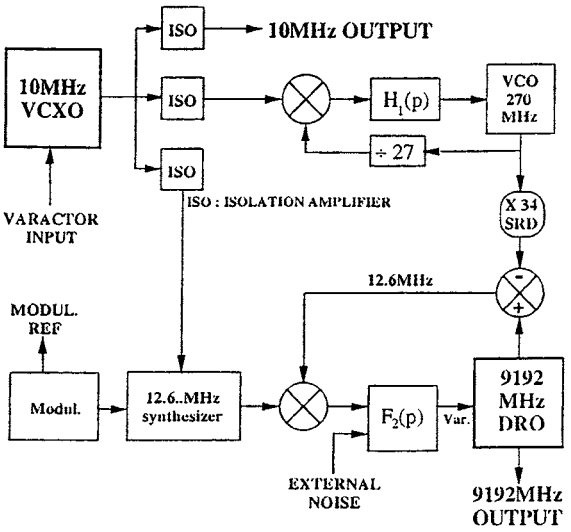


Figure 4 : Microwave generators DRO1,2

IV. Experimental results

The following experimental results have been obtained with a square wave frequency modulation at f_M = 35 Hz in the probe signal. Various additive noises have been injected in the PLL of the DRO.

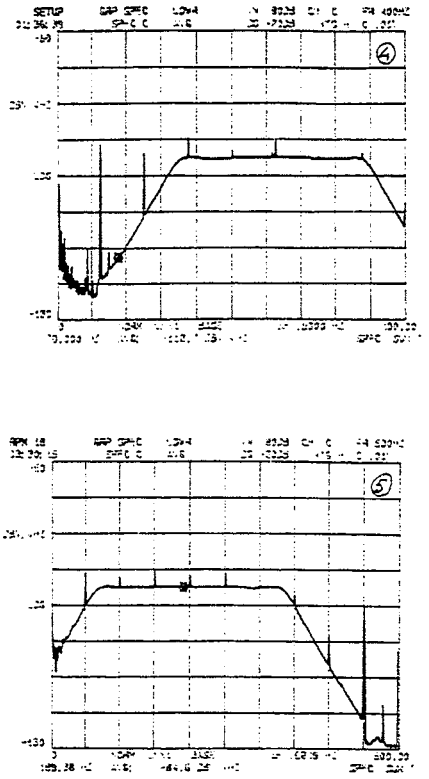
IV.1. Preliminary measurements

We have observed that the transient signals ($85mV_{p-p}$) occuring during the frequency switching were larger than the amplitude required for a linear behavior of the 1496 circuit. With the low modulation frequency that is used (35Hz), the corresponding disturbance is a very small fraction of the modulation period. Nevertheless we have verified that these transient signals do not significantly disturb the linearity of the analog demodulator. For that, we have measured the (disturbed) frequency stability of the locked VCXO for τ = 9.17s :

- in the usual conditions -that is with a selective filter at the output of the resonator- by means of a Brookdeal 9503 synchronous detection

- then without a selective filter at the output of the resonator, by means of the analog demodulator driven by a sine wave demodulation function, which is theoretically equivalent to the Brookdeal measurement.

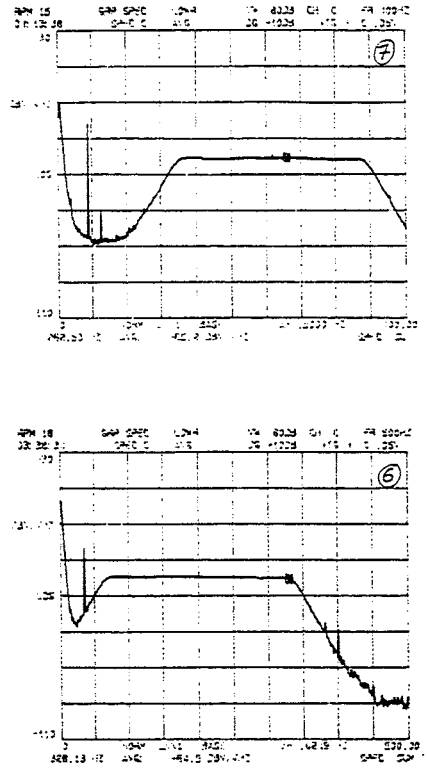
With the upper noise profile shown in figures 5-6 and mentioned just before table 2, the frequency stability for τ = 9.17s is 1.74 (10^{-12}) with the Brookdeal synchronous detection, 1.70 (10^{-12}) with the analog demodulator. With the noise profile mentioned just before table 3, the result is 2.85 (10^{-12}) with the Brookdeal synchronous detection, 2.76 (10^{-12}) with the analog demodulator. Taking into account the number of samples (90), we obtain error bars of about 10% : the previous results may be considered as identical and hence the behavior of the analog demodulator is satisfactory in the noise conditions of tables 2 and 3.



Rejection $2f_M$

Normal

Fig.5 : injected additive noise



Rejection $2f_M$

Normal

Fig.6 : measured phase noise (9192MHz)

IV.2. Effect of the third harmonics in the demodulation

We have measured the frequency stability of the locked VCXO in presence of the noise profile shown in figures 5-6. The phase noise is flat in the range 140 Hz-350 Hz, where its power spectral density is -50 dBc/Hz. The rejection measured at $2f_M$ is ≈ 24 dB, except for the measurement reported in the last line of table 2 (**) where the phase noise at 9192MHz is -50 dBc/Hz from f = 70 Hz to 330 Hz (no rejection).

We have changed the spectrum of the demodulation waveform particularly by means of its third harmonics rate. Three interesting results are observed in table 2 :

- when the rejection at $2f_M$ is used, the corresponding improvement is only a factor 1.4, instead of the theoretical value 1.87 or the simulation result 1.7. It confirms that with a square wave frequency modulation, it is not possible to strongly reduce the intermodulation effect only by a dual notch filter. [As the additive noise spectra in the usual case and in the rejection condition are not quite identical near $10f_M$, the real improvement should be a little bit higher than 1.4, which could explain the previous moderate discrepancy]

- the 3rd harmonics rate $C_3 = -0.2$ induces a significant improvement probably limited by the natural performance of the Cs3 resonator (≈ $0.8\ 10^{-12}$ for $\tau = 9.17$s). It is interesting to notice that, in spite of a limited additive noise spectrum, this first optimization is obtained for a value C_3 very similar to that of the theoretical results (-0.193) with a quasi infinite white phase noise with rejection at $2f_M$

- the use of a pseudo-square demodulation waveform is not favorable at all. A similar observation has been made with the Brookdeal synchronous detection when the output of the atomic resonator is not filtered, leading to a frequency stability $2.74\ 10^{-12}$ for $\tau = 9.17$s.

- Square wave frequency modulation -

Demodulation	$\sigma_y(\tau = 9.17s)$ $/10^{-12}$	N (samples)
$C_3 = 0$	1.70	90
$C_3 = -0.2$	1.01	90
$C_3 = -0.4$	1.70	90
$C_3 = +0.4$	3.74	90
Pseudo square $C_{2n+1} = \frac{1}{2n+1}$ up to n=4	5.61	90
C_3=0, no rejection at $2f_M$ **	2.38	90

Table 2 - Effect of various demodulation waveforms

IV.3. Effect of the optimized demodulation

In order to verify more precisely the theoretical predictions in our operating conditions, we have taken into account the limited spectrum of the additive noise and calculated the corresponding optimized demodulations with (C_3,C_5), then (C_3,C_5,C_7). These values are indicated in table 3. Various C_3 configurations have also been checked. Furthermore, we have increased the power of the additive noise up to the highest possible level in order to strongly degrade the frequency noise of the probe signal. In those conditions, the phase noise at 9192MHz is -46 dBc/Hz from f = 140 Hz to 350 Hz and the rejection at $2f_M$ (70Hz) is approximately 27dB. The shape of the noise spectrum at 9.2GHz is similar to that of figure 6.

Corresponding experimental results are shown in table 3.

- Square wave frequency modulation -

Demodulation	$\sigma_y(\tau = 9.17s)$ $/10^{-12}$	N
$C_3 = 0$	2.76	90
$C_3 = -0.1$	1.75	90
$C_3 = -0.2$	1.31	90
$C_3 = -0.3$	1.50	90
$C_3 = -0.18745$ and $C_5 = -0.02310$	1.07	270
$C_3 = -0.19037$ and $C_5 = -0.02460$ and $C_7 = -0.00661$	1.02	180

Table 3 - Effect of optimized demodulation waveforms

The performance of the system is improved when we use a larger number of optimized harmonics. In the last configuration, the frequency stability without an additive noise is 8.10^{-13}. Hence, the corresponding contribution of the additive noise may be estimated about 6.10^{-13}. In this case, the spurious effect related to the phase noise of the local oscillator is reduced by a factor of 4.5 compared to the case of the dual notch filter alone. The expected theoretical improvement given by table 1 is 14.17/1.89 ≈ 7.5, whilst our experimental estimate is 4.5. It is possible that the quasi-static assumption limits the precision of our determination (waveform and expected improvement). In any case, the vicinity of the noise floor of the experiment does not enable us to have a precise estimate of the obtained improvement, but these measurements show that it is possible to significantly improve the NIST method [3] by means of a proper signal processing, according to the theory developped in [4], [5] and [7].

V. Conclusion

Experimental results undoubtedly show the efficiency of a proper demodulation coupled with the use of a dual notch filter in the case of a square wave frequency modulation. It is now possible to obtain the stability of the atomic resonator itself with a relatively noisy local oscillator in the case of a square wave frequency modulation. A more interesting improvement is expected with a sine wave modulation, but the direct measurement of this improvement will be difficult.

VI. Acknowledgements

We are very grateful to **DRET** (Direction des Recherches, Etudes et Techniques) who has supported this work and made it possible to obtain the previous results.

We also are grateful to **OMEGA TECHNOLOGIES** deeply involved in our studies.

We are happy to gratefully acknowledge Bertrand **BOUSSERT** for his excellent adjustment of the Cs3 resonator and **Benoit CARDOSO** whose technical contribution has been very important in the experimental apparatus.

References

[1] G. Kramer "Noise in Passive Frequency Standards." CPEM 1974 Digest, IEEE Conference Publication n°113, IEEE Catalog n° CH 0770-8IM (unpublished).

[2] C.Audoin, V. Candelier, and N. Dimarcq "A Limit to the Frequency Stability of Passive Frequency Standards due to an Intermodulation Effect". IEEE Trans. on Instr. and Measur. vol 40, n°2, april 1991, pp 121-125.

[3] C. Szekely, F.L. Walls, L.P. Lowe, R.E. Drullinger and A. Novick. "Reducing Local Oscillator Phase Noise Limitations on the Frequency Stability of Passive Frequency Standards : Tests of a new concept"- IEEE Trans. UFFC, Vol 41 N°4, july 1994, pp 518-521.

[4] D. Venot, R. Barillet and C. Audoin. "On a Possibility of Reducing the Frequency Stability Limitation by the Interrogation Noise in Passive Frequency Standards" Proceedings of the 1994 EFTF, Munich, march 94, pp 656-670

[5] D. Venot, R. Barillet and C. Audoin. "A possibility of reducing the effect of interrogation phase noise on the frequency stability in passive frequency standards". CPEM 95, june 1995, Boulder, USA (unpublished)

[6] G. Santarelli, P. Laurent, S.N.Lea, A. Nadir and A. Clairon."Preliminary results in a cesium atomic fountain frequency standard". Proceedings of the 1994 EFTF, Munich, march 94, pp 46-57

[7] D.Venot. "Démodulation synchrone et effet d'intermodulation"- Rapport interne - LHA, février 1994

[8] J.Vanier and C. Audoin. "The quantum physics of atomic frequency standards", Adam Hilger, Vol.1 p238

[9] R. Barillet, V. Giordano, J. Viennet and C. Audoin."Limitation of the clock frequency stability by the interrogation frequency noise : experimental results". IEEE Trans. on Instr. and Meas., vol 42, N°2, april 1993, pp 276-280.

Confidence on the Modified Allan Variance And the Time Variance*

Marc A. Weiss[1], F. L. Walls[1], Chuck Greenhall[2], Todd Walter[3]

[1]Time and Frequency Division, National Institute of Standards and Technology, 325 Broadway, Boulder, Colorado 80303, USA

[2]Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr., Pasadena, California 91109, USA

[3]HEPL Gravity Probe B, Stanford University, Stanford, California 94305-4085, USA

ABSTRACT

This paper presents tabulated factors for calculating confidence intervals for the square root of the Modified Allan Variance and the related Time Variance when the analysis uses full overlapping of the data. Confidence intervals are determined by multiplying factors from the appropriate table times sample deviations. These factors have been calculated from the equivalent degrees of freedom (edf) for the five common power law noise types. Closed form expressions for calculating the edf for these variances and noise types have only recently become available for fully overlapped data analysis. We compute the edf using two different closed form expressions as well as simulations to ensure correctness of the results. We also show the relative advantage of using the modified Allan variance instead of the Allan variance to evaluate frequency or time stability in the presence of white phase modulated noise.

INTRODUCTION

The focus of this paper is the discussion and presentation of factors for determining confidence intervals for the Modified Allan Variance (MVAR) [1] and the related Time Variance (TVAR) [2]. Like the Allan Variance (AVAR) [3], these variances have become standard tools for analyzing high-performance signal sources, signal processing devices, and measurement systems. AVAR, MVAR, and TVAR have been adopted by the IEEE and the ITU for characterizing such systems. This is because dominant noise processes in those systems have nonwhite spectra sometimes called colored noise. The standard sample variance typically is dependent on the data length for nonwhite noise types and is not defined in the limit as the data length approaches infinity. Hence, the standard sample variance is not satisfactory for characterizing the frequency or time stability of high-performance frequency sources and associated equipment.

Five power law noise types are commonly found in such equipment: white phase noise modulation (PM), flicker PM, random walk PM, flicker frequency modulation (FM), and random walk FM. The spectra associated with these noise types obey power laws in the sense that the spectrum of a measurement of time offset x(t) dominated by one of these noise types is of the form

$$S_x(f) = gf^{\beta}, \qquad (1)$$

where $\beta = 0, -1, -2, -3, -4$, respectively, for the five noise types. We may also consider the spectrum of the frequency offset y(t), the derivative of x(t), which would be of the form

$$S_y(f) = hf^{\alpha}, \qquad (2)$$

where $\alpha = +2, +1, 0, -1, -2$, respectively, for the five noise types. A random walk modulation of phase is the same as a white noise modulation of frequency.

Each of the three variances, AVAR, MVAR, and TVAR, has a different dependence on averaging time τ for the five different noise types. For a given noise type, each variance has a power-law dependence τ. This is illustrated in Figure 1. Thus, using these variances it is possible to measure the dominant noise type for different averaging times, if the spectrum is a sum of power-law spectra.

---

*Contribution of U. S. Government, not subject to copyright.

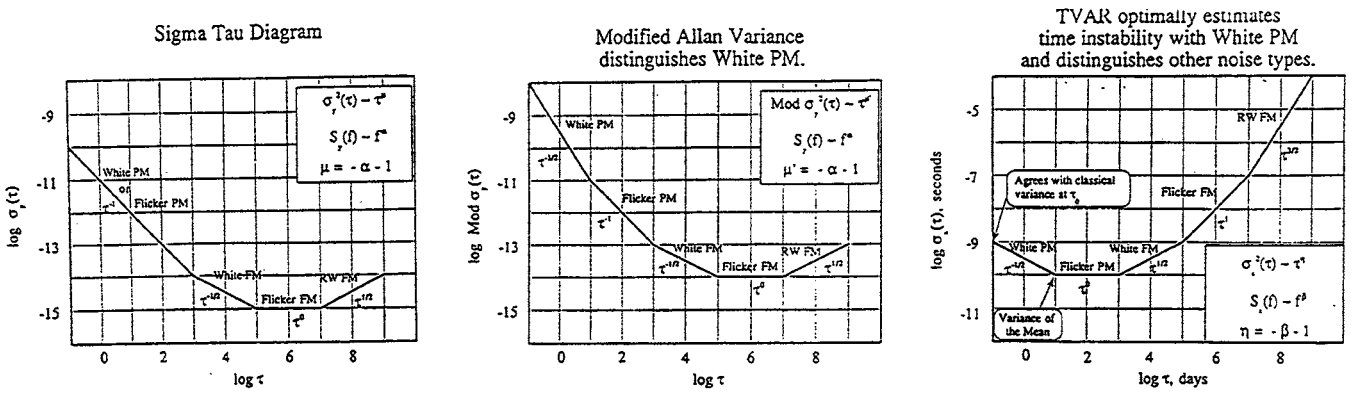


Figure 1 Illustration of the three variances, the Allan, Modified Allan, and Time Variances, and their relation to the power spectrum.

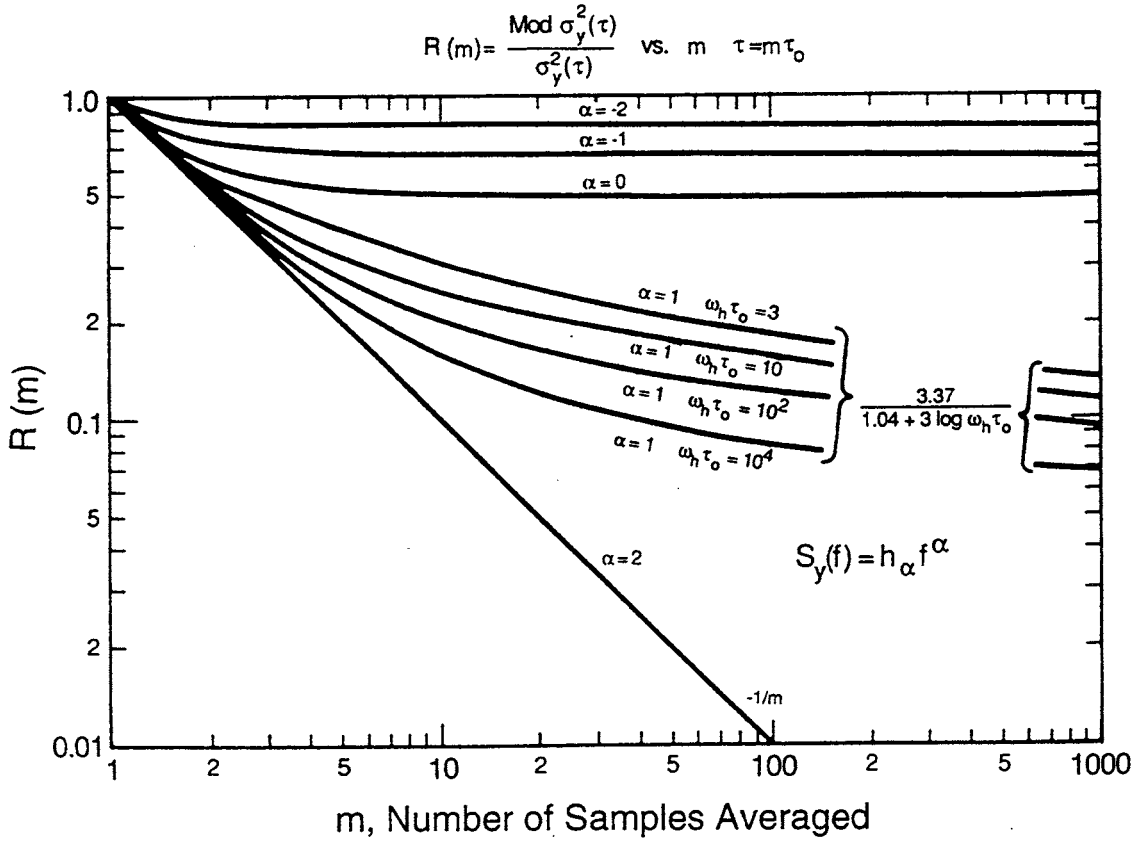


Figure 2 The ratio of the Modified Allan Variance to the Allan Variance as a function of the number of samples in the estimate for different noise types and bandwidths.

There is an advantage to using MVAR over AVAR in the presence of white PM, since MVAR falls off faster for white PM than AVAR. Figure 2 shows the ratio of MVAR to AVAR. We see that for white PM the ratio drops with a constant slope, unlike for the other noise types. This allows MVAR to measure the levels of flicker PM, white FM, flicker FM and random walk FM in the presence of white PM for shorter values of τ than AVAR. Though this advantage is reduced somewhat because the confidence intervals are larger for MVAR than for AVAR, MVAR still provides an improvement over AVAR. Additionally, MVAR can distinguish between white and flicker PM, whereas AVAR cannot. TVAR has all the same advantages as MVAR, the only difference being its power law dependence on τ as indicated in Figure 1 and its level. Often white PM comes from a different part of the system being measured than other noise types. Thus it is valuable to be able to average down white PM optimally to characterize the other noise types.

COMPUTATION OF CONFIDENCE INTERVALS

If s^2 denotes the usual sample variance of n independent and identically distributed Gaussian measurements (i.e., white noise) with actual variance σ^2, then it is well-known that

$$U = \frac{s^2}{\sigma^2}\cdot\nu. \qquad (3)$$

has a chi-square distribution with $\nu = n-1$ degrees of freedom [4]. In the classical situation, the degrees of freedom associated with σ^2, are integer values depending only on the number of measurements, and exact confidence limits on the measurement variance are easily calculated using percentiles of the appropriate chi-square distribution. For example, figure 3 shows the chi-square distribution with 10 degrees of freedom, and also depicts the percentiles a and b that are needed to calculate a uncertainty bounds on σ^2 at the $p = 0.95$ confidence level from a particular s^2 based on 11 Gaussian measurements.

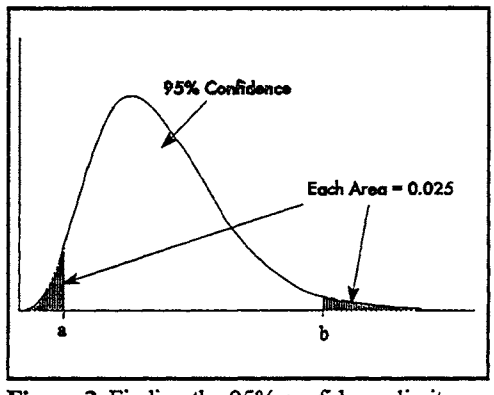


Figure 3 Finding the 95% confidence limits under the chi-squared distribution with 10 degrees of freedom.

A 95% confidence interval is obtained as follows. First the probability 0.95 that U of equation (3) lies between a and b. This condition is equivalent to the following:

$$\begin{aligned} & a < \frac{s^2}{\sigma^2}\cdot\nu < b \\ & \textit{if and only if} \\ & \frac{\nu}{b} < \frac{\sigma^2}{s^2} < \frac{\nu}{a} \\ & \textit{if and only if} \\ & \frac{\nu}{b}\cdot s^2 < \sigma^2 < \frac{\nu}{a}\cdot s^2. \end{aligned} \qquad (4)$$

While the lower and upper bounds in the final inequality are confidence limits on the actual sample variance, note that the confidence factors ν/b and ν/a needed in the calculations are indepent of the actual data. They give the magnitude of the confidence interval as a function of the number of points used to compute the variance. Hence we can compute these confidence factors for each noise type and various data lengths. The factors $1-\nu/b$ and $\nu/a-1$ give the multipliers for the magnitude of the lower and upper confidence intervals, respectively, and these values are presented at the end of this for each of the five noise types and for varous data lengths. We also use these values to characterize the relative merits of AVAR and MVAR for estimating noise processes in the presence of white PM.

Since the common time and frequency stability measures (AVAR, MVAR, TVAR) are calculated from data arising from non-white noise processes, the confidence limit procedure outlined above is an approximate method [5]. The method is based on approximating the

distribution of U in (3) with the chi-square distribution with degrees of freedom

$$\nu = \frac{2(\sigma^2)^2}{Var\ (s^2)}. \qquad (5)$$

where σ^2 represents the appropriate stability measure (e.g., MVAR), s^2 represents its corresponding estimator, and $Var(s^2)$ is the expected variance of the s^2 estimators. The quantity ν here is called the equivalent degrees of freedom, edf, since it need not be integer-valued. Equations for the edf of MVAR and TVAR have only recently been published in closed form [6], and other papers have been submitted for publication [7],[8].

We generated tables of the factors $1-\nu/b$ and $\nu/a-1$ by computing the edf using two different formulas and verified the theoretical results by simulating corresponding noise processes. Both theoretical solutions express MVAR as a function of the discrete autocorrelation function, and evaluate (5) for autocorrelation functions associated with respective noise types. One method uses the conventional definition of MVAR as a variance of second differences of averages, whereas the other technique uses an equivalent definition of MVAR in terms of third differences [9]. While the definition of MVAR is equivalent in the two formulations, the autocorrelation functions differ, so the numerical solutions for calculating the edf differ as well. The two methods agreed in all results.

SIMULATION

To check the theory we simulated power-law noise series and calculated estimated edf's by substituting appropriate estimates of σ^2 and $Var(s^2)$ in (5). We generated data consistent with a model of white PM using a random number generator, computed cumulative sums of the data sequence to simulate white FM, then again computed cumulative sums to simulate random walk FM. We used the Barnes-Jarvis filter [10],[11] to simulate flicker PM, then computed cumulative sums to simulate flicker FM. Series lengths N appearing in tables at the end of this paper were generated in each case.

For each noise type and series length N, we generated 10,000 time series and computed MVAR for each series. For every set of 10,000 MVAR estimates, corresponding sample mean and sample variance of the 10,000 obverved values of MVAR were substituted, respectively, for $\sigma^2$and $Var(s^2)$ in equation (5), thereby producing an estimated value of the edf, ν. Estimated values of the edf. from the simulation agreed with the numerically computed values within a few percent for white PM, white FM, and random walk FM noises. For flicker noise, the match was not as close for small values of m, where $\tau=m\cdot\tau_0$. This is consistent with recent publications showing that the effects of discrete sampling, as reflected for example in the autocorrelation function, produce different effects than the output of the Barnes-Jarvis filter [6].

The edf from the two theoretical methods and from simulation are presented in Table 1 for N=1025 and for white PM (whpm) and flicker PM (flpm) noise, respectively. The table illustrates the agreement.

Table 1
Comparison of Degrees of Freedom
From Three Methods
N = 1025

	Theoretical Methods		Simulation	
m	whpm	flpm	whpm	flpm
1	526.4	589.9	527.66	630.38
2	477.4	497.0	493.24	496.87
4	298.7	263.1	294.80	257.91
8	158.2	128.4	156.68	126.50
16	78.96	62.33	80.94	61.84
32	38.15	29.85	38.80	29.96
64	17.62	13.73	17.69	13.69
128	7.40	5.74	7.29	5.71
256	2.85	2.07	2.73	2.07

USE OF OVERLAPPING ESTIMATES

There are two common methods for the estimation of the three variances, AVAR, MVAR, and TVAR. One method uses fully overlapping and the other nonoverlapping time intervals. We will explain the meaning of these terms, then discuss their relative advantages. To understand the potential for overlapping estimates we will use the definitions of AVAR, MVAR, and TVAR, and use another common notation for them, $\sigma^2_y(\tau)$, $\mathrm{mod}.\sigma^2_y(\tau)$, and $\sigma^2_x(\tau)$, respectively.

Theoretically we can define these variances independent of sampling. Suppose there is a time-difference signal, x(t), between two clocks. Then, at least conceptually, for AVAR we would apply an operator which finds one-half the ensemble average of the mean second difference squared. For MVAR the operator would extract one-half the ensemble average of the square of the time averaged second-differences. This process is illustrated in Figure 4. TVAR is simply τ^2·MVAR/3. Thus,

$$\sigma^2_y(\tau) = \frac{1}{2\tau^2}\left\langle \left(\Delta^2_\tau x\right)^2 \right\rangle$$

$$\text{mod}\ .\sigma^2_y(\tau) = \frac{1}{2\tau^2}\left\langle \left(\Delta^2_\tau \bar{x}^\tau\right)^2 \right\rangle \quad (6)$$

$$\sigma^2_x(\tau) = \frac{1}{6}\left\langle \left(\Delta^2_\tau \bar{x}^\tau\right)^2 \right\rangle$$

Here, Δ^2_τ is the second difference operator over the interval τ, $<\ldots>$ is the infinite ensemble average, and $\bar{x}^\tau$ is generated from x by averaging over an interval of length τ.

For computing estimates of these variances, we assume stationarity of the second difference of the data and compute time averages of the square of the second differences. In forming the second differences in AVAR, as well as the $\bar{x}^\tau$ in MVAR and TVAR, we can choose where to start the next second difference relative to the last one. This allows the option of overlapping the second differences. Figure 4 illustrates a second difference starting with $\bar{x}^\tau_n$. The next second difference can begin from a point either τ_0 or τ increased from this one. Equations (7) and (8) below give the computational definitions.

If there are N data points with an even spacing of τ_0, then the variances are functions of $\tau = m \cdot \tau_0$. Suppose we have sampled our time difference signal x(t) with a time series $\{x_n\}$. Then the estimates of AVAR, MVAR, and TVAR using fully overlapping samples would be computed as in equation (7) below.

The computational equations for estimates of AVAR, MVAR, and TVAR using nonoverlapping samples take the form in equation (8) below where [A] is the greatest integer less than A.

The difference between the fully overlapping and non-overlapping estimates is in how the summation parameter n increases. Either we would increase by the minimum spacing, or by m intervals. The latter means that we would use different data for intervals of length τ included in each second difference. Figure 5 plots the ratio of the confidence factors for the use of non-overlapping over fully overlapping samples for the Allan Variance. The table shows that the use of fully overlapping estimates is better primarily for noise types of white and flicker PM.

ADVANTAGES OF MVAR OVER AVAR

AVAR and MVAR are standard tools for the characterization of frequency standards and measurement systems. Often we measure the frequency stability between frequency standards through measurement systems whose noise may exceed the performance of the standards in short-term. The noise types of measurement systems are typically white and flicker PM, whereas the noise types of the frequency standards are often white, flicker and random walk FM. If the noise type of the measurement system is white PM, we can average it down faster to potentially determine the performance of frequency standards if we use MVAR instead of AVAR. This is because the ratio of the Modified Allan Deviation over the Allan Deviation for white PM is proportional to $\tau^{-1/2}$, where τ is the averaging time, the Modified Allan Deviation (MDEV) is the square root of MVAR, and the Allan Deviation (ADEV) is the square root of AVAR. Figure 2 shows a plot of the ratio of MVAR over AVAR for different noise types and bandwidths. Note that flicker PM is bandwidth dependent. This Figure shows that the advantage of using MVAR is primarily in the case of white PM. Flicker PM shows some advantage, though the loss of confidence, as discussed below, destroys any improvement of MVAR over AVAR in this case.

We illustrate the advantage of MVAR over AVAR in Table 3 for the case of a data set of length N=1025 dominated by a white PM noise, using a confidence interval for a probability of 68%. This advantage should be understood as the combination of two effects. MVAR has an advantage over AVAR because it drops faster for white PM, but it loses some of that advantage because the confidence intervals for MVAR are poorer than for AVAR. The advantage of MVAR over AVAR due to its faster averaging of white PM is presented for our particular example as the ratio of the deviations, ADEV over MDEV, in column 2 of Table 3. The advantage of AVAR over MVAR is due to its smaller confidence intervals. We define σ(ADEV) and σ(MDEV) as the root-mean-square of the upper and lower confidence intervals for ADEV and MDEV respectively, using a probability of 0.68. Thus, σ(ADEV) and σ(MDEV) are the equivalent of standard deviations around ADEV and MDEV respectively. The ratio of σ(ADEV) over σ(MDEV) is shown in column 3 of Table 3, illustrating the advantage of using AVAR. The advantage of MVAR over AVAR is presented in column 4. This is the ratio of column 2 over column 3.

$$\sigma^2_y(\tau) = \frac{1}{2(N-2m)\tau^2}\sum_{n=1}^{N-2m}(x_{n+2m} - 2x_{n+m} + x_n)^2,$$

$$\mathrm{mod}.\sigma^2_y(\tau) = \frac{1}{2(N-3m+1)\tau^2}\sum_{n=1}^{N-3m+1}\left(\frac{1}{m}\sum_{k=0}^{m-1}(x_{n+k+2m} - 2x_{n+k+m} + x_{n+k})\right)^2, \quad (7)$$

$$\sigma^2_x(\tau) = \frac{\tau^2}{3}\mathrm{mod}.\sigma^2_y(\tau).$$

$$\sigma^2_y(\tau) = \frac{1}{2[(N-2m-1)/m]\tau^2}\sum_{n=0}^{[(N-2m-1)/m]}(x_{nm+1+2m} - 2x_{nm+1+m} + x_{nm+1})^2,$$

$$\mathrm{mod}.\sigma^2_y(\tau) = \frac{1}{2[(N-3m)/m]\tau^2}\sum_{n=0}^{[(N-3m)/m]}\left(\frac{1}{m}\sum_{k=0}^{m-1}(x_{nm+1+k+2m} - 2x_{nm+1+k+m} + x_{nm+1+k})\right)^2, \quad (8)$$

$$\sigma^2_x(\tau) = \frac{\tau^2}{3}\mathrm{mod}.\sigma^2_y(\tau).$$

Table 3
Comparison of AVAR and MVAR
Noise Type = White PM
Confidence = 68%, Data Length = 1025

m	$\frac{\mathrm{ADEV}}{\mathrm{MDEV}}$	$\frac{\sigma(\mathrm{ADEV})}{\sigma(\mathrm{MDEV})}$	$\frac{\text{column 2}}{\text{column 3}}$
1	1.00	1.00	1.00
2	1.41	1.09	1.30
4	2.00	1.33	1.51
8	2.83	1.83	1.54
16	4.00	2.59	1.55
32	5.66	3.78	1.50
64	8.00	6.02	1.33
128	11.31	10.24	1.10
256	16.00	35.08	0.46

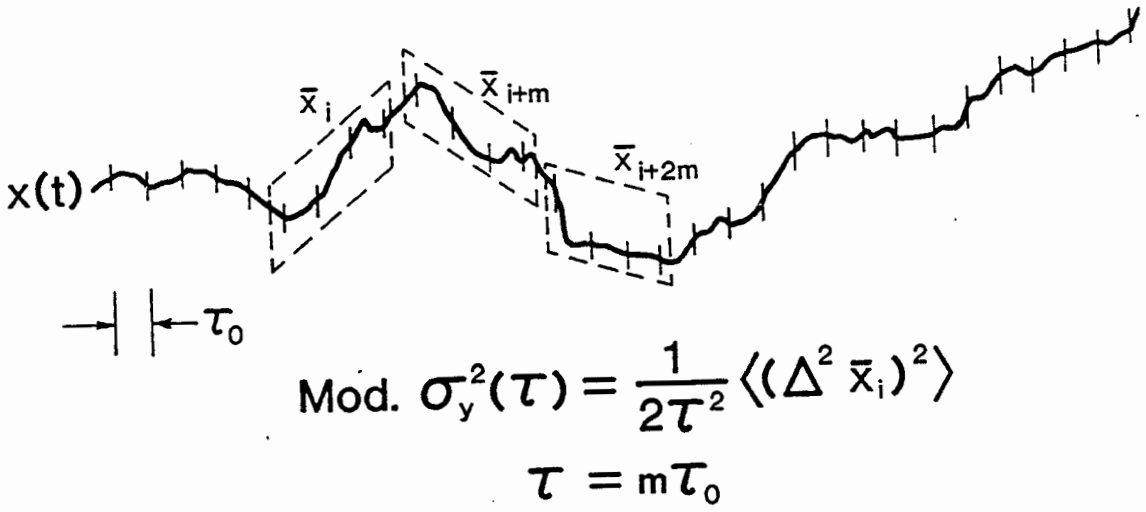


Figure 4 An illustration of the computation of the Modified Allan Variance.

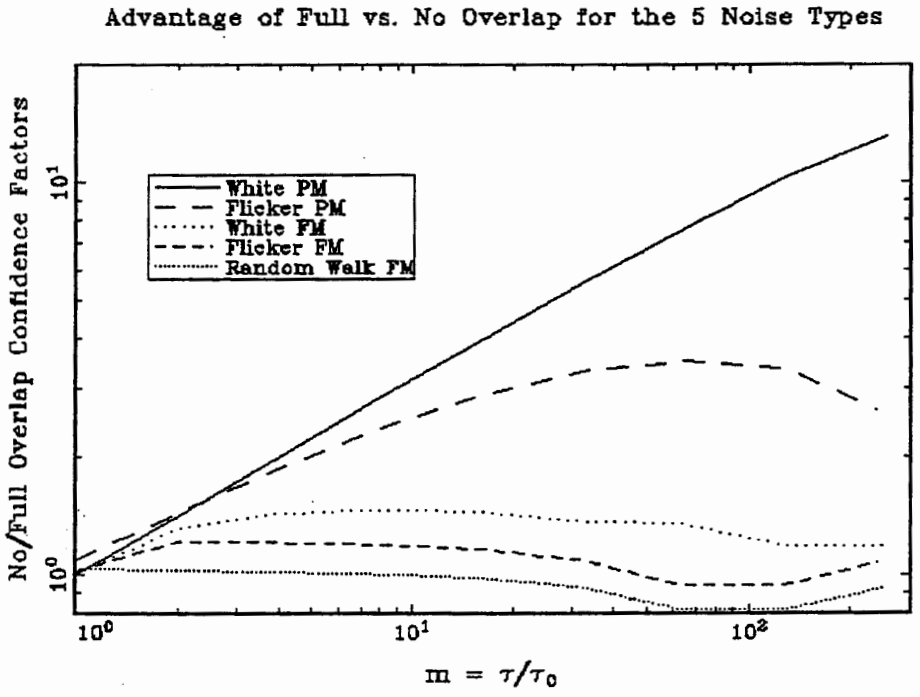


Figure 5 Ratio of the confidence factors for the Allan Variance of data with N= 1025 points using no overlapping data divided by factors using full overlapping data. The factors are larger using no overlapping data leading to poorer confidence limits for the use of no overlap primarily for white and flicker PM.

TABLES OF CONFIDENCE LIMITS

Tables 4-8 give confidence limits for each of the five noise types respectively: white PM, flicker PM, white FM (the same as random walk PM), flicker FM, and random walk FM. In each table we consider cases for data lengths N= 5, 9, 17, 33, 65, 129, 257, 513, and 1025. We cover the possible values of m as a power of 2, where the averaging time is $\tau=m\cdot\tau_0$. Thus m is the number of the minimum intervals between data for the averaging time. Column 3 of each table lists the degrees of freedom from theoretical derivations. Columns 4 and 5 list the lower and upper confidence limits for 68% probability, respectively. Columns 6 and 7 list the same for 95% probability. The values we present in the tables for the lower limits are 1.0 minus the factor needed to multiply times the sample deviation to give the lower limit value as a percent. For the upper limits we present the multiplicative factor minus 1.0. Thus we give the percent change from the sample deviation to give the lower and upper limits corresponding to the given probability. That is, if A and B are the factors such that $A\cdot s$ is the upper limit and $B\cdot s$ is the lower limit for probability p, then the values we present are σ_l for the lower limit and σ_u for the upper limit where

$$\begin{aligned} \sigma_l &= 1.0 - A, \\ \sigma_u &= B - 1.0. \end{aligned} \qquad (9)$$

CONCLUSIONS

MVAR shows benefits over AVAR when the dominant noise type is white PM. In particular, MVAR averages down the white PM faster than AVAR, revealing other noise types for shorter τ values. Even though the confidence is generally poorer for MVAR than for AVAR, the advantage in averaging white PM still prevails, though its main benefit occurs for τ values more toward the middle of the range from the minimum data spacing to the maximum allowable. In addition, while the use of fully overlapping samples always improves the degrees of freedom in estimating AVAR, MVAR, and TVAR, the improvement is best for noise types of white and flicker PM. For white, flicker, and random walk FM, the advantage in the use of fully overlapping samples is insignificant.

REFERENCES

[1] D. W. Allan and J. A. Barnes, "A Modified Allan Variance with increased Oscillator characterization ability," in *Proc. 35th Ann. Frq. Contr. Symp.*, pp. 470-474, May, 1981.

[2] D. W. Allan, M. A. Weiss, J. L. Jespersen, "A frequency-domain view of time-domain characterization of clocks and time and frequency distribution systems," *Proc. of the IEEE Frequency Control Symposium 1991*, pp. 667-678.

[3] D. W. Allan, "Statistics of atomic frequency standards," in *Proc. IEEE*, vol.54. pp.221-230, Feb.1966.

[4] P. G. Hoel, *Introduction to Mathematical Statistics*, New York: John Wiley & Sons, 1971.

[5] *Characterization of Clocks and Oscillators*, Eds. D. B. Sullivan, D. W. Allan, D. A. Howe, and F. L. Walls, National Institue of Standards and Technology Tech Note 1337, 1990.

[6] T. Walter, "Characterizing frequency stability: a continuous power-law model with discrete sampling," *IEEE Trans. on Instrum. and Meas.*, vol.43, pp. 69-79, Feb. 1994.

[7] C. A. Greenhall, "Estimating the modified Allan variance," submitted to the *1995 IEEE Frequency Control Symposium*.

[8] C. A. Greenhall, "The third difference approach to the modified Allan variance," in preparation for the *IEEE Trans. on Instrum. and Meas.*

[9] C. A. Greenhall, "A shortcut for computing the modified Allan variance", *Proc. IEEE Frequency Control Symp.*, 1992, pp. 262-264.

[10] J. A. Barnes, *The Analysis of Frequency and Time Data*, Austron, Inc., 1992, pp.43-44.

[11] C. A. Greenhall, "Initializing a flicker-noise generator," *IEEE Trans. on Instrum and Meas.*, vol. IM-35, no.2, 1986.

Table 4
White PM Confidence Factors
for MDEV and TDEV
Confidence factors times sample deviations
give magnitudes of
lower and upper confidence intervals
for 68% and 95% probabilities

N	m	edf	lower 68%	upper 68%	lower 95%	upper 95%
5	1	1.862	26.50%	152.4%	48.73%	606.7%
9	1	3.885	22.20%	69.42%	40.42%	194.1%
9	2	2.636	24.57%	101.8%	44.82%	330.4%
17	1	7.988	17.70%	38.29%	32.47%	91.70%
17	2	6.216	19.24%	46.49%	35.17%	116.1%
17	4	2.951	23.90%	90.54%	43.54%	279.3%
33	1	16.21	13.67%	23.32%	25.40%	51.69%
33	2	13.66	14.59%	26.14%	27.02%	58.81%
33	4	7.293	18.25%	41.03%	33.44%	99.64%
33	8	3.011	23.78%	88.71%	43.31%	271.4%
65	1	32.67	10.32%	14.99%	19.42%	31.84%
65	2	28.61	10.90%	16.25%	20.47%	34.74%
65	4	16.64	13.53%	22.92%	25.16%	50.70%
65	8	7.508	18.07%	40.13%	33.13%	97.00%
65	16	2.970	23.86%	89.95%	43.47%	276.8%
129	1	65.58	7.656%	9.956%	14.57%	20.61%
129	2	58.53	8.046%	10.63%	15.29%	22.07%
129	4	35.43	9.974%	14.27%	18.80%	30.20%
129	8	17.49	13.27%	22.18%	24.70%	48.89%
129	16	7.488	18.09%	40.21%	33.16%	97.24%
129	32	2.920	23.96%	91.52%	43.66%	283.6%
257	1	131.4	5.607%	6.748%	10.77%	13.75%
257	2	118.4	5.879%	7.146%	11.28%	14.59%
257	4	73.04	7.300%	9.363%	13.92%	19.32%
257	8	37.57	9.731%	13.78%	18.36%	29.09%
257	16	17.64	13.23%	22.06%	24.62%	48.59%
257	32	7.444	18.13%	40.39%	33.22%	97.77%
257	64	2.885	24.04%	92.66%	43.80%	288.6%
513	1	263.1	4.066%	4.634%	7.863%	9.342%
513	2	238.1	4.261%	4.889%	8.232%	9.867%
513	4	148.3	5.305%	6.315%	10.20%	12.84%
513	8	77.76	7.101%	9.037%	13.55%	18.62%
513	16	38.07	9.677%	13.67%	18.26%	28.84%
513	32	17.64	13.23%	22.06%	24.62%	48.59%
513	64	7.414	18.15%	40.52%	33.26%	98.13%
513	128	2.864	24.08%	93.36%	43.88%	291.7%
1025	1	526.4	2.928%	3.212%	5.693%	6.430%
1025	2	477.4	3.068%	3.381%	5.960%	6.773%
1025	4	298.7	3.831%	4.331%	7.416%	8.718%
1025	8	158.2	5.150%	6.097%	9.910%	12.38%
1025	16	78.96	7.053%	8.959%	13.46%	18.45%
1025	32	38.15	9.669%	13.66%	18.25%	28.81%
1025	64	17.62	13.24%	22.08%	24.63%	48.63%
1025	128	7.396	18.17%	40.59%	33.29%	98.35%
1025	256	2.854	24.10%	93.70%	43.92%	293.2%

Table 5
Flicker PM Confidence Factors
for MDEV and TDEV
Confidence factors times sample deviations
give magnitudes of
lower and upper confidence intervals
for 68% and 95% probabilities

N	m	edf	lower 68%	upper 68%	lower 95%	upper 95%
5	1	2.020	26.08%	137.8%	47.82%	518.7%
9	1	4.300	21.56%	63.35%	39.27%	172.0%
9	2	2.676	24.49%	100.2%	44.65%	322.9%
17	1	8.903	17.04%	35.33%	31.33%	83.34%
17	2	6.427	19.03%	45.27%	34.81%	112.3%
17	4	2.560	24.75%	105.1%	45.15%	345.8%
33	1	18.12	13.09%	21.68%	24.38%	47.65%
33	2	14.18	14.39%	25.49%	26.66%	57.15%
33	4	6.394	19.06%	45.45%	34.86%	112.9%
33	8	2.362	25.21%	114.8%	46.06%	393.9%
65	1	36.56	9.844%	14.01%	18.56%	29.60%
65	2	29.74	10.73%	15.87%	20.16%	33.86%
65	4	14.63	14.22%	24.96%	26.37%	55.81%
65	8	6.066	19.39%	47.40%	35.44%	118.9%
65	16	2.223	25.55%	123.1%	46.75%	436.7%
129	1	73.45	7.282%	9.333%	13.88%	19.26%
129	2	60.89	7.909%	10.39%	15.04%	21.55%
129	4	31.18	10.52%	15.42%	19.78%	32.82%
129	8	14.17	14.39%	25.50%	26.67%	57.18%
129	16	5.882	19.59%	48.59%	35.78%	122.6%
129	32	2.145	25.75%	128.3%	47.15%	465.0%
257	1	147.2	5.323%	6.341%	10.24%	12.89%
257	2	123.2	5.774%	6.991%	11.08%	14.26%
257	4	64.31	7.722%	10.07%	14.69%	20.85%
257	8	30.48	10.62%	15.64%	19.96%	33.32%
257	16	13.90	14.49%	25.84%	26.86%	58.03%
257	32	5.796	19.68%	49.18%	35.94%	124.5%
257	64	2.104	25.86%	131.2%	47.37%	481.4%
513	1	294.8	3.855%	4.362%	7.461%	8.780%
513	2	247.8	4.182%	4.785%	8.083%	9.653%
513	4	130.6	5.622%	6.771%	10.80%	13.79%
513	8	63.12	7.785%	10.18%	14.81%	21.09%
513	16	30.03	10.68%	15.78%	20.08%	33.64%
513	32	13.78	14.54%	25.99%	26.94%	58.41%
513	64	5.757	19.72%	49.45%	36.02%	125.3%
513	128	2.084	25.91%	132.7%	47.47%	489.8%
1025	1	589.9	2.773%	3.026%	5.395%	6.052%
1025	2	497	3.010%	3.310%	5.849%	6.629%
1025	4	263.1	4.066%	4.634%	7.863%	9.342%
1025	8	128.4	5.666%	6.834%	10.88%	13.93%
1025	16	62.33	7.828%	10.25%	14.89%	21.25%
1025	32	29.85	10.71%	15.84%	20.13%	33.78%
1025	64	13.73	14.56%	26.05%	26.97%	58.58%
1025	128	5.740	19.74%	49.57%	36.05%	125.7%
1025	256	2.073	25.94%	133.6%	47.53%	494.5%

Table 6
White FM Confidence Factors for
MDEV and TDEV
Confidence factors times sample deviations
give magnitudes of
lower and upper confidence intervals
for 68% and 95% probabilities

N	m	edf	lower 68%	upper 68%	lower 95%	upper 95%
5	1	2.250	25.49%	121.4%	46.61%	427.6%
9	1	4.900	20.73%	56.55%	37.80%	148.5%
9	2	2.703	24.43%	99.17%	44.54%	318.0%
17	1	10.23	16.23%	31.96%	29.90%	74.08%
17	2	6.618	18.85%	44.24%	34.49%	109.2%
17	4	2.356	25.23%	115.2%	46.09%	395.6%
33	1	20.89	12.38%	19.78%	23.11%	43.06%
33	2	14.67	14.20%	24.92%	26.34%	55.70%
33	4	6.096	19.36%	47.22%	35.39%	118.3%
33	8	2.097	25.88%	131.8%	47.40%	484.3%
65	1	42.22	9.262%	12.86%	17.51%	27.01%
65	2	30.82	10.57%	15.53%	19.88%	33.07%
65	4	14.01	14.45%	25.70%	26.78%	57.68%
65	8	5.759	19.72%	49.43%	36.01%	125.3%
65	16	1.951	26.26%	143.8%	48.21%	553.9%
129	1	84.89	6.829%	8.601%	13.04%	17.68%
129	2	63.14	7.784%	10.17%	14.81%	21.08%
129	4	29.89	10.71%	15.82%	20.12%	33.75%
129	8	13.52	14.65%	26.33%	27.12%	59.28%
129	16	5.612	19.88%	50.49%	36.30%	128.6%
129	32	1.875	26.47%	151.1%	48.65%	598.4%
257	1	170.2	4.979%	5.859%	9.589%	11.88%
257	2	127.8	5.678%	6.852%	10.90%	13.96%
257	4	61.69	7.864%	10.31%	14.95%	21.38%
257	8	29.11	10.82%	16.08%	20.33%	34.34%
257	16	13.32	14.73%	26.60%	27.27%	59.96%
257	32	5.545	19.96%	50.99%	36.43%	130.2%
257	64	1.836	26.57%	155.2%	48.89%	623.9%
513	1	340.9	3.600%	4.038%	6.976%	8.116%
513	2	257.1	4.110%	4.692%	7.947%	9.460%
513	4	125.3	5.729%	6.927%	11.00%	14.12%
513	8	60.31	7.942%	10.45%	15.10%	21.68%
513	16	28.82	10.87%	16.18%	20.41%	34.57%
513	32	13.24	14.76%	26.70%	27.33%	60.24%
513	64	5.513	19.99%	51.23%	36.49%	131.0%
513	128	1.817	26.63%	157.2%	49.00%	637.2%
1025	1	682.2	2.586%	2.805%	5.036%	5.604%
1025	2	515.7	2.957%	3.247%	5.748%	6.500%
1025	4	252.5	4.145%	4.737%	8.013%	9.554%
1025	8	122.7	5.784%	7.007%	11.10%	14.29%
1025	16	59.85	7.968%	10.49%	15.15%	21.78%
1025	32	28.72	10.88%	16.22%	20.44%	34.65%
1025	64	13.21	14.77%	26.75%	27.35%	60.35%
1025	128	5.498	20.01%	51.35%	36.52%	131.4%
1025	256	1.807	26.65%	158.4%	49.06%	644.3%

Table 7
Flicker FM Confidence Factors for
MDEV and TDEV
Confidence factors times sample deviations
give magnitudes of
lower and upper confidence intervals
for 68% and 95% probabilities

N	m	edf	lower 68%	upper 68%	lower 95%	upper 95%
5	1	2.606	24.64%	103.1%	44.95%	336.3%
9	1	5.843	19.63%	48.86%	35.85%	123.5%
9	2	2.615	24.62%	102.7%	44.91%	334.5%
17	1	12.33	15.16%	28.04%	28.02%	63.69%
17	2	6.656	18.82%	44.04%	34.43%	108.6%
17	4	2.086	25.90%	132.6%	47.46%	488.9%
33	1	25.29	11.46%	17.54%	21.48%	37.74%
33	2	14.85	14.14%	24.72%	26.22%	55.19%
33	4	5.883	19.58%	48.59%	35.78%	122.6%
33	8	1.814	26.63%	157.6%	49.02%	639.3%
65	1	51.23	8.525%	11.48%	16.16%	23.95%
65	2	31.27	10.51%	15.39%	19.76%	32.76%
65	4	13.61	14.61%	26.21%	27.06%	58.97%
65	8	5.563	19.94%	50.85%	36.39%	129.8%
65	16	1.683	27.00%	173.8%	49.85%	747.6%
129	1	103.1	6.259%	7.716%	11.98%	15.79%
129	2	64.12	7.732%	10.08%	14.71%	20.89%
129	4	29.10	10.82%	16.09%	20.33%	34.35%
129	8	13.14	14.80%	26.84%	27.40%	60.60%
129	16	5.426	20.09%	51.91%	36.67%	133.2%
129	32	1.619	27.18%	183.2%	50.28%	813.7%
257	1	206.9	4.549%	5.272%	8.777%	10.66%
257	2	129.8	5.638%	6.793%	10.83%	13.84%
257	4	60.11	7.953%	10.47%	15.12%	21.72%
257	8	28.37	10.94%	16.34%	20.54%	34.94%
257	16	12.97	14.88%	27.08%	27.53%	61.21%
257	32	5.363	20.17%	52.41%	36.80%	134.8%
257	64	1.587	27.27%	188.3%	50.50%	850.9%
513	1	414.4	3.282%	3.642%	6.369%	7.306%
513	2	261.3	4.079%	4.651%	7.888%	9.377%
513	4	122.1	5.797%	7.026%	11.12%	14.33%
513	8	58.83	8.028%	10.60%	15.26%	22.00%
513	16	28.13	10.98%	16.42%	20.60%	35.14%
513	32	12.90	14.91%	27.18%	27.58%	61.47%
513	64	5.333	20.20%	52.66%	36.86%	135.6%
513	128	1.571	27.32%	190.9%	50.61%	870.7%
1025	1	829.4	2.354%	2.534%	4.589%	5.056%
1025	2	524.1	2.935%	3.219%	5.704%	6.445%
1025	4	246.2	4.195%	4.802%	8.107%	9.687%
1025	8	119.8	5.848%	7.100%	11.22%	14.49%
1025	16	58.46	8.050%	10.63%	15.30%	22.09%
1025	32	28.04	10.99%	16.46%	20.63%	35.21%
1025	64	12.87	14.92%	27.23%	27.60%	61.58%
1025	128	5.318	20.22%	52.78%	36.89%	136.0%
1025	256	1.564	27.34%	192.1%	50.66%	879.7%

Table 8

Random Walk FM Confidence Factors for MDEV and TDEV
Confidence factors times sample deviations give
magnitudes of lower and upper confidence intervals
for 68% and 95% probabilities

N	m	edf	lower 68%	upper 68%	lower 95%	upper 95%
5	1	3	23.80%	89.04%	43.35%	272.9%
9	1	7	18.50%	42.35%	33.88%	103.5%
9	2	2.067	25.95%	134.0%	47.57%	497.2%
17	1	15	14.08%	24.55%	26.13%	54.77%
17	2	5.492	20.02%	51.39%	36.54%	131.5%
17	4	1.603	27.23%	185.7%	50.39%	832.0%
33	1	31	10.55%	15.48%	19.83%	32.95%
33	2	12.41	15.12%	27.91%	27.96%	63.37%
33	4	4.677	21.03%	58.85%	38.33%	156.3%
33	8	1.429	27.73%	218.6%	51.64%	1090.%
65	1	63	7.791%	10.19%	14.82%	21.11%
65	2	26.26	11.29%	17.14%	21.17%	36.80%
65	4	10.98	15.82%	30.40%	29.18%	69.89%
65	8	4.400	21.41%	62.07%	39.01%	167.5%
65	16	1.353	27.94%	237.3%	52.22%	1252.%
129	1	127	5.694%	6.875%	10.93%	14.01%
129	2	53.96	8.336%	11.14%	15.82%	23.19%
129	4	23.61	11.79%	18.31%	22.06%	39.56%
129	8	10.57	16.04%	31.22%	29.56%	72.10%
129	16	4.285	21.58%	63.54%	39.31%	172.7%
129	32	1.318	28.04%	247.0%	52.50%	1341.%
257	1	255	4.126%	4.713%	7.977%	9.503%
257	2	109.4	6.093%	7.466%	11.68%	15.26%
257	4	48.89	8.699%	11.80%	16.48%	24.65%
257	8	22.93	11.93%	18.65%	22.30%	40.36%
257	16	10.42	16.12%	31.54%	29.71%	72.96%
257	32	4.233	21.66%	64.24%	39.45%	175.2%
257	64	1.301	28.09%	252.0%	52.64%	1388.%
513	1	511	2.970%	3.262%	5.773%	6.532%
513	2	220.2	4.419%	5.098%	8.531%	10.30%
513	4	99.44	6.361%	7.872%	12.18%	16.13%
513	8	47.67	8.794%	11.97%	16.66%	25.04%
513	16	22.72	11.97%	18.76%	22.38%	40.62%
513	32	10.36	16.15%	31.67%	29.77%	73.31%
513	64	4.208	21.70%	64.58%	39.51%	176.4%
513	128	1.292	28.11%	254.8%	52.72%	1415.%
1025	1	1023	2.128%	2.273%	4.151%	4.530%
1025	2	441.8	3.183%	3.521%	6.181%	7.060%
1025	4	200.5	4.616%	5.362%	8.903%	10.85%
1025	8	97.16	6.428%	7.975%	12.30%	16.34%
1025	16	47.34	8.820%	12.02%	16.70%	25.14%
1025	32	22.65	11.98%	18.80%	22.41%	40.70%
1025	64	10.33	16.17%	31.74%	29.80%	73.48%
1025	128	4.196	21.71%	64.74%	39.55%	177.0%
1025	256	1.288	28.12%	256.0%	52.75%	1427.%

Systematic Errors in Cesium Beam Frequency Standards Introduced by Digital Control of the Microwave Excitation

W. D. Lee, J. H. Shirley, F L. Walls, and R. E. Drullinger

National Institute of Standards and Technology
Time and Frequency Division
Boulder, Colorado, USA

Abstract— **As part of the accuracy evaluation of the NIST-7 primary frequency standard we have investigated potential sources of frequency offsets due to the electronic servo system. We will present results from investigations of three major sources of frequency bias in slow square wave FM servo systems: RF spectral purity, synthesizer switching transients, and phase-locked-loop settling time.**

Introduction

The Ramsey technique of separated oscillating fields is widely used in the field of atomic frequency standards. We have investigated potential sources of frequency shifts associated with the microwave radiation interrogating the atoms when using the Ramsey technique. We first measured the frequency bias that results from an interrogation signal whose spectrum contains imbalanced sidebands. We also investigated the frequency offsets that arise in determining the center of the Ramsey fringe using slow square wave digital servos. These servos include a blanking interval to allow the atomic beam tube to reach steady state during the frequency modulation cycle. In addition to atomic beam transit time, there are additional time delays that should be included in the servo's blanking interval. These delays are caused by anomolous switching transients in direct digital frequency synthesizers, and the settling time of phase-locked loops in the microwave synthesis chain.

RF Sideband Pulling

The presence of spurious, discrete sidebands in the spectrum of the microwave radiation can lead to errors in the determination of the center of the Ramsey peak. While these effects are not unique to slow square-wave digital frequency servos, they may be more pronounced in systems employing direct digital synthesis (DDS) due to phase accumulator truncation [1], and electromagnetic interference from high speed digital electronics.

The frequency shift due to a discrete sideband has been calculated theoretically [2]. An independent calculation by one of us (J.S.) has confirmed their result, except that in their equation (9) the drift time T should be replaced by $T+\tau$, where τ is the transit time through one interaction region. For sideband detunings of the order of the Ramsey linewidth, we obtain shifts like those shown by the solid line in figure 2 when we average over the velocity distribution. The curve is antisymmetric about zero detuning; hence symmetric sidebands produce shifts that cancel each other.

Contribution of the U. S. Government, not subject to copyright.

The calculations were verified experimentally using the United States primary frequency standard NIST-7 by intentionally modulating the microwave signal in both amplitude and phase to produce a well defined spectrum. Pure amplitude modulation (AM) generates symmetric sidebands about the carrier and does not lead to a frequency shift. Similarly, neither does pure phase modulation (PM). The amplitude and phase of the carrier must be modulated coherently to yield imbalanced sidebands. A carrier wave that is modulated in both amplitude and phase can be represented by

$$E(t) = \mathrm{Re}\left\{Ae^{i\theta}[1+\epsilon\cos(\Omega t+\phi)]e^{i\eta\sin(\Omega t+\psi)}e^{i\omega t}\right\}, \tag{1}$$

where A is the carrier wave amplitude, ω is the carrier frequency, θ is the carrier phase, ϵ is the amplitude modulation index ($\epsilon \leq 1$), Ω is both the amplitude and phase modulation frequency, ϕ is the phase of the amplitude modulation, η is the phase modulation index, and ψ is the phase of the phase modulation. If $\epsilon \ll 1$ and $\eta \ll 1$ this can be rewritten in three terms representing the carrier and the upper sideband and lower sideband, respectively:

$$\begin{aligned} E(t) &= \mathrm{Re}\{Ae^{i\theta}e^{i\omega t} \\ &+ \frac{1}{2}Ae^{i\theta}(\epsilon e^{i\phi}+\eta e^{i\psi})e^{i(\omega+\Omega)t} \\ &+ \frac{1}{2}Ae^{i\theta}(\epsilon e^{-i\phi}-\eta e^{-i\psi})e^{i(\omega-\Omega)t}\}. \end{aligned} \tag{2}$$

The upper and lower sideband relative amplitudes can then be written

$$S_+ = \frac{1}{2}(\epsilon e^{i\phi}+\eta e^{i\psi}), \tag{3}$$

$$S_- = \frac{1}{2}(\epsilon e^{-i\phi}-\eta e^{-i\psi}). \tag{4}$$

If we choose $\epsilon = \eta$, then the relative sideband intensities are

$$|S_+|^2 = \frac{1}{2}\epsilon^2[1+\cos(\phi-\psi)], \tag{5}$$

$$|S_-|^2 = \frac{1}{2}\epsilon^2[1-\cos(\phi-\psi)]. \tag{6}$$

The upper and lower sidebands have equal intensity for

$$(\phi-\psi) = \frac{\pi}{2} \pm n\pi, \text{ for } n = 0, \pm1, \pm2, \ldots. \tag{7}$$

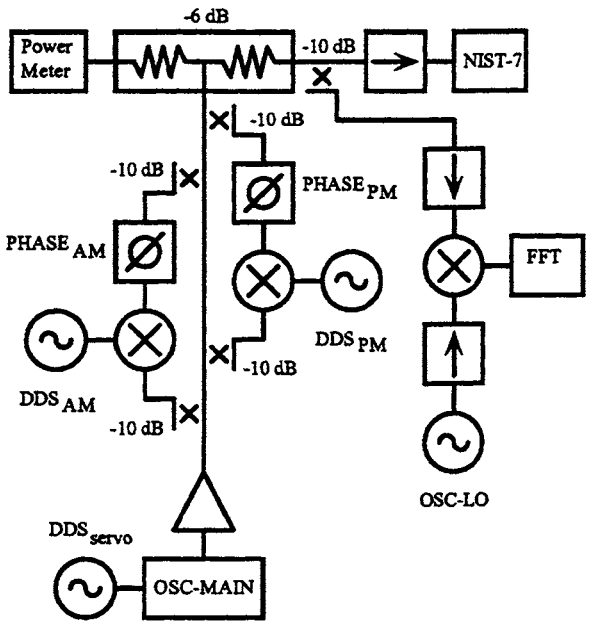


Figure 1: Single Sideband Modulator.

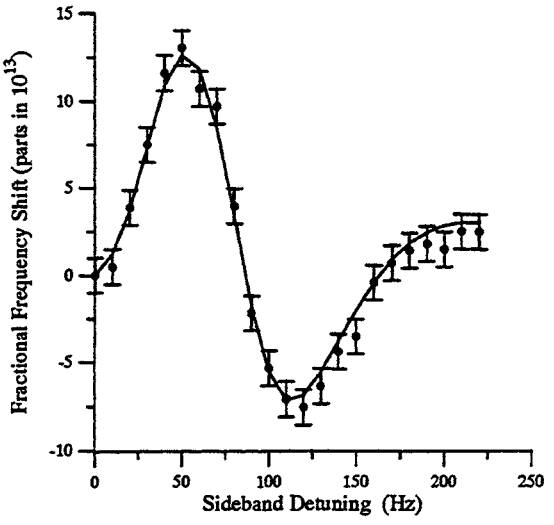


Figure 2: Frequency Shift Due to Sideband Pulling. Solid line is the theoretical calculation. Dots are the experimental measurements. Carrier power: -2.5 dB relative to optimum power; lower sideband: -29 dBc; upper sideband: -66 dBc. Experiment duration: 300 seconds per point.

To measure the frequency pulling with high accuracy and within a short time, we chose to introduce the maximum sideband asymmetry (a single sideband). This is achieved by setting

$$(\phi - \psi) = 0 \pm n\pi, \text{ for } n = 0, \pm 1, \pm 2, \ldots. \quad (8)$$

Such a single sideband modulator is realized with the circuit of figure 1. A low phase noise, 9 GHz microwave source (OSC-MAIN) is modulated in both amplitude and phase by two independent modulators of the type described in [4]. One modulator is optimized to produce high purity AM (by tuning PHASE_{AM}), while the other is optimized to generate only PM (by tuning PHASE_{PM}). Two phase-locked direct digital synthesizers (DDS_{AM} and DDS_{PM}) provide the audio frequency modulation signals ($\cos(\Omega t+\phi)$ and $\sin(\Omega t+\psi)$) to two mixers. The synthesizers are adjusted (under computer control) so that $\epsilon = \eta$ and $\phi = \psi + \pi$. This is verified by monitoring the microwave spectrum with a spectrum analyzer (FFT) at an intermediate frequency of 50 kHz, using a local oscillator (OSC-LO). The amplitudes of the carrier, lower, and upper sidebands are then recorded by the computer.

The frequency bias introduced by this single sideband is determined by measuring the center of the Ramsey fringe using a slow square-wave frequency servo (computer controlled by synthesizer $\text{DDS}_{\text{SERVO}}$). Figure 2 shows the measured frequency pulling (dots) as a function of the sideband detuning Ω. The error bars represent $\pm 1\sigma$ statistical uncertainty. Since the experimental velocity distribution [3], sideband intensities, and detunings were used in calculating the theoretical curve, no free parameters were adjusted to obtain the agreement shown in figure 2. We have obtained similar results for detunings comparable to the width of the Rabi pedestal.

Switching Transients in Slow Square Wave FM Servos

The use of direct digital synthesis in microwave synthesis chains has grown in recent years. This is due to improvements in the frequency resolution and agility of such synthesizers. Slow square-wave frequency servos are readily implemented when DDS instruments are computer controlled. The task of evaluating the frequency biases in such a servo is greatly simplified if the beam tube and electronics are allowed to reach steady state before the clock signal is measured. Thus, after a frequency step command is sent to the DDS, a blanking interval τ_b is established, during which the clock signal is discarded. After the blanking interval, the clock signal is recorded during the acquisition interval τ_a. It is desireable to have τ_a occupy a large fraction of the modulation period in order to enhance frequency stability. Simply reducing the modulation frequency is an unacceptable approach to increasing the ratio τ_a/τ_b, because it increases the system's sensitivity to noise that exhibits $1/f$ power spectra. The blanking interval can be written as the sum

$$\tau_b = \tau_r + \tau_\phi + \tau_{\text{pll}} + \tau_\rho, \quad (9)$$

where τ_r (10 ms) is the response time of the DDS, that is, the time from the transmission of a frequency step command to the change in frequency at the DDS output. τ_ϕ is phase-settling time of the DDS, τ_{pll} is the settling time of the phase-locked loop (PLL), and τ_ρ (25 ms) is the transit time of the slowest atoms of interest across the Ramsey cavity and to the detection region. The PLL is shown in detail in figure 3. There is

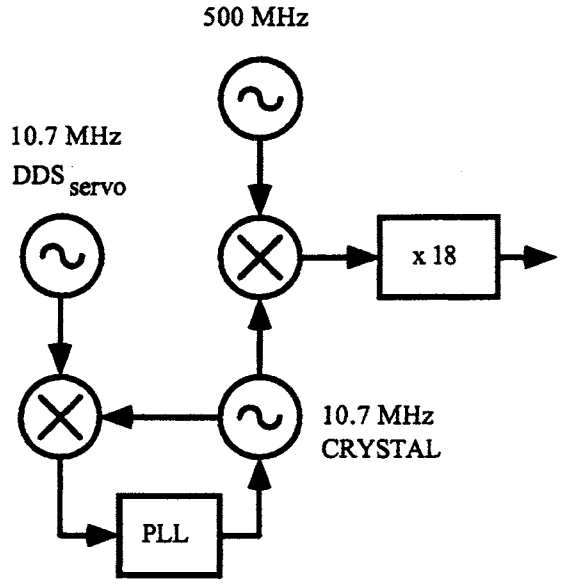


Figure 3: Details of Microwave Synthesis Chain. A 10.7 MHz "clean-up" oscillator is phase locked to the DDS unit in order to reduce the levels of spurious sidebands.

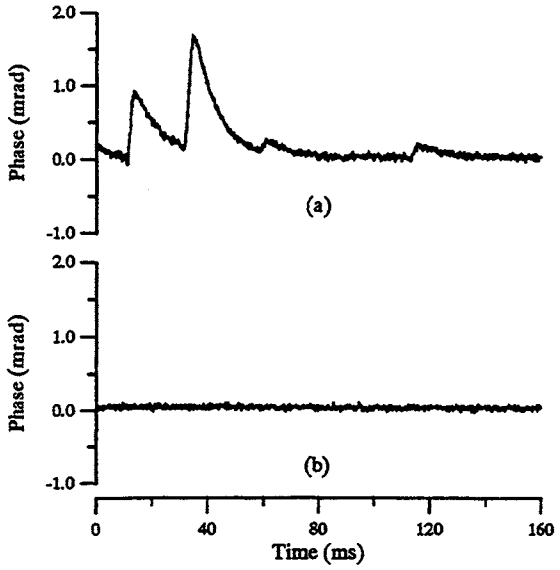


Figure 4: Phase Transients Due to Computer Control. Trace (a): phase transients occuring after frequency control command is sent to synthesizer at $t = 0$ ms. Trace (b): supression of phase transients when phase locked loop is disabled.

a lower limit placed on τ_b by τ_r and τ_ρ. We have, however, discovered techniques whereby τ_ϕ may be eliminated, improving frequency stability and eliminating a potential frequency bias.

Synthesizer Switching Transients

We have discovered a potential frequency bias due to phase transients in a commercial direct digital synthesizer. The transients occur whenever the synthesizer is sent frequency control instructions by the computer. These transients are therefore synchronous with the frequency modulation, some occurring tens, and even hundreds of milliseconds after the frequency step is performed. Since the position of the Ramsey fringe is sensitive to the phase difference between the two ends of the cavity as seen by atoms taking several milliseconds to traverse the cavity, such phase transients can give rise to significant frequency offsets. The transients were measured by synchronously recording the relative phase of two synthesizers while one was sent frequency control commands by a computer. To simplify the measurement, the same frequency control command was sent to the synthesizer repeatedly. Figure 4(a) shows the resulting phase transients. The peak phase deviation of 1.7 mrad occurs 35 ms after the instruction is sent to the synthesizer.

When a DDS unit is used in a microwave synthesis chain, it is common to provide the device with an external frequency reference. The DDS then phase-locks its own internal crystal oscillator to the external signal. We think that the "housekeeping" operations of the microprocessor within the DDS are inadvertently coupled to this phase-locked loop, generating phase transients that are synchronous with its operation. We were able

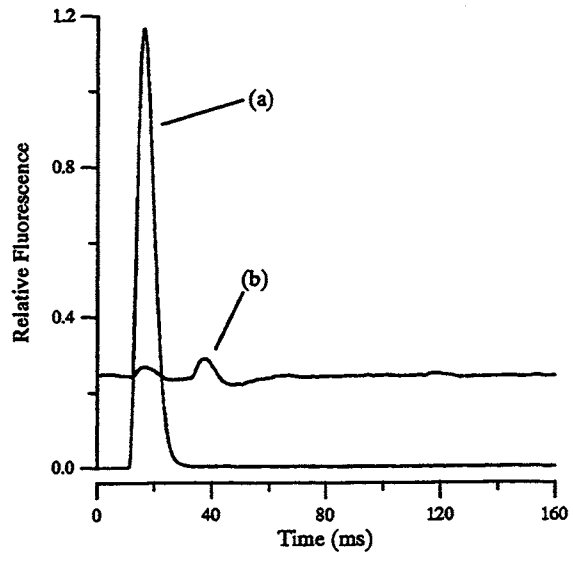


Figure 5: Effect of Synthesizer Phase Transients. Trace (a): Normal slow square wave FM transient due to the atomic transit time distribution. This trace was taken with the frequency servo enabled, thus the steady state fluorescence levels before and after the transient are equal. Trace (b): Effect of DDS phase transients (figure 4(a)) on atomic beam fluorescence. Trace (b) has been offset for clarity.

to eliminate the phase transients by removing the instrument's internal crystal oscillator and instead injecting a suitable reference frequency derived from an external source. Figure 4(b) shows the dramatic improvement in phase stability.

Figure 5 shows the effect of the phase transient upon the atomic beam fluorescence as well as the normal modulation transient. If we were to determine a blanking interval based solely upon the response time of the synthesizer and the atomic beam transit time distribution, the effect of the synthesizer switching transient would still introduce a significant contribution to the measured fluorescence. The size of the frequency bias due to the switching transient was determined by measuring the center of the Ramsey fringe using two different direct digital synthesizers. Both units were identical except that one unit was modified to eliminate the phase transient. With a blanking interval of 40 ms, the fractional frequency bias produced by the transient was $(50 \pm 1) \cdot 10^{-13}$. Although the blanking interval might be extended to avoid the initial switching transients, periodic transients (visible in figure 4(a) at 120 ms) are virtually unavoidable. We emphasize that these phase transients preserve their sign regardless of the sign of a frequency step performed by the DDS. Thus, the resulting bias cannot be eliminated through complex or pseudorandom modulation sequences. By removing these transients, we were able to transfer at least 20 ms from the blanking interval to the acquisition interval in each half of the modulation cycle.

Phase Locked Loop Settling Time

We have also investigated the contribution to the blanking interval of τ_{pll}, the phase-locked loop's settling time.[5] The microwave synthesis chain used in our experiments contains only one PLL that is affected by the slow square-wave frequency modulation. This PLL (figure 3) controls a 10.7 MHz crystal that serves as a "clean-up" oscillator following the DDS. In selecting the loop gain, there is a compromise between the desire for fast settling time (high gain) and good sideband supression (low gain). The settling time of the PLL is show in figure 6. The time constant for the exponential fit in figure 6 is 0.5 ms. Hence τ_{pll} is small compared to τ_r and τ_ρ in equation (9). Nevertheless, we have used the value $\tau_{\text{pll}} = 5$ ms in arriving at our total blanking time of 40 ms.

Conclusion

We have investigated three sources of bias in digitally controlled microwave synthesis chains: spurious RF sidebands, synthesizer switching transients, and the phase-locked loop's settling time. The experimental results obtained for RF sideband pulling are in excellent agreement with theory. They apply not only to digital servo systems but also to traditional analog servos. While the synthesizer switching transients we observed were unique to a specific model and might be avoided by selecting a different DDS, our modifications yielded a DDS with phase stability superior

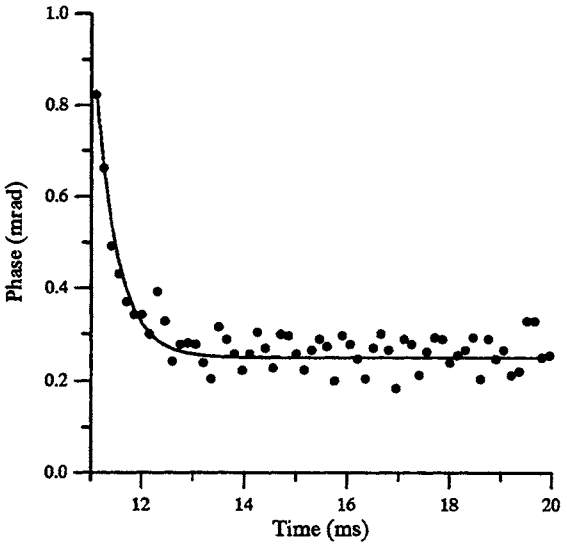


Figure 6: PLL Error Signal During Frequency Step. Dots: Measured phase error after the DDS was instructed to perform a 54 Hz step (the peak to peak modulation amplitude). Solid Line: Exponential fit to the data.

to other models we have tested. Our analysis of the phase-locked loop transient in the microwave-synthesis chain shows that it settles rapidly compared to the other delay times.

References

[1] Nicholas, H. T., Samueli, H., "An Analysis of the Output Spectrum of Direct Digital Frequency Synthesizers in the Presence of Phase-Accumulator Truncation," 41st Annual Frequency Control Symposium, 1987, pp. 495-502.

[2] Audoin, C., Jardino, M., Cutler, L. S., and Lacey, R. F., "Frequency Offset Due to Spectral Impurities in Cesium-Beam Frequency Standards," IEEE Trans. Inst. Meas. vol. IM-27, 1978, pp. 325-329.

[3] Lee, W. D., Shirley, J. H., and Drullinger, R. E., "Velocity Distributions of Atomic Beams by Gated Optical Pumping," Proceedings of the 1994 IEEE International Frequency Control Symposium, pp. 658-661.

[4] Walls, F. L., "Extending the Range and Accuracy of Phase Noise Measurements," Proceedings of the 42nd Annual Frequency Control Symposium, 1988, pp. 432-441. U. S. Patent 4,968,908, "Method and apparatus for wide band phase modulation." (1990).

[5] Schroder, R., "Frequency Synthesis in Primary Cesium Clocks," Proc. 5th EFTF, Besancon (1991).

The Effect of Harmonic Distortion on Phase Errors in Frequency Distribution and Synthesis

F.L. Walls* and F. G. Ascarrunz+
*Time and Frequency Division, NIST, +SpectraDynamics Inc. Boulder CO, USA

Abstract

This paper explores the effect of harmonic distortion on phase errors in frequency distribution and synthesis. Harmonic distortion gives rise to large timing errors in phase or zero crossing detectors. These large timing errors may cause unacceptable sensitivity to environmental conditions or signal fluctuations. Two types of phase detectors are considered in this paper. The timing errors in an ideal zero crossing detector is calculated as a function of harmonic number and the phase of the harmonic relative to the fundamental. The result is $\delta t_{max} = \sum_{n=2}^{\infty} v_n \Big/ v_1 \omega_1$, where v_1 is the amplitude of the carrier frequency ω_1, and v_n is the amplitude of the nth harmonic. The timing errors of double balanced mixers are in principle insensitive to even harmonics of the RF signal, and their sensitivity to higher odd harmonics fall roughly as 1/n. Experimental data from several mixers at 5 MHz and different drive levels show that this model is a first-order approximation to real devices. The low sensitivity to the second harmonic is an important advantage over true zero-crossing detectors because it is difficult to significantly reduce the second harmonic without introducing large phase shifts due to filtering networks. Balanced mixers offers approximately 20 dB improvement over the true zero-crossing detector when used in the linear range of operation. Harmonic distortion is an important parameter in the design of systems with timing stability less than about 100 ps.

Introduction

With the potential of new frequency standards with short-term frequency stability less than $1\times10^{-14}\,\tau^{-1/2}$, it has been necessary to re-examine the origin of timing errors in measurement systems as well as frequency distribution and synthesis building blocks. It is desirable to have these subsystems maintain timing errors of less than approximately 1 ps. General criteria for frequency metrology to support these standards are discussed in [1], where the requirements on voltage-standing-wave-ratio (VSWR), the temperature coefficients of the cable delay, and errors in phase detectors for the design of ultra stable synthesis and distribution systems, are discussed. References [2,3], demonstrate that distribution amplifiers which satisfy the requirements set forth in [1] can be made at 5 to 10 MHz and at 100 MHz.

This paper explores the one remaining problem that has not been considered: the effect of harmonic distortion on phase errors in frequency synthesis and distribution systems. Harmonic distortion can cause large timing errors, which may give rise to unacceptable sensitivity to environmental conditions or signal fluctuations. The harmonic distortion may be present in the source or arise within the distribution amplifiers. Harmonic distortion is converted to timing errors by the phase detector. It is not the timing error itself which causes problems but the changes in this error due to signal amplitude fluctuations or phase fluctuations. These amplitude and phase fluctuations are generally due to environmental effects but could have other systematic origins.

Two types of phase detectors are considered in this paper. The first is a true zero-crossing detector and the second is a balanced mixer. To simplify the discussion, the local oscillator signal is assumed to be

Work of U.S. Government. Not Subject to U.S. copyright.

an ideal sine wave and all the harmonic distortion is in the RF input signal. Clearly if the local oscillator signal also has harmonic content there will be additional terms in the timing error which can be treated as discussed below.

True zero-crossing detector

In many timing and frequency synthesis systems, a digital zero-crossing based phase detector is used. These devices generally trigger some type of counter which essentially counts the time between zero crossings. Harmonic distortion causes the zero crossing to be a function of the phase and amplitude of the harmonics as well as the fundamental. Fluctuations in the time of the zero crossing therefore compromise the accuracy of the system. The analysis to determine the timing error for an ideal zero crossing detector is straightforward. The input signal V can be represented as a summation of the carrier with frequency ω_1 and amplitude v_1 with harmonics $n\omega$ of amplitude v_n and phase ϕ_n relative to the fundamental:

$$V = v_1 \cos(\omega_1 t) + \sum_{n=2}^{\infty} v_n \cos(n\omega_1 t + \phi_n). \quad (1)$$

Since we are interested in what happens at the zero crossings when the phase of the fundamental is a multiple of $\pi/2$, we linearize about that point with a Taylor series expansion:

$$V = v_1 \cos(\omega_1 t_0) + \sum_{n=2}^{\infty} v_n \cos(n\omega_1 t_0 + \phi_n) - \left[v_1\omega_1 \sin(\omega_1 t_0) + \sum_{n=2}^{\infty} v_n n\omega_1 \sin(n\omega_1 t_0 + \phi_n) \right](t - t_0) \quad (2)$$

At the zero crossing $V = 0$ and $\omega_1 t_0 = \pi/2$. After simplification the time error Δt, due to the harmonic content of the signal can be expressed as

$$\Delta t = (t - t_0) = \frac{\sum_{n=2}^{\infty} v_n \cos\left(\phi_n + n\frac{\pi}{2}\right)}{v_1\omega_1 + \sum_{n=2}^{\infty} v_n n\omega_1 \sin\left(\phi_n + n\frac{\pi}{2}\right)} \quad (3)$$

As Eq. (3) shows, the timing error depends on the amplitude and phase of the harmonics. The worst case timing error is

$$\Delta t_{max} = \frac{\sum_{n=2}^{\infty} v_n}{v_1\omega_1}, \quad (4)$$

which occurs when the phase of the nth harmonic is an integral multiple of $\pi/2$ relative to the phase of the fundamental. In frequency distribution applications, the amplitude and phase of the harmonics may exhibit fluctuations due to environmental effects, changes or mismatch in the load impedance or dispersion along a transmission line. These timing errors are also important in frequency synthesis or phase-locking applications since they are converted to frequency errors by the phase-locked loops. For a signal with a carrier frequency of 5 MHz and a second harmonic of -25 dBc the timing error could be as much as 1.8 ns depending on the phase of the harmonic relative to the fundamental. A 10% change in the amplitude of the harmonic; that is from -25 dBc to -25.5 dBc would cause a 180 ps error. These errors could be greater if other harmonics are present as well. At a carrier frequency of 100 MHz the timing error due to a second harmonic of -25 dBc is 9 ps.

Double Balanced Mixers

Balanced mixers used as phase detectors are not so easily analyzed so we have experimentally measured the phase errors due to harmonic distortion up to the ninth harmonic. A 5 MHz source is split the first branch is filtered to suppress the harmonic content to well below -60 dBc.

This signal is applied to the LO port of the mixer. The harmonic frequency to be tested is generated at -30 dBc and offset by 3 kHz from $n\omega_1$. This signal is then summed to the second branch as shown in Figure 1. Variable attenuators are used to adjust RF and LO power at the mixer ports for the various phase error measurements. The low harmonic power assures linear conversions within the phase detector. The 3 kHz IF signal is low pass filtered and analyzed on an FFT analyzer. This approach assures that measurements yield the worst case timing error.

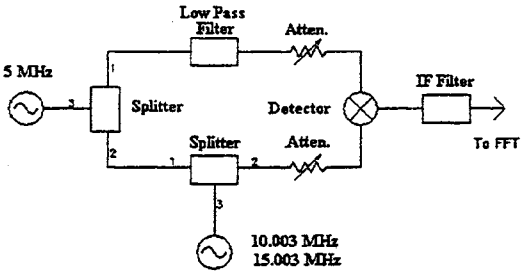


Figure 1. The setup for measuring the timing error due to distortion in double balanced mixer phase detectors.

The timing errors due to harmonic distortion in balanced mixers are different than for the zero crossing detector. In principle balanced mixers are insensitive to even harmonics of the reference oscillator and the sensitivity to the odd harmonics fall roughly as 1/n. This trend is apparent in Figs. 2, 3, 4 and 5, where there is suppression of the timing error due to the even harmonics. However leakage from port to port causes even and odd harmonics to mix down to DC causing at the IF port, additional DC components which are interpreted as phase shifts or timing errors. These timing errors are about 15-25 dB smaller than for the zero-crossing detector.

Four different low frequency (.1-500 MHz) mixers commonly used in our synthesizers and measurement equipment, were tested at LO power ranging from 4 to 13 dBm and RF power of 1 to 10 dBm. The timing errors normalized to 0 dBc are displayed in Figs. 2-5 for four different combinations of LO and RF power at the mixer ports. For a specific harmonic power the timing error is obtained by multiplying the time error read from the chart by v_n/v_1 where v_n is the amplitude of the harmonic and v_1 is the amplitude of the carrier. The data for LO ranges of 4 to 13 dBm and RF ranges of 1 to 10 dBm are displayed in tabular form in Tables 1-4.

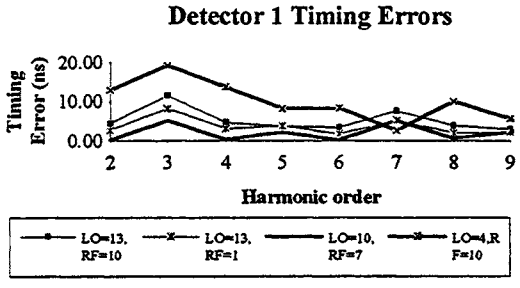


Figure 2. The timing error for a single balanced mixer phase detector w/ nominal LO of +10 dBm normalized to a harmonic level of 0 dBc for a carrier frequency of 5 MHz. For a specific harmonic level the timing error is obtained by multiplying the time error read from the chart by v_n/v_1.

Figure 2 displays the timing errors for Detector #1, a single balanced mixer optimized for use as a phase detector (it has very low conversion loss) for a carrier frequency of 5 MHz. The timing error for the even harmonics is generally much lower than for the odd harmonics. Even in the worst case, the timing error at 5 MHz for this phase detector is lower than 31.8 ns. Recall that for a true zero-crossing detector the normalized error for each harmonic is 31.8 ns for a carrier frequency of 5 MHz. For a 100 MHz carrier the timing error of a true zero-crossing detector sensitivity is 1.6 ns. The lowest

sensitivity to harmonic distortion occurs when the power at the LO port is +10 dBm and the power at the RF port is +7 dBm. Detector #2 is a general purpose double balanced mixer.

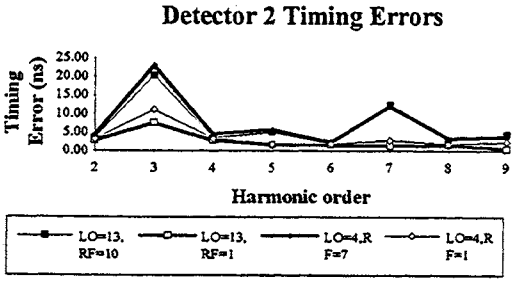


Figure 3. The timing error for a double balanced mixer w/ nominal LO power of +7 dBm normalized to a harmonic level of 0 dBc for a carrier frequency of 5 MHz. For a specific harmonic level the timing error is obtained by multiplying the time error read from the chart by v_n/v_1.

Figure 3 shows that the best performance is attained when the mixer is driven at +13 dBm on the LO port and +1 dBm on the RF port. This power level configuration corresponds to the mixer being driven in its linear operating range; the RF power level is well below the LO power level. High power levels at both the RF and LO ports is the configuration commonly used when a double balanced mixer is used as a phase detector. In this configuration the mixer is operating in saturated mode. Operating in saturated mode may provide the lowest conversion loss but the sensitivity to harmonic distortion increases by an order of magnitude. For detector 3 the sensitivity to the third harmonic can be reduced by having a high LO power and a low RF power at the expense of a slightly higher sensitivity to the other harmonics. In general detector 3 and detector 4 have trends similar to the previous two mixers.

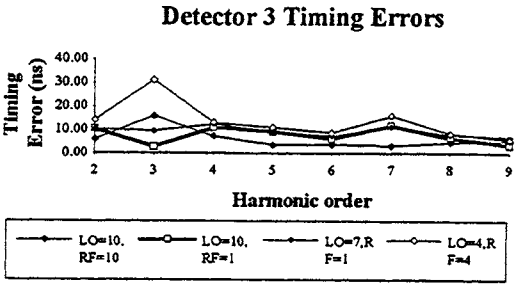


Figure 4. The timing error for a double balanced mixer w/ nominal LO power of +15 dBm normalized to a harmonic level of 0 dBc for a carrier frequency of 5 MHz. For a specific harmonic level the timing error is obtained by multiplying the time error read from the chart by v_n/v_1.

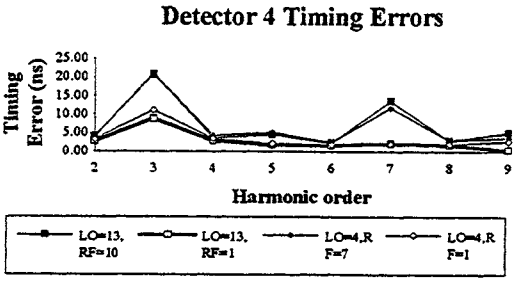


Figure 5. The timing error for a double balanced mixer w/ nominal LO of +7dBm, normalized to a harmonic level of 0 dBc for a carrier frequency of 5 MHz. For a specific harmonic level the timing error is obtained by multiplying the time error read from the chart by v_n/v_1.

Discussion

Harmonic distortion changes the zero crossing of a signal by an amount which depends on the amplitude and phase of each harmonic present in the signal. The timing error which results becomes problematic with signal and phase fluctuations due to environmental effects such as temperature and humidity, aging in crystal oscillators, aging in the bias components of active devices or changes in load impedance and transmission line characteristics. A balanced mixer phase detector offers a significant improvement

over a true zero-crossing detector. The timing error due to the even harmonics present in the signal are suppressed and if the mixer is operated linearly, the sensitivity to the 3rd harmonic decreases 10 to 20 dB. Increasing the carrier frequency from 5 to 100 MHz should help to reduce the effects by about a factor of 20.

If a timing stability of 1 ps is required and one assumes that the harmonic content is stable to approximately 3%, then the harmonic content of the signal should not exceed -60 dBc when a true zero-crossing detector is used to determine time or phase. If a balanced mixer is used in place of the true zero-crossing detector, these requirements can be reduced to a harmonic content of approximately -45 dBc at 5 MHz or -20 dBc at 100 MHz. The distribution amplifiers described in [3,4] easily satisfy these requirements.

References

[1] F.L. Walls, L.M. Nelson, and G.R. Valdez, "Designing for frequency metrology at the 10-18 level", Proc. 6th EFTF, pp 477-481, 1992.

[2] A. DeMarchi, F. Mussion, and M. Siccardi, "A high isolation low noise amplifier with near unity gain up to 100 MHz," Proc. 1993 IEEE Symp. on Freq. Contr., 1993, pp 216-219.

[3] C.W. Nelson, F.L. Walls, M. Siccardi, and A. DeMarchi, "A new 5 and 10 MHz high isolation distribution amplifier," Proc. 1994 IEEE Symp. on Freq. Contr., 1994, pp. 567-571.

[4] M. Siccardi, S. Romisch, F.L. Walls, and A. DeMarchi, "A low noise 100 MHz distribution amplifier for precision metrology," to be published in Proc. 9th Eur. Freq. and Time Forum, 1995.

Table 1.

5 MHz Carrier	Timing Error in (ns) of the nth harmonic							
Power at Detector #1	2nd	3rd	4th	5th	6th	7th	8th	9th
LO=13dBm RF=10dBm	4.55	11.56	4.71	3.78	3.45	7.64	3.92	3.01
LO=13dBm RF=7dBm	4.01	11.56	4.65	2.31	2.62	4.82	2.94	1.31
LO=13dBm RF=4dBm	1.08	13.27	1.18	1.22	0.49	2.90	0.23	2.39
LO=13dBm RF=1dBm	2.77	8.18	3.11	3.87	1.83	5.34	2.08	2.28
LO=10dBm RF=10dBm	0.70	8.28	0.67	2.15	0.91	6.42	0.89	1.36
LO=10dBm RF=7dBm	0.28	5.16	0.34	2.25	0.30	5.16	0.74	2.28
LO=10dBm RF=4dBm	3.11	10.07	3.37	3.49	1.92	4.44	2.36	4.20
LO=10dBm RF=1dBm	7.90	6.50	8.57	7.64	4.93	8.28	5.34	4.20
LO=7dBm RF=10dBm	5.73	10.54	5.60	2.74	3.08	2.56	4.05	0.81
LO=7dBm RF=7dBm	4.01	2.80	4.01	3.65	2.10	5.93	1.99	3.83
LO=7dBm RF=4dBm	6.21	9.84	6.65	5.16	3.83	6.42	4.29	5.34
LO=7dBm RF=1dBm	10.18	8.57	10.91	8.87	6.42	8.47	6.88	4.99
LO=4dBm RF=10dBm	12.97	19.18	13.74	8.18	8.47	2.74	10.07	5.66
LO=4dBm RF=7dBm	15.59	13.89	15.41	9.84	8.67	8.18	9.18	4.44
LO=4dBm RF=4dBm	14.38	13.74	15.59	10.79	8.87	12.67	9.50	8.37
LO=4dBm RF=1dBm	15.77	14.89	17.09	12.53	9.84	11.42	10.54	7.04

Table 2.

5 MHz Carrier	Timing Error in (ns) of the nth harmonic							
Power at Detector #2	2nd	3rd	4th	5th	6th	7th	8th	9th
LO=13dBm RF=10dBm	3.57	20.32	3.49	5.04	1.99	12.38	2.90	4.44
LO=13dBm RF=7dBm	3.26	20.55	3.37	4.20	1.81	6.00	2.36	2.47
LO=13dBm RF=4dBm	2.87	15.59	3.01	2.28	1.67	2.25	1.75	4.55
LO=13dBm RF=1dBm	2.74	7.55	2.77	1.52	1.63	1.34	1.63	0.49
LO=10dBm RF=10dBm	3.70	21.03	3.65	5.34	2.10	14.22	3.11	5.79
LO=10dBm RF=7dBm	3.22	21.52	3.33	4.44	1.79	7.90	2.36	2.84
LO=10dBm RF=4dBm	2.90	17.70	3.08	2.59	1.71	2.47	1.77	5.47
LO=10dBm RF=1dBm	2.84	9.39	2.90	1.54	1.73	2.06	1.75	0.30
LO=7dBm RF=10dBm	4.34	22.80	4.29	6.07	2.47	16.90	3.65	7.29
LO=7dBm RF=7dBm	3.29	21.27	3.45	4.65	1.85	9.39	2.50	3.04
LO=7dBm RF=4dBm	2.97	18.11	3.18	2.90	1.77	2.39	1.85	5.66
LO=7dBm RF=1dBm	2.84	10.42	2.94	1.56	1.75	2.65	1.71	1.28
LO=4dBm RF=10dBm	4.87	21.03	4.76	6.28	2.84	16.90	4.01	7.55
LO=4dBm RF=7dBm	4.29	23.06	4.50	5.73	2.44	12.10	3.29	3.92
LO=4dBm RF=4dBm	3.70	19.18	3.96	3.78	2.20	3.53	2.31	5.93
LO=4dBm RF=1dBm	3.15	11.16	3.29	1.83	1.94	2.90	1.81	2.36

Table 3.

5 MHz Carrier	Timing Error in (ns) of the nth harmonic							
Power at Detector #3	2nd	3rd	4th	5th	6th	7th	8th	9th
LO=13dBm RF=10dBm	4.60	14.38	5.28	1.85	2.62	2.59	3.22	2.62
LO=13dBm RF=7dBm	1.31	21.03	1.73	3.57	0.96	8.28	0.89	6.88
LO=13dBm RF=4dBm	6.88	8.87	7.13	4.20	3.92	7.46	4.29	3.92
LO=13dBm RF=1dBm	22.53	17.29	23.87	14.89	13.42	7.21	14.72	6.07
LO=10dBm RF=10dBm	6.00	15.95	7.38	3.74	3.87	3.18	4.65	5.47
LO=10dBm RF=7dBm	10.30	10.42	11.16	8.97	6.07	11.83	6.21	5.47
LO=10dBm RF=4dBm	17.29	15.95	17.90	12.53	10.07	7.38	11.04	5.60
LO=10dBm RF=1dBm	10.42	2.74	10.91	9.29	6.21	11.83	6.57	3.45
LO=7dBm RF=10dBm	1.47	24.99	2.36	2.20	0.95	5.66	1.03	3.70
LO=7dBm RF=7dBm	5.73	18.74	5.73	6.28	1.49	9.72	3.45	6.81
LO=7dBm RF=4dBm	10.18	19.18	11.42	9.50	6.81	7.04	7.64	8.37
LO=7dBm RF=1dBm	10.30	9.61	12.53	9.39	7.64	11.16	8.09	6.96
LO=4dBm RF=10dBm	8.87	23.06	5.04	2.71	3.01	5.93	3.01	2.18
LO=4dBm RF=7dBm	10.07	35.72	10.79	7.38	6.14	5.79	7.38	12.97
LO=4dBm RF=4dBm	14.22	31.11	13.12	11.16	8.87	15.95	8.47	6.00
LO=4dBm RF=1dBm	19.63	15.59	11.56	10.54	7.72	13.27	8.09	5.93

Table 4.

5 MHz Carrier	Timing Error in (ns) of the nth harmonic							
Power at Detector #4	2nd	3rd	4th	5th	6th	7th	8th	9th
LO=13dBm RF=10dBm	3.96	20.79	3.61	4.65	2.28	13.74	2.94	5.16
LO=13dBm RF=7dBm	2.87	19.63	3.01	3.78	1.69	6.42	2.03	1.56
LO=13dBm RF=4dBm	2.77	16.32	2.97	2.36	1.67	1.75	1.81	4.55
LO=13dBm RF=1dBm	2.68	8.87	2.87	1.63	1.60	2.15	1.63	0.40
LO=10dBm RF=10dBm	3.65	20.55	3.57	4.99	2.15	14.55	2.80	6.21
LO=10dBm RF=7dBm	3.01	20.32	3.18	4.10	1.79	8.09	2.23	2.23
LO=10dBm RF=4dBm	3.01	18.32	3.29	2.84	1.85	1.77	2.10	5.16
LO=10dBm RF=1dBm	2.87	9.95	3.08	1.75	1.71	2.08	1.71	1.42
LO=7dBm RF=10dBm	4.24	21.77	4.15	5.60	2.50	16.51	3.22	7.21
LO=7dBm RF=7dBm	3.33	21.03	3.53	4.60	1.99	9.95	2.53	2.90
LO=7dBm RF=4dBm	2.97	17.49	3.29	3.04	1.83	2.33	1.99	4.93
LO=7dBm RF=1dBm	2.71	10.07	2.94	1.73	1.61	1.94	1.63	2.03
LO=4dBm RF=10dBm	5.04	21.52	4.87	6.07	2.94	17.49	3.78	8.28
LO=4dBm RF=7dBm	4.05	21.03	4.29	5.16	2.41	11.83	3.04	3.83
LO=4dBm RF=4dBm	3.78	18.74	4.10	3.92	2.31	4.05	2.53	4.93
LO=4dBm RF=1dBm	3.26	11.16	3.49	2.13	1.92	2.01	1.99	2.80

On the dependence of spatial phase variations with the RF source in cavities for Cesium fountain standards

A. Khursheed, G. Vecchi, A. De Marchi
Politecnico di Torino, Torino, Italy

Abstract

In this paper, a numerical implementation of an electromagnetic field approach is used to evaluate spatial phase variation in a cavity in the presence of sources. The finite element method is used, in association with an efficient perturbation-like formulation, to calculate the spatial phase distribution for a two dimensional approximation of the fountain cavity operating in the cylindrical TE_{011} mode. A physical interpretation of the phase gradient in the cavity is presented, which shows it to be proportional to power flow.

Introduction

Accuracies of the order of 10^{-15} are expected to be obtainable in the near future in frequency standards based on laser assisted Cesium fountains [1]. This figure is limited by the uncertainty that remains in the theoretical description and in the actual estimation of collisional spin-exchange shift. Improvements to come in theoretical understanding, and in the experimental ability to guarantee accurate and reproducible densities of the atomic clouds launched in the fountain, may later allow even better accuracies in the evaluation of the spin-exchange shift.

It appears therefore important, in this perspective, to be able to guarantee that the evaluation accuracy of all remaining systematic effects be in the low 10^{-16} range, so that the overall accuracy of the standard may be not worse than that of the spin-exchange shift.

A non trivial task appears then to be the assessment of the uncertainty introduced by the end-to-end phase shift. The latter in fact depends on the atomic cloud's density distribution as it flies through the cavity on both its way up and its way down. Because the phase of the relevant rf magnetic field component varies spatially within the microwave cavity, due to losses on the walls, different density distributions in the two passes may result in a different average phase, and produce as a consequence a net end-to-end phase shift. If ϕ_1 and ϕ_2 are the average phases in the first and second pass respectively, the usual phase shift formula reads

$$\frac{\delta\nu}{\nu} = \frac{\phi_1 - \phi_2}{\pi Q} \tag{1}$$

with

$$\phi_i = \frac{1}{\Sigma_i}\int_{\Sigma_i} \rho_i(P)\phi(P)\mathrm{d}\sigma, \tag{2}$$

where Σ_1 and Σ_2 are the total cloud sections in the first and second pass respectively, and the $\rho_i(P)$ are their respective total atomic flux densities, relative to the total number of detected atoms. The atomic flux density can be reasonably assumed to be uniform when the cloud falls down, and yield therefore a value for ϕ_2 which is just the average phase over the cavity aperture. The latter can be computed once the phase distribution is known. Evaluation of ϕ_1 requires knowledge of the total flux density distribution. This is not easy to obtain at the present development stage of MOTs or optical molasses. However, limits to $\phi_1 - \phi_2$ can be stated by worst case reasoning on calculated phase maps, and compared to the maximum uncertainty acceptable for a given accuracy level of end-to-end phase shift.

It can be easily seen from (1) that the condition $|\phi_1 - \phi_2| < 3\mu$rad must be satisfied for a claim of zero end-to-end phase shift down to an accuracy of 10^{-16} in a fountain standard with a line Q of 10^{10}.

The actual phase distribution can only be computed until better control of a point source of slow atoms may allow experimental mapping via frequency shift measurements (which in any case will be very difficult). A first attempt at analysing spatial phase variations of the relevant H component in the cylindrical TE_{011} preselected cavity mode was made in [2] with an analytical approach in the absence of source and external coupling, and later extended with the help of numerical FEM calculations to include external coupling in [3], though still in the absence of sources.

In this paper a 2D calculation of spatial phase variations is implemented, which takes into account external coupling and includes consideration of the sources.

Assumptions and Theory

We consider here a two dimensional cavity structure as in Fig. 1, with cross section in the x-y plane, and consistently delete any dependence on the longitudinal (axial) coordinate z. Wall losses are accounted for by surface impedence boundary conditions [6]. The entire structure is composed of the fountain cavity itself and of the access waveguide that feeds power into the cavity. The presence of a source is accounted for by specifying appropriate boundary conditions for the fields at a reference plane in the waveguide, here called the source plane.

We will consider a purely TE field (with respect to z); the coupling generated by wall losses between TE and TM modes will be neglected. For the fountain cavity, operating in TE_{011} mode, this is a second-order effect [2]. In other cases this assumption might lead to an underestimation of the effect of wall losses; however, the present analysis will still highlight the dominant physical phenomena. With a purely TE field ($E_z = 0$), all fields are derived from $H_z(x, y)$; because of the invariance in the z direction, the magnetic field will be purely longitudinal (e.g. see [8, Chap. 2]). We will be concerned with the spatial phase variation of $H = H_z$, that is, in the quantity $\phi(x, y)$ defined through

$$\tan\phi = \frac{\mathrm{Im}(H)}{\mathrm{Re}(H)} \tag{3}$$

The complex fields are governed by a driven electromagnetic wave equation. A direct numerical solution to this equation suffers from numerical precision problems. In fact, as mentioned in the introduction, we are interested in μrad phase accuracy, and we need therefore accurate evaluation of the field down to 10^{-6}H. A perturbation-like approach is taken, in which the fields in the cavity are broken down into a *primary* part $(\underline{E}_0, H_0)$ that corresponds to the fields in absence of losses, and a (small) remainder $(\underline{e}, h)$, to be called *secondary* part, that is directly generated by the losses

$$\begin{aligned} H(x,y) &= H_0(x,y) + h(x,y), \\ \underline{E}(x,y) &= \underline{E}_0(x,y) + \underline{e}(x,y) \end{aligned} \tag{4}$$

In general, these fields will be complex. To the first order in the small normalized wall resistance R_s/Z_0 ($Z_0 = \sqrt{\mu_0/\epsilon_0}$ being the intrinsic vacuum impedance), the magnitude of h/H_0 will be of the order of R_s/Z_0. Upon insertion of (4) into the driven form of the wave equation for H, one obtains the usual (Helmholtz) equation for both H_0 and h

$$[\nabla^2 + k^2]H_0(x,y) = 0, \tag{5}$$

$$[\nabla^2 + k^2]h(x,y) = 0 \tag{6}$$

where $k = \omega/c$.

The normal derivative of H_0 must be zero at the wall boundary.

As for the secondary fields, the surface impedance conditions relate the (total) tangential electric field to the (total) magnetic (longitudinal) field component. Considering that the primary fields satisfy the lossless problem, the following condition at the metal boundary γ_m is found

$$\left.\frac{\partial h}{\partial n}\right|_{\gamma_m} = (1+j)\frac{R_s}{Z_0}k\,(H_0 + h)|_{\gamma_m} \tag{7}$$

which shows that the primary field acts as an (equivalent) boundary source term on the secondary field.

For the purpose of calculating the spatial phase distribution, the real part of h can be neglected because, up to first order in R_s/Z_0, $\tan\phi$

$= \mathrm{Im}(h)/H_0$. This approximation is valid everywhere, with the exception of the vicinity of points where $H_0 = 0$, which are not of interest here because the atoms pass where H_0 is maximum. In the following analysis, only the calculation of $\mathrm{Im}(h)$ will be considered.

The boundary condition for $\mathrm{Im}(h)$ at the source plane is obtained here on the basis of energy conservation, enforcing the balance between the power flowing from the source into the structure (cavity plus input lossy waveguide) and the power dissipated at the lossy metal walls. The per unit length (PUL) power is expressed as the the real part of the PUL flux of the Poynting vector $\underline{S} = \underline{E} \times \underline{H}^*$. If the source plane is taken far enough from the waveguide-cavity discontinuity, the field at the source plane is very well approximated by that of the dominant waveguide mode, and it can be shown [7] that the power balance condition will provide an explicit value for $\mathrm{Im}(h)$ at an electric plane, or for the normal derivative of $\mathrm{Im}(h)$ at a magnetic plane.

Phase gradient and power flow

Phase variations are associated with losses (in a lossless, closed cavity there is no phase variation, aside from jumps), and losses generate a power flow from the source to the dissipating walls. It is therefore relevant to investigate the relationship between the phase gradient and the real part of the Poynting vector. For the two-dimensional, TE case examined here ($\underline{H} = \hat{z} H_z$), on taking the gradient of (3), and relating the fields as per Maxwell's equations (in terms of transverse and longitudinal components) one finds [7] that

$$\mathrm{Re}(\underline{S}) = -Z_0 |H|^2 \frac{\nabla \phi}{k} \qquad (8)$$

(the above equation holds everywhere, including the turning points). The phase gradient turns out to be proportional to the PUL power flow, a circumstance that allows one to give an estimate of the phase variation in a complex structure by energy considerations.

Results obtained for the fountain cavity in the absence of sources [3] are clear in the light of this physical interpretation: when no power is fed from the outside, it will flow from the center toward the walls, giving a negative phase slope along its path. Results obtained with an external source (see next section) validate this interpretation.

Finite Element Solution

The finite element method was used to numerically solve for the primary and secondary field distributions within the cavity; since the formulation is standard, only the relevant part of the method as it relates to this specfic application will be summarised.

In this work, both the basis functions used to expand the solution, and the weighting functions used to transform the equation into the integral (weak) form are piece-wise linear functions on triangular domains.

The numerical mesh was automatically fitted to curved metal boundaries with the boundary fitted coordinate method [4]. Solution to the eigenvalue form of the wave equation is necessary to know resonance frequency and field distribution of the resonant mode of the cavity. Because the operating frequency will be close to resonance, the eigen-mode solution of the primary field (sourceless) equation is a good approximation to the primary field distribution, and thus a useful starting point for finding the actual primary field distribution (with sources).

For the eigenvalue problem, the different modes of the cavity were found by using an alternating iterative technique reported in [5].

For the secondary field distribution, the system matrix of the FEM (stiffness matrix) was not found to be positive definite, and hence prohibited the use of standard iterative techniques such as successive overrelaxation or the conjugate gradient method. This restriction imposed the use of the of Gaussian Elimination method. For a mesh having 10000 nodes, it took around 3 minutes on a 486 PC running at 66 MHz to solve for the secondary field distribution.

In order to test the formulation and numerical implementation, we first considered a test structure consists of parallel straight lossy metal sidewalls, terminated by a lossy short at one end; at the other end of the structure, a source boundary plane is specified. The the structure can be analytically solved for as in [7], thus allowing a check of the numerics. Analytical and numerical results show good agreement: the maximum difference between them is around 9% of their peak

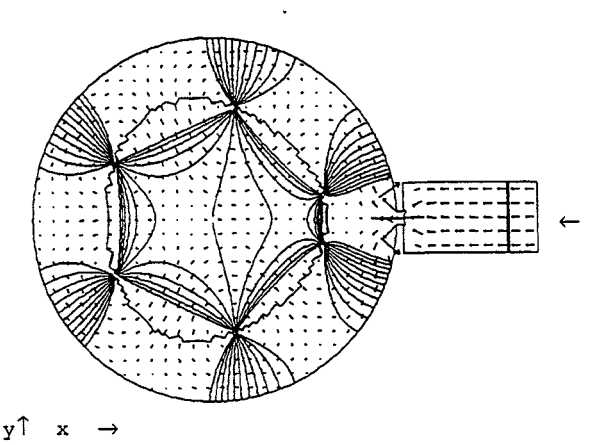


Figure 1: *Power flow (arrows) and equiphase lines in the asymmetrically fed fountain cavity. (the jagged line appears as an artifact of the discretization close to the turning points where $H_0 \approx 0$). Phase lines plotted in the range -500 to 500 μrad.*

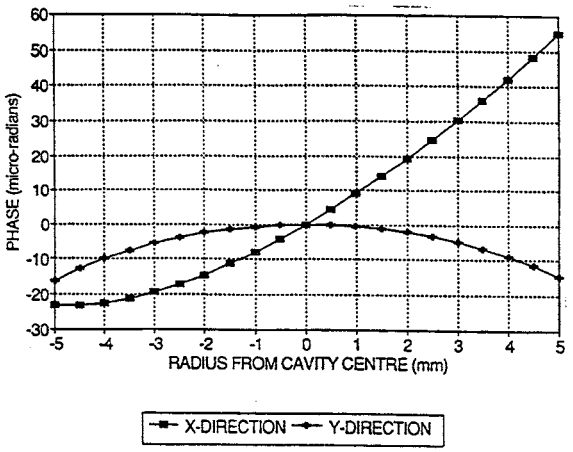


Figure 2: *Radial variation of phase distribution in the asymmetrically fed fountain cavity along the horizontal ($\hat{x}$) and vertical ($\hat{y}$) directions.*

values, using a mesh resolution of 9211 nodes; in general, the number of mesh nodes needs to be gradually increased until no observable changes in the field distribution is produced. The above mentioned agreement can thus be taken as a validation of the numerical electromagnetic approach taken in this paper.

Results and Discussion

In the calculations described here, the source plane is located far enough from the cavity-waveguide discontinuity so that any coupling between the fields at this plane and the fields inside the cavity can be neglected.

The spatial phase map of H_z around the centre of the fountain cavity was not found to be sensitive to the width of the iris separating the cavity from the input wavguide; likewise, spatial phase variations around the cavity centre were also unaffected by the type of source boundary (electric or magnetic) specified.

Fig. 1 depicts maps of the power flow (arrows) in the asymmetrically fed fountain cavity cross-section, superimposed on the equiphase contour lines. Power flow appears normal to equiphase lines, which constitutes a consistency check of the numerical solution with respect to the general relationship in (8).

The radial phase variation of the magnetic field is shown in Fig. 2 along the horizontal ($y = 0$ cut) and vertical ($x = 0$ cut) direction. from the centre of the cavity up to 5mm off-axis.

The effect of the source is mostly evident along the x-direction, where the phase steadily decreases from the source plane to the centre of the cavity. Consistent with the relationship in (8), the phase passes through a turning point some 5mm beyond the centre of the cavity, precisely where the power flow reverses direction. On the other hand, the zero phase gradient and power reversal points along the vertical direction occur at the centre line of the cavity, reflecting the symmetry of this plane with respect to the source of power.

The above result suggests that a phase uniformity improvement is possible by symmetrizing the feed: therefore, the spatial phase variations inside the fountain cavity were next investigated for symmetric waveguide coupling, where identical sources are used at opposite sides of the cavity. Since all fields will be symmetric about the centre lines of such a cavity, it is necessary only to consider the fields in a quarter section of the total cavity domain.

Fig. 3 displays the power flow and equiphase lines in the symmetrically fed fountain cavity.

The spatial phase variations around the centre of the cavity are shown in Fig. 4. Over a 5mm radius from the centre of the cavity, they indicate a reduction in the maximum phase variation by a factor of 3 as compared to the asymmetric case: the maximum phase variation fell from 54 μrad to below 14 μrad. Moreover, the phase variations

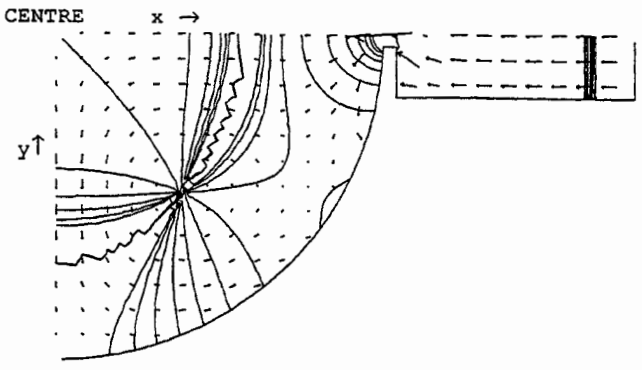


Figure 3: *Power flow (arrows) and equiphase lines in the symmetrically fed fountain cavity (only quarter section of the cavity shown). Phase lines range from -500 to 500 μrad.*

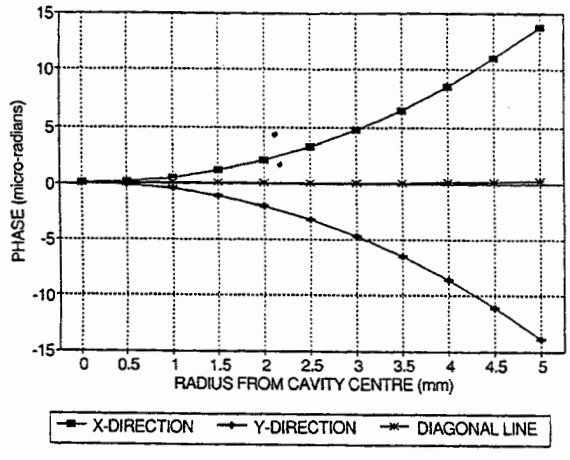


Figure 4: *Radial variation of phase distribution in the symmetrically fed fountain cavity along the horizontal ($\hat{x}$) and vertical ($\hat{y}$) directions, and along the diagonal line through the centre.*

have an odd plane of symmetry along the diagonal line (45° gradient), where their maximum value reaches 0.3 μrad at a radius of 5mm. This plane of symmetry only exists close to the cavity centre, within approximately a radius of 5mm, beyond which it rapidly rises in value.

Unlike the asymmetric case, the change in direction of the power flow is at the cavity centre, in both horizontal and vertical directions. This null power flow point coincides precisely with a zero phase gradient point, providing further illustration of the relationship between power flow density and spatial phase gradient given by (8).

An appraisal of the mismatch of power levels between the two sources that can be tolerated can be based upon the following considerations. A total mismatch, one where the power level of one source is forced to be zero, will produce the asymmetric phase response shown in Fig. 2, while a perfectly balanced source arrangement will produce the symmetric phase response shown in Fig. 4. Asymmetry in power levels between the sources will produce asymmetries in the spatial phase variations which lies between these two extremes. Therefore power imbalance is not expected to have a significant effect on the spatial phase distribution within a cavity fed from both sides. A phase difference between the two sources is likewise expected to give no effect, as it would just affect the coupling of the two sources with the resonant field in the cavity.

Conclusions

The general conclusion that can be drawn from these results is that a symmetrically fed cylindrical cavity, resonant on Cesium frequency in the TE_{011} mode, shows spatial phase variations around its centre, where atoms pass, which appear acceptable for an accuracy of 10^{-16} on the end-to-end phase shift. However, in drawing this conclusion, several cautions must be put forward:

1. These results are based on a 2D analysis. which allows no variations in the z direction. A complete 3D analysis would be needed to confirm their validity for the real cavity, particularly in consideration of the first order Doppler effect that atoms can see if they do not fly exactly vertically, parallel to the axis of the cavity.

2. The surface conductivity value used in the calculations reported in this paper is what can be found in any reference material handbook for OFHC Copper, whereas it is well known that the actual value not only varies with temperature, surface finish and fabrication technology, but can be a factor of 2 worse than what is quoted in the technical literature [9]. Furthermore, it is not obvious at this point that it be known exactly how stable this conductivity can be expected to be with time, and how much Cesium deposition or diffusion into Copper may change it. In the present paper a uniform conductivity was assumed throughout the microwave

structure walls.

3. As indicated in the paper, the sole TE_{011} mode was assumed to be excited in the cavity, both because the coupling to the source is supposed to discriminate against TM modes and because mode filters remove the degeneracy of the TM_{111} mode. Exactly to what extent this is true was not analyzed here, and may need to be investigated if the claimed μrad accuracy is seriously targeted for the phase.

4. This analysis was carried out with the assumption that the cavity be empty, which is not the case when Cesium atoms go through it. The effect of the interaction with the cloud of resonant atoms should be also investigated.

Acknowledgements
The work reported in this paper was carried out as an EEC funded Human Capital and Mobility research project, under contract number ERBCHBICT941115.

References

[1] S.Ghezali, P.Laurent, S.Lea, G.Santarelli, M.Bahour, K.Szymaniec, A.Clairon, "A Cesium fountain frequency standard: Present status and future improvements", these Proceedings.

[2] G. Vecchi and A. De Marchi, "Spatial phase variations in a microwave cavity for use in a cesium fountain primary frequency standard", *IEEE Trans. Instruments and Measurments*, Vol. 42, No. 2, pp. 434-438, April 1993.

[3] A. Khursheed and A. De Marchi, "Spatial phase distributions within cavities for cesium fountain standards", *Proc. of the 8th European frequency and time forum (EFTF 94)*, Vol. 1, pp. 65-70, March 1994.

[4] A. Khursheed, "The boundary fitted coordinate method to improve the numerical analysis of electron optical systems", *Scanning*, Vol. 16, pp. 201-208, 1994.

[5] R. Parodi and P. Fernandes, "Higher order modes in RF cavities", *IEEE Trans. Magnetics*, Vol. 24, pp. 154-157, January 1988.

[6] J.Van Bladel: *Electromagnetic Fields*, Washington: Hemisphere Publishing Corporation, 1985; Sec. 10.4

[7] A.Kursheed, G.Vecchi, A.De Marchi, "Spatial variations of field polarization and phase in microwave cavities: application to the Cesium fountain cavity", to be published on *IEEE Trans. on Ultrason. Ferroel. and Freq. Control*

[8] N. Marcuvitz, *Waveguide Handbook*. New York: McGraw-Hill, 1951.

[9] A.Bauch, T.Heindorff, R.Schroeder, "Measurement of the frequency-shift due to distributed cavity phase difference in an atomic clock", *IEEE Trans. Instruments and Measurments*, Vol. 34, No. 2, pp. 136-138, June 1985.

SESSION 9

SAW I

Chairman : V. PLESSKY

SAW FILTERS FOR MOBILE COMMUNICATION APPLICATION

G. Riha, K. Anemogiannis, J. Bauregger , J. Machui
Siemens Matsushita Components, Munich, Germany

Abstract

Cordless and cellular phone systems are one of the most rapidly growing areas of telecommunication and have become part of modern life. In addition personal communication with global access and wireless data networks will emerge. The development of such systems concentrates on both technological evolution and on benefits in functionality.

Handheld phones provoked a continuously increasing demand on miniaturised high performance components and led to the establishment of SAW filters in their radio circuits. Inversely the specific performance of SAW components already influenced the architecture of the receiver circuit. The radio quality could be increased, weight, size and total circuit cost were reduced. The present paper will concentrate on representative solutions to make the benefits of SAW devices transparent.

I. Introduction

Modern mobile communication systems have radio frequencies between 800 MHz and 2.0 GHz. Cellular systems like the *Advanced Mobile Phone System* (AMPS), the *Global System for Mobile Communication* (GSM) and the *Nippon Total Access Communication System* (NTACS) have an RF band in the 900 MHz range and a bandwidth of 25 MHz. Recently introduced systems like the *Personal Communication System* (PCS) and the *Personal Communication Network* (PCN) have an RF band in the 1.8 GHz range and a bandwidth of up to 75 MHz. Finally, cordless phone systems like CT1, CT1+ and CT2 operate in the 900 MHz range with bandwidths between 1 MHz and 4 MHz. Advanced digital cordless phone systems like the *Digital European Cordless Telephone* (DECT) operate at 1.9 GHz with a bandwidth of 20 MHz.

For about two decades now SAW filters are standard components in every TV IF stage. In radio link systems SAW channel and spectrum shaping filters are being preferred due to the excellent reproducibility of passband and stopband characteristics /1/. Soon after the definition of the first standards for digital mobile phone systems in Europe it became obvious, that also in this area only SAW technology could provide the high-performance IF-filtering required in low-cost circuit solutions /2/. A variety of new SAW techniques developed in order to achieve moderately low loss and high suppression of adjacent interferers in minimum size devices. Today, almost every GSM and DECT phone uses a SAW filter in the IF stage.

In the RF stage SAW devices compete with other filtering techniques, such as Micro Wave Ceramic (MWC). The main system demand on RF-filters in today's mobile phones besides high-preformance frequency characteristics is low cost. Yet, the growing importance of handheld phones in the past couple of years led to strong demands on the size and especially the height of components. IC technology already announced a maximum height of 1.6 mm. In addition, the radio stage area on the PCB must further shrink in order to allow for additional hardware bringing new features into the digital part of the phone. And finally, in order to reduce the size and weight of the phone battery, the power consumption of the circuits has to be minimized. Therefore components must offer high efficiency and low loss.

Over the last years a number of design techniques for low-loss SAW RF filters were developed and brought to market providing steep passband skirts and high stopband rejection. The insertion attenuation is now comparable to Micro Wave Ceramic (MWC) filters. New substrate materials with high electromechanical coupling coefficients allow for self matched operation at 50 Ω. Furthermore, the use of standardized materials and processes decreased fabrication costs and simultaneously increased product quality. Finally, the small size of RF SAW devices brought a break through in circuit design leading to the use of SAW instead of MWC filters.

Falling market prices in mobile communication exert considerable pressure on components and manufacturing cost. Low cost components help to achieve cost reduction. In addition cost efficient circuit architectures must be developped. The latter can be achieved by reduction of the number and of the cost of components. Eventually, the application of automatic PCB assembling requires SMD capability of all components.

This paper presents filter solutions for system architectures based on SAW components. Section II gives detailed information on the structure of modern mobile communication systems. Section III introduces current cordless phone standards and describes a particularly cost-effective SAW based solution for the Rx path of CT1+ phones. A summary of cellular phone standards will be given in section IV and its particular challenges to SAW RF filter techniques will be explained on the examples of AMPS and GSM requirements. Finally, GSM and PCN IF filters will be delt with in section V.

II. System architecture

The radio parts of wireless communication systems consist of the transmit path (Tx) and the receive path (Rx). In addition, one or more oscillators generate the signals for up- and down conversion in the mixers. Filtering is distributed to six functionalities with strongly differing performance requirements (Fig. 1).

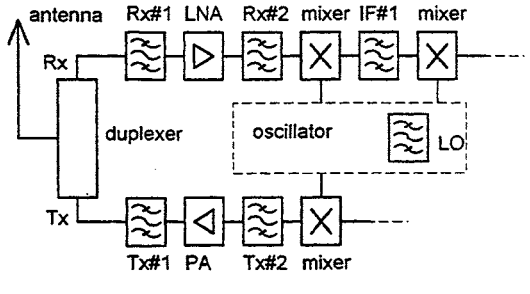


Fig 1: *The RF circuit of a wireless communication system demands six different filtering functionalities.*

Rx#1:	*receiver front-end filter*
Rx#2:	*receiver interstage filter*
IF#1:	*receiver IF filter*
Tx#1:	*transmitter front-end filter*
Tx#2:	*transmitter interstage filter*
LO:	*local oscillator clean-up filter*
LNA:	*low noise amplifier*
PA:	*power amplifier*

In the *transmit path* the modulated signal is converted up to the radio frequency. In general, in order to convert a modulated signal up to a radio frequency of 1 GHz double conversion is required. Recently developped IC technology allows direct conversion to the RF band. Modulator, amplifiers, converter and power control circuits are integrated in a single IC. Direct conversion transmitters are already state-of-the art for systems operating at 1 GHz like AMPS, NTACS and CT1+. Direct conversion for systems at 2 GHz is expected to come out soon.

The *power amplifier (PA)* boosts the signal up to the appropriate level. Cellular hand held units transmit a power of between 600 mW to 2 W, cordless phones offer some 10 mW. Depending on the transmit power level the PA may consist of several stages. In order to save battery life time the amplifiers are designed for low current C-class operation. Consequently, noise and higher harmonics appear at the PA output. Two filters are necessary to prevent performance degradation due to strong non-linearities of the PA. The *transmitter front-end filter (Tx#1)* after the PA suppresses higher harmonics and noise in the Rx band. The *transmitter interstage filter (Tx#2)* between mixer and PA suppresses close-in noise to prevent its convolution by the PA into the Tx-band.

In the *receive path* the weak incoming radio signal has to be amplified and converted down to the base band for demodulation. At first, the weak antenna signal has to be cleaned of strong spurious, e.g. signals from other radio systems and noise. It is equally important to suppress the strong Tx signal applied at the antenna. The *receiver front-end filter (Rx#1)* immediately after the antenna must prevent an overdrive of the sensitive low-noise amplifier (LNA) and assure its linear operation. The receiver must process signals within a dynamic range of up to 120 dB with high linearity. For that purpose the received signal must be down-converted through one or more IF stages. The *receiver interstage filter (Rx#2)* suppresses unwanted signals in particular at the image frequency.

In the intermediate frequency (IF) stages of the receive path the signal must be cleared of strong closely adjacent interferers stemming from other users of the same communication system. Hence, in general the *receiver IF filter (IF#1)* must provide high selectivity and narrow bandwidth while preserving a high degree of linearity.

The *antenna duplexer* is a three port element connecting the antenna with receiver and transmitter. In frequency domain multiple access (FDMA) systems like AMPS and NTACS transmission and reception occurs simultaneously and the duplexer consists of a passive matching network using either $\lambda/4$ transmission lines or discrete reactive elements. In time domain multiple access (TDMA) systems like GSM, PCN, PCS and DECT transmitting and receiving occurs in different time slots. In this case, the duplexer consists of a diode network, switching the antenna either to the receiver or the transmitter. Pin-diode pairs tend to be used for this purpose due to their low switch transit time at high frequencies.

Pure time domain duplexing systems use the same frequency band for transmission and reception. All RF filters have the same passband and in many cases the filters Tx#1 and Rx#1 can be substituted by a single filter (RxTx) placed between antenna and duplexer. The duplexer consists of a switch as shown in fig. 2.

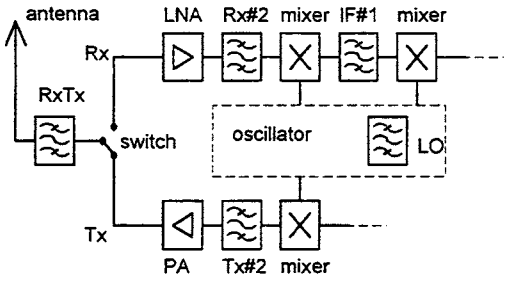


Fig 2: *In pure time multiplexing systems the filter RxTx may substitute the front-end filters Rx#1 and Tx#1.*

Radio circuits of mobile phones must conform to a fair number of performance and quality requirements. The main technical targets are the following:

1) High receive sensitivity.
The receive sensitivity is mainly influenced by the signal-to-noise (S/N) ratio. The noise figure F of the receiver is given by:

$$F = 1 + IL_{Rx\#1}\,(F_{LNA}-1) + (F_{mixer}-1)\,IL_{Rx\#1}\,IL_{Rx\#2} / G_{LNA} + \ldots$$

$IL_{Rx\#1}$ and $IL_{Rx\#2}$ are the insertion attenuations of Rx#1 and Rx#2 respectively. F_{LNA} and F_{mixer} are the noise figures of the LNA and the mixer. Finally, G_{LNA} is the gain of the LNA. According to the above formula and for high values of G_{LNA} the S/N ratio of the entire receiver is primarily given by the noise figure of the LNA and the attenuation of filter Rx#1. Hence, the insertion attenuation of Rx#1 has to be extremely low.

2) Low interference level.
Several factors influence the interference level. Adjacent channels must be suppressed in the IF stage in order to prevent co-channel interference in the demodulator. The rejection requirements for adjacent interferers will depend on system regulations, on additional filtering in the baseband and on the dynamic range of the demodulator. In the RF stage the suppression of the image frequency, of the strong Tx signal and of the oscillator signal are important parameters for low interference. The image frequency suppression in analog systems like AMPS and CT1+ has to be about 80 dB. In digital systems the image frequency has to be suppressed by some 70 dB. The suppression of the Tx signal has to be better than 40 dB in FDMA and better than 20 dB in TDMA systems in order to prevent an overdrive of the LNA. Finally the suppression of the LO signal depends on the system architecture and should be between 40 dB and 80 dB.

3) Low spurious radiation.
The suppression of noise, mixing spurious and higher harmonics in the transmitter prevents interference with other communication systems. The maximum level of the radiated spurious depends on the system and is specified by the regulations of the national and international communication commissions.

4) Low power consumption.
Losses in the transmitter filters Tx#2 and in particular Tx#1 mean additional power consumption and reduce battery life time. As a good deal of the total current is used to power up the transmit signal the insertion attenuation of Tx#1 is crucial.

5) Large operation temperature range.
Most cellular phones are designed for operation in a temperature range of -20°C to +75°C. However, in some cases an extended range between -30°C and +85°C is required. Of course filter specifications must be fulfilled in the entire temperature range. This requirement is particularily felt in the case of narrow band IF filters.

All above requirements are similar in the different cellular and cordless phone systems. However, there are particularities in the specific applications. In sections III to V some solutions are described in detail.

III. SAW based architecture for cordless phones

Established cordless phone systems operate in the range of 900 MHz and occupy a radio bandwidth of between 1 MHz (CT1) and 4 MHz (CT2). They still represent the largest market segment. Lately digital systems operating in the range of 1.9 GHz with radio bandwidths of some 20 MHz have gained much attention. DECT in Europe and PHS in Japan are expected to win considerable market shares. Table 1 shows the radio frequencies allocated to cordless phone standards.

Standard	Tx band [MHz]	Rx band [MHz]	Bandwidth [MHz]
CT1	914 - 915	959 - 960	1
CT1+	885 - 887	930 - 932	2
CT2	864 - 868		4
CT ISM (1)	902 - 928		26
DECT	1880 - 1900		20
PHS	1895 - 1918		23

Table 1: *Frequency allocation of cordless phone systems. (1) free Tx/Rx allocation within the 26 MHz band.*

CT ISM in the USA represents an interesting new development. Here the allocation of Rx and Tx frequencies within the 26 MHz bandwidth and the choice of data transmission modus is largely up to system providers. One approach uses *Code Division Multiple Access (CDMA)* to increase the maximum number of subscribers: the original digital information is modulated using orthogonal codes which allow to distiguish between subscribers. The modulation leads to an essentially *spread frequency spectrum*. On the other hand a huge number of subscriber may use the same radio band for transmission.

According to their market importance we will now concentrate on 900 MHz standards and more specifically on the *receiver circuit of CT1+* telephones. System specifications can be fulfilled using a double heterodyne receiver architecture as sketched in fig. 3. It consists of a front-end filter, a low-noise preamplifier and an interstage filter in the RF stage. The signal is then mixed down to the 1st IF. After appropriate image frequency suppression a second mixer converts the signal to the 2nd IF for channel filtering and demodulation. Usually, the 2nd IF is at 455 kHz. Narrow-band ceramic filters can provide sufficient channel selectivity for the following demodulation.

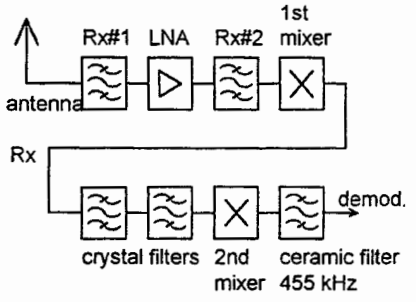


Fig 3: *Double heterodyne receiver architecture for analog CT1+*

In order to achieve a decent radio quality analog systems are specified to have very strong co-channel interference suppression. This implies high image-frequency rejection of up to 80 dB before every conversion. Since filter technology and PCB isolation do not allow such high selectivity with a single filter, two crystal filters are used in the 1st IF. In the RF stage the image-frequency rejection is given by both the front end and the interstage filters.

	SAW	MWC 2 pole	MWC 3 pole
insertion attenuation	3.5 dB	3.0 dB	5.3 dB
suppression at ± 21.4 MHz	45 dB	17 dB	31 dB
suppression at ± 42.8 MHz	45 dB	26 dB	45 dB
area in mm^2	11.3x4.8	12x10	12x15
height in mm	3.3	4.5	4.5
cost relation	1	0.8	1.2

Table 2: *Characteristics of SAW and MicroWave Ceramic filters, typical values.*

In the past, *Microwave Ceramic Filters (MWC)* were used in the RF stage. Yet, 3 pole MWCs are bulky (see table 2). For this reason, most systems apply 2 pole MWCs. However, in order to achieve the 80 dB image frequency suppression and due to the moderate close-in selectivity of 2 pole MWCs the first IF must be chosen at some 50 MHz or higher. At 50 MHz and above crystal filters are expensive. Furthermore, the connection of two filters in cascade requires discrete matching components, which must often be tuned in high-volume manufacturing.

SAW RF-filters offer the chance to decrease IF-filter cost and to avoid tuning. Such a filter solution must provide sufficient close in stop band rejection.In addition it must offer low insertion attenuation, sufficient temperature stability and low cost. These requirements can be achieved using dual-track filters with transducer coupling on a $LiTaO_3$ 36°rotY/X with its moderate temperature coefficient of -30 ppm/K. SIP4M leaded metal housing is preferred to smaller SMD ceramic packages in order to keep costs low.

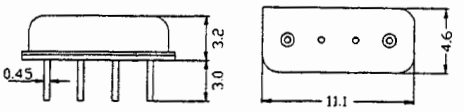


Figure 4: SIP4M leaded metal package

Dual-track filters consist of two acoustic tracks electrically connected in series /8/ (fig. 5). Coupling of the two tracks is governed by impedance matching of two identical coupling transducers placed in the center of each track. By appropriate dimensioning of the electrode configuration a purely real impedance of the track-coupling transducers in the passband can be achieved. In this case the electrical matching condition between the two tracks is fulfilled leading to tight track-coupling. In the stopband, however, the coupling transducers show strongly reactive impedances. The two tracks are electrically mismatched yielding high stopband rejection. The shape of the filters passband can be optimized using a mixture of solid and split electrodes in the coupling transducers. The outer transducers serve to convert external electrical signals into acoustics and vice versa. Their dimensions can be chosen independently, e.g. to achieve filters self-matched to a 50 Ω environment. Dual-track filters are particularily suited for narrow band applications with high close-in rejection requirements.

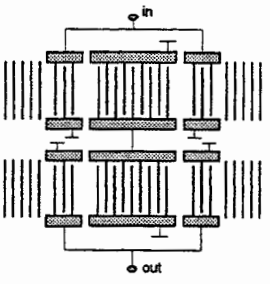


Fig 5: *Dual-track filter with transducer coupling: configuration*

Fig. 6 shows the frequency response of a SAW dual-track filter developed for CT1+. Its centre frequency is 886 MHz and the usable bandwidth 2 MHz. Minimum insertion loss and the amplitude ripple in the passband are typically 2.5 dB and 1 dB, respectively. The selectivity at ±21.4 MHz and ±42.8 MHz is typically 50 dB. There is no matching required for 50 Ω operation. Its fabrication demands line widths of down to 0.5 µm. This can be achieved using i-line steppers and a lift-off process /9, 10/.

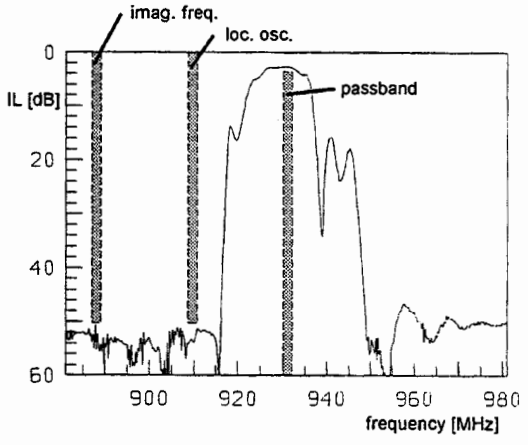


Fig. 6: *Dual-track filter for CT1+: frequency response*

Cordless phone standards using larger relative bandwidths will need different filtering techniques. Solutions for cellular phones as sketched in the next paragraph can fairly readily be adapted to the needs of DECT, PHS or CT ISM.

VI. SAW RF filters for Cellular Phones

Table 3 shows the radio frequencies used in common cellular phone systems. Established systems like AMPS in the USA are mostly analog and operate in the 900 MHz range. They occupy radio bandwidths between 25 MHz and 33 MHz. During the last couple of years two digital systems have been introduced in the same frequency range. GSM and JDC started operation in Europe and Japan respectively. GSM and JDC have a radio bandwidth of 25 MHz and 20 MHz respectively. In the meantime, the GSM recommendations have been reviewed and an extended GSM (EGSM) radio band of 35 MHz was defined corresponding to a subscriber capacity increase of 40%.

System	Tx band [MHz]	Rx band [MHz]	Bandwidth [MHz]
AMPS	824 - 849	869 - 894	25
ETACS	872 - 905	917 - 950	33
NTACS	898 - 925	843 - 870	27
GSM	890 - 915	935 - 960	25
EGSM	880 - 915	925 - 960	35
PCN	1715 - 1790	1805 - 1880	75
PCS	1850 - 1910	1930 - 1990	60
JDC-800	940 - 960	810 - 830	20
JDC-1.5	1430 - 1460	1480 - 1510	30

Table 3: *Frequency allocation of cellular phone systems.*

The demand for more subscriber capacity and new services in the cellular phone market brought the allocation of additional radio bands in the 1.8 GHz range. In Europe the PCN/DCS1800 system started operation in 1994. In the USA the PCS system will begin soon and in Japan JDC-1.5 GHz has already been introduced. They occupy radio bandwidths of up to 75 MHz and are designed for digital transmission.

In spite of their fair number cellular phone systems show a remarkable amount of similarities, at least with respect to requirements on filter design and fabrication. In particular relativ bandwidths lie in the range of 2.5 % to 4.5 % and to a large extent the same design approaches may be used for all standards /5/. A similar statement holds for housing where hermetically sealed *SMD-type ceramic packages* are standard. Size and more particularily height is crucial. Fig. 7 shows the DCC6, the more recent 3.8x3.8 mm^2 successor of the early 5x5 mm^2 QCC8 package. Again we follow market shares and take filter solutions for AMPS and GSM in order to explain the two most popular design techniques for SAW RF filters. We shall show that these approaches demand solid electrodes only. Hence a resolution of 0.5 µm is sufficient to fabricate devices operating in frequency ranges up to 2 GHz.

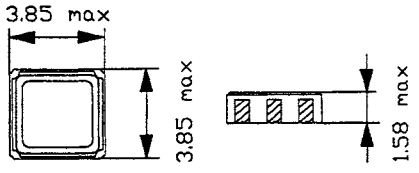


Figure 7: *Hermetically sealed SMD-type ceramic package DCC6.*

As a first example we take a transmitter filter for the analog AMPS system. It serves to suppress noise and spurious signals generated in the Tx path. AMPS operates in FDMA modus and the filter must offer particularily high suppression in the Rx-band. In addition front-end filters require very low loss in the passband and excellent power handling capabilities. Resonator network filters are best suited for this task. The substrate material is again $LiTaO_3$ 36°rotY/X demanding only about ±1 MHz tolerance for temperature drift in the range of -20°C to +80°C.

Multiple-resonator filters consist of a number of one-port SAW resonators integrated on one chip /6, 7/. Their frequency dependent impedances are suitably combined to form the desired filter characteristic. Fig. 8 shows configuration and equivalent circuit of a single one-port resonator. An interdigital transducer (IDT) is placed between two reflectors on a piezoelectric substrate. All acoustic structures use only solid electrodes. The equivalent circuit consists of an electrostatic capacitance C_{st}, defined by the electrode configuration of the IDT and the permittivity of the substrate. In parallel to C_{st} there is a series resonance circuit L_a-C_a, describing acoustic resonances in the SAW resonator.

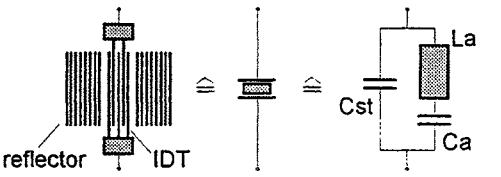


Fig. 8: *Configuration of a one-port SAW resonator and its equivalent circuit*

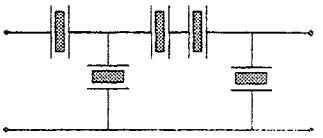


Fig. 9: *Multiple resonator technique: scheme of a filter with five resonators*

In order to get a filter characteristic a number of resonators must be electrically connected to form a network. Fig. 9 schematically shows a filter configuration with five resonators. The high electroacoustical coupling coefficient of $LiTaO_3$ 36°rotY/X (6.6%) allows to achieve wide-band 50 Ω solutions with this design approach. Stopband rejection level, roll-off factor and passband ripple can be optimised by means of LC-filter design rules.

The main advantage of the multiple-resonator technique is low insertion attenuation in combination with a wide passband characteristic. On the other hand stopband rejection remains moderate even in frequency ranges far from the passband. Fig 10 shows the frequency characteristic of the *Tx filter for AMPS*. The minimum insertion loss is 1 dB and the suppression of the Rx band is better than 20 dB.

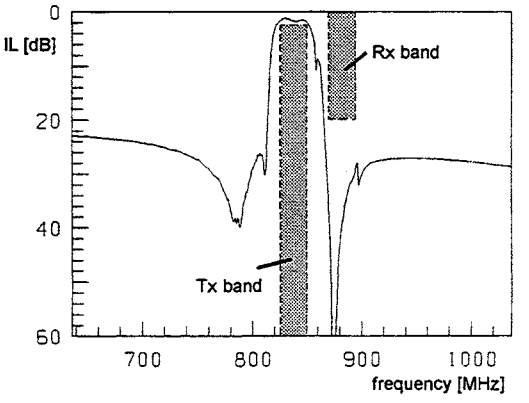


Fig 10: *Multiple-resonator filter for AMPS-Tx: frequency response*

The moderate stopband suppression of multiple resonator filters is not always sufficient for its application in the receiver circuit of mobile phones. The receiver filter must handle strong incoming spurious signals and it must suppress the image frequency before the signal can be converted down to a first IF. In case of GSM receiver filters must also provide strict uniformity over the passband to allow for transmission power management. The dual-mode resonator technique provides these features. But it demands a substrate material with even stronger electroacoustical coupling in order to achieve the desired wide-band 50 Ω solution. $LiNbO_3$ 64°rotY/X with a coupling coefficient of 11% fulfills these needs but demands relatively large temperature tolerances of ±2.6 MHz.

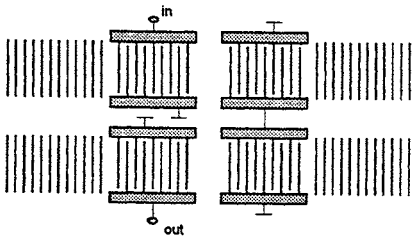


Fig 11: *Dual-mode resonator filter: configuration*

Dual-mode resonator filters /11, 12/ consist of two resonant acoustic tracks electrically connected in series as schetched in Fig 11. They use two longitudinal resonance modes in each track to achieve a flat passband. One resonance stems from the eigen-mode of the solid electrode transducer gratings and the second one corresponds to a standing wave in the cavity between the two transducers /13/. Again only solid electrodes are needed. The limited bandwidth of acoustic transmission leads to very attractive stopband rejection. Dual mode resonator filters may be operated with balanced or unbalanced termination at input and output. This feature allows to omit a balun e.g. in case of an interstage filter in combination with a Gilbert mixer.

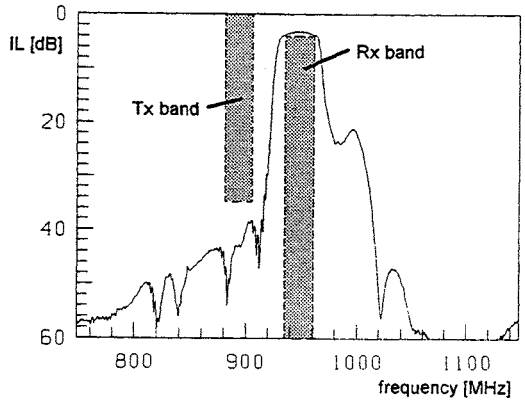


Fig 12: *Dual-mode resonator filter for GSM-Rx: frequency response*

Fig 12 shows the frequency response of a *GSM Rx filter*. The minimum insertion loss and the passband ripple are typically 3 dB and 0.5 dB respectively. Attenuation in the Tx-band is better than 35 dB. High rejection in the lower stopband is appropriate for high image frequency suppression in a wide frequency range. This gives additional flexibility for a low cost design of the IF and the oscillator stages. The filter offers sufficient image frequency supression even for an IF of 45 MHz.

V. Channel selection in GSM and PCN phones

Ideally SAW IF filters in the receive path contribute a good deal to the total channel selection required by the system. They must also suppress strong interferers from other subscribers to prevent saturation of the IF amplifier and provide good selectivity for image frequencies if a second IF is present. For filter technology channel selection is the most demanding task. Again following market importance typical solutions will be highlighted on the example of GSM. For PCN and PCS almost the same requirements hold and PHS and DECT show very similar features. Generally speaking, SAW devices provide IF filter solutions for the complete range from very narrow (i.e. AMPS 15 kHz) to large (DECT 1.2 MHz) channel bandwidths.

GSM uses Gauss Minimum Shift Keying (GMSK) modulated signals with a channel to channel spacing of 200 kHz. Its main part is concentrated in a frequency band of some 160 kHz, but its skirts reach much further out such that adjacent signals overlap. Adjacent interferers can be quite a bit stronger than the desired signal and hence the GSM recommendations demand a total suppression of 18 dB for the first, 50 dB for the second and 58 dB for the third neighbouring channel. Total blocking requirements for frequencies further than 600 kHz from the signal carrier reach from 69 dB to ultimately 84 dB. To achieve significant suppression of the first adjacent interferer filter selection must start at ±110 kHz. The SAW IF filter cannot fulfill the entire selection requirements but it must contribute significantly to channel selection in order to reduce additional filtering in a second IF or in the baseband to save cost and current.

Frequencies for the first IF in GSM go from 45 MHz up to over 100 MHz, with 71 MHz being a particularly popular choice. Filters for these applications must provide moderate insertion attenuation (typically in the range of 8 dB) to maintain a low noise level and to limit power consumption. To permit an undisturbed passage of the signal low ripple in amplitude and group delay are required. For handheld phones size and, of course, cost of filter components are crucial and they must offer temperature stability in the range of -20°C to 80°C. The obvious choice of substrate material for these requirements is quartz. And figure 13 shows two out of a fair number of different space saving SMD packages for GSM IF filters that can be found on the market.

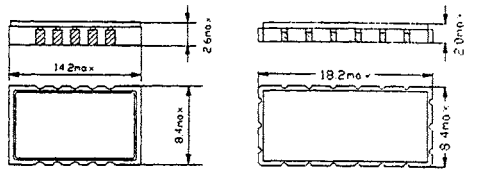


Figure 13: *Hermetically sealed SMD-type ceramic packages DCC14 and QCC18 for GSM IF filters.*

The demand for low loss in connection with good design possibilities triggered the development of *Single Phase UniDirectional Transducers (SPUDT)* /14, 15/. The directionality of SPUDT transducers reduces losses and allows for low echos and hence for proper passband behaviour in filters that are electrically matched to their environment in order to achieve low insertion attenuation. To first order their design follows established rules for transversal filters, yielding in particular an estimate of the length of the acoustically active track. For GSM channel selection transversal filters need large time windows i.e. acoustic lengths corresponding to outside dimensions of over 30 mm in length which is not acceptable for handheld phones.

Hence SAW designers had to provide for longer acoustic paths on limited substrates. A considerable number of innovative resonating and multi-track solutions was suggested /e.g. 16, 17/. Here we present *Z-path filters* /3/ as they permit to design for excellent selectivity in comparably small packages (fig. 14). Z-path filters achieve longer time domains by folding the acoustic propagation into a zig-zag path. Four selecting elements (two directional transducers and two inclined reflectors) provide steep skirts and good stopband rejection. Weakly inclined reflectors preserve reasonably low loss and maintain the quartz substrate's excellent temperature characteristics.

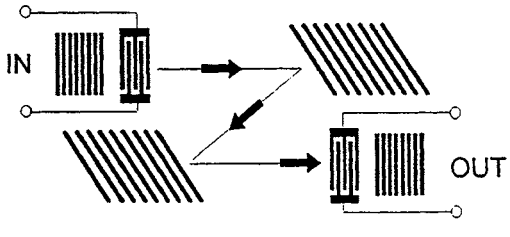


Fig. 14: *Configuration of a Z-path filter: arrows indicate the path of acoustic waves from input to output transducer.*

Fig. 15 shows the frequency response of a Z-path filter operating at 45 MHz. Its typical minimum insertion attenuation is 8 dB and the group delay deviation in the passband is better than 1 µs. Selectivity at ±200 kHz is better than 20 dB, while the suppression of the first adjacent channel is only 12 dB. The second and third adjacent interferers are suppressed by at 29 dB and more than 37 dB respectively. Due to its low frequency and the use of quartz the temperature tolerance this filter demands is only 5 kHz. Consequently additional selection in baseband or in a second IF may be limited to simple low order filters.

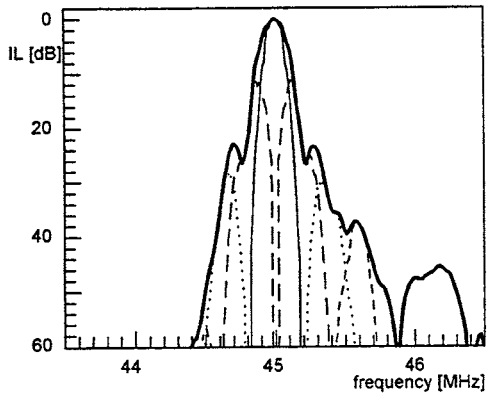


Fig 15: *Z-path filter: frequency response (fat solid line) and GSM signal with its first three adjacent interferers on both sides after filtering (fine solid and fine broken and dotted lines).*

With communication systems like PCN and PCS moving to higher frequencies and with their larger bandwiths intermediate frequencies also had to go up. Popular frequencies lie in the range of 250 MHz. Channel width and spacing remain the same as for GSM. The resulting smaller relative bandwidths allow for resonant SAW filter solutions. Resonator filters fit into the small QCC10 package (fig. 16) and provide additional savings in space and cost. On the other hand tolerances rise with higher frequencies requiring in particular even tighter control of all fabrication processes.

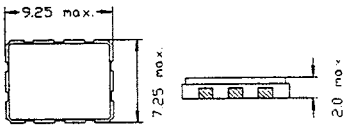


Figure 16: *Hermetically sealed SMD-type ceramic package QCC10.*

The most widely used design approach is the *transverse-mode coupled resonator filter* /18, 19/. The filter consists of double-track structures (fig. 17). Two distinct transverse modes propagate in a double-track, They provide acoustic coupling from input track to output track at their respective excitation frequencies. Using two double-tracks in series provides the required selectivity. Together with an external matching network transverse-mode coupled resonator filters achieve bandwidths of up to about 0.8 ‰ which is just sufficient to provide the required passband width. Larger bandwidths demand an additional external coil, the expansion coil, to allow coupling between the two double tracks.

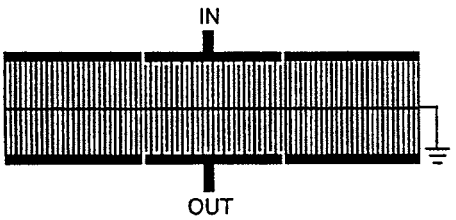


Fig. 17: *Configuration of transeverse mode coupled filter: structure of one double track. The complete filter consists of two to three double tracks in series.*

Fig. 18 shows the frequency response of a transverse-mode coupled resonator filter on quartz with expansion coil operating at 246 MHz. Its typical minimum insertion attenuation is 4 dB and the group delay deviation in the passband is better than 2 µs. Selectivity at ±200 kHz is around 5 dB, while the suppression of the first adjacent channel is 4 dB. The second and third adjacent interferers are suppressed by 25 dB and 37 dB respectively. Stopband rejection is typically 50 dB.

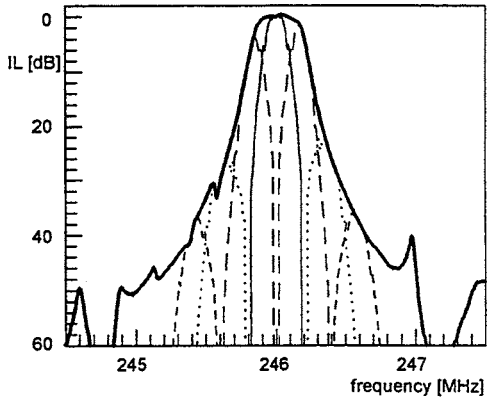


Fig 18: *Transverse mode coupled resonator filter for IF channel selection: frequency response (fat solid line) and GSM/PCN signal with its first three adjacent interferers on both sides after filtering (fine solid and fine broken and dotted lines).*

VI. Conclusion

SAW filters offer substantial advantages to the architecture of cordless and cellular phone circuits. Design approaches for lowloss components to be applied in RF and IF stages have matured over the last years and today SAW devices offer the performance required in mobile communication systems. Particularily favourable features are their small size and low cost. Low insertion attenuation combined with high close-in selectivity offers the possibility to choose cost effective system architectures and to reduce filter reqirements in low frequency stages.

References

/1/ C. Ruppel et al., "SAW Devices for Consumer Communication Applications", IEEE Trans. Ultras., Ferroel. and Freq. Control UFFC, 1993, pp. 438-452.

/2/ G. Riha, K. Anemogiannis, R. Dill, C. Kappacher, "SAW Low-Loss Filters for Mobile Communication Systems", Proc. Intern. Symp. SAW Devices for Mobile Commun. 1992, pp. 113-120.

/3/ J. Machui, W. Ruile, "Z-path IF-Filters for Mobile Telephones", IEEE Ultrasonics Symposium Proc. 1992, pp. 147-150.

/4/ K. Anemogiannis, J. Bauregger, G. Riha, " Cordless Phone Architectures based on SAW Filters", IEEE Ultrasonics Symposium Proc. 1993, pp. 85-88.

/5/ K. Anemogiannis, J. Bauregger, J.Machui, G. Riha, " NEW System Applications of Modern RF SAW Filters", IEEE NTC Mai 1995, to be published.

/6/ Y. Satoh, O. Ikata, T. Matsuda, T. Nishihara, T. Miyashita, "Resonator-Type Low-Loss Filters", Proc. Intern. Symp. SAW Devices for Mobile Commun. 1992, pp. 179-185.

/7/ T.Matsuda, U.Uchishiba, O.Ikata, T.Nishihara, Y.Satoh, "L and S Band Low-Loss Filters Using SAW Resonators", IEEE Ultrasonics Symp. 1994,to be published.

/8/ K. Anemogiannis, F. Mueller, W. Ruile, G. Riha, "High Performance Low-Loss SAW Filters for Mobile Radio with Improved Stopband Rejection", IEEE Ultrasonics Symposium Proc. 1988, pp. 77-82.

/9/ S. Berek, U. Knauer, H. Zottl, "I-line Lithography for Highly Reproducible Fabrication of Surface Acoustic Wave Devices", SPIE Symp. on Microlithography, Proc. Vol. 1463, 1991.

/10/ K. Anemogiannis, S.Berek, H.Zottl: "GHz Low-Loss SAW Filters Employing Standard Fabrication Process", IEEE Ultrasonics Symposium Proc. 1990, pp. 129-133.

/11/ T. Morita, Y. Watanabe, M. Tanaka, Y. Nakazawa, "Wideband Low Loss Double Mode Filters", IEEE Ultrasonics Symposium Proc. 1992, pp. 95-104.

/12/ T. Morita, Y. Watanabe, M. Tanaka, Y. Nakazawa, "SAW resonator filters for mobile communication", Proc. Intern. Symp. SAW Devices for Mobile Commun. 1992, pp. 186-193.

/13/ Y.Yamamoto, R.Kajihara, "SAW Composite Longitudinal Mode Resonator Filters and ...",IEEE Ultrasonics Symposium Proc. 1993, pp. 47-51.

/14/ C.S.Hartmann, P.V.Wright, R.J.Kansy, E.M.Garber, "An Analysis of SAW Interdigital Transducers with Internal Reflection and the Application to the Design of Single-Phase Unidirectional Transducers", IEEE Ultrasonics Symposium Proc. 1982, pp. 40-45.

/15/ C.Kappacher, O.Männer, W.Ruile, R.Dill, "Design and Analysis of Single Phase Unidirectional Transducers", IEEE Ultrasonics Symposium Proc. 1991, pp. 1-4.

/16/ P.Ventura, M.Solal, P.Dufilie, S.Chamaly, "A Global Optimization Prozedure for SPUDT Filters", IEEE Ultrasonics Symposium Proc. 1993, pp. 5-8.

/17/ M.Solal, J-M.Hode, "New Compact SAW Low Loss Filter for Mobile Radio", IEEE Ultrasonics Symposium Proc. 1993, pp. 105-110.

/18/ C.S.Hartmann, D.P.Chen, J.Heighway, "Modeling og SAW Transversely Coupled Resonator Filters Using the Coupled Modes Modeling Technique", IEEE Ultrasonics Symp. Proc. 1992, pp. 39-43.

/19/ G.Sholl, W.Ruile, P.H.Russer, "P-Matrix Modeling of Transverse -Mode Coupled Resonator Filters", IEEE Ultrasonics Symp. Proc. 1993, pp. 41-46.

Equivalent Circuit Configurations for SAW Coupled Resonator Filters and the Experimental Determination of their Elements

D.P. CHEN, C. LAMBERT, M-A. SCHWAB and C. HARTMANN*

Advanced SAW Products SA, CH-2022 Bevaix, Switzerland
*Hartmann Research Inc. Dallas, TX 75243, USA

ABSTRACT

A variety of modeling methods for SAW resonator and SAW coupled resonator filters have been successfully established. However, most of them are not compatible with CAD tools normally used by electronic circuit design engineers. Much effort has been devoted to modeling equivalent circuits of SAW single mode resonator filter, but not enough to SAW coupled resonator filters where usually there is more than one mode involved in the passband. Coupled resonator filters have been widely used in the rapidly growing mobile phone market. An accurate electrical circuit simulation approach for these types of filters would be very useful for system design because it could be directly compiled into CAD tools. This paper will present an unique way to construct the equivalent circuits for transversely or longitudinally coupled multiple mode coupled resonator filters and introduce an accurate approach for experimental determination of the elements in the equivalent circuit. The determined elements include both the dynamic (motional) elements and the electrostatic parasitic elements. The comparison between the theory and the measured results will be made for Rayleigh wave type (either transverse or longitudinal) of coupled resonator filters on Quartz. A good agreement is shown.

1. INTRODUCTION

Due to their good performances: low insertion loss, narrow transition band, high out-of-band rejection and miniaturized size, SAW coupled resonator filters are drawing more and more attention from system design engineers and are going to play bigger role in compact system applications. A variety of modeling methods for SAW coupled resonator filters have been successfully established. However, most of them are not compatible with CAD tools normally used by electronic circuit design engineers. Much effort has been devoted to modeling equivalent circuit of SAW single mode resonator or resonator filter but not enough to SAW coupled resonator filters where usually there are more than one mode involved in the passband [1-2]. The two mode (either transverse or longitudinal) coupled resonator filters are good examples. The established analysis and synthesis tools widely used by SAW design engineers are based on modeling techniques such as coupling-of-mode modeling technique [3-4] and P-matrix modeling technique [5]. These models work very accurately and efficiently but cannot be directly built in the commonly used electrical circuit CAD tools. The only way to connect the designed device to the system network is through S, Z or Y-matrix. Usually it is not a convenient and efficient approach which the system design engineers would like to use. This paper will present an unique way to construct the equivalent circuits for transversely or longitudinally coupled multiple mode resonator filters and introduce a simple and accurate approach for extraction of the values of the elements in the equivalent circuit from either the measured or the simulated performances of the devices. The determined elements include both the dynamic (motional) elements and the electrostatic parasitic elements. The former are governed by acoustic behavior and the latter are only dependent on the structure geometry. The basic assumption in this approach is that all the modes involved are the guided cavity eigenmodes and that they can be decoupled from the measured or the simulated S-parameters. The parameter extraction technique established for single mode resonator can then be applied to these eigenmodes separately. Afterwards they will be appropriately reconstructed to lead to an equivalent circuit model for the device. This process is discussed in the next section in details. In sections 3 and 4 the methods we use for the parameters' extraction is presented for both motional elements and electrostatic elements. In order to verify the accuracy of this approach, the comparison between theory and experiment is made for both transversely and longitudinally coupled resonator filters on Quartz. In the transversely coupled resonator filter only two fundamental modes are concerned. The behavior of the longitudinally coupled two port resonator seems more complicated. It is shown that more modes have to be taken into account in order to obtain a good description of the device performance. Apart from the two fundamental modes, three other symmetric modes are included to verify the modeling approach. The comparison is done based on the full S- or Y-parameters and is discussed in section 5. A good agreement can be seen. This approach can be readily built in the electronic CAD software and will be very helpful for system design engineers. It should be noted that this approach is valid assuming that the SAW filter works with an unique type of wave and that no wave mode convertion is involved over the frequency range concerned. Therefore it would not be suitable for devices where there are more than one type of waves coupling to each other over the studied frequency range.

2. DECOUPLING OF EIGEN MODES

In this section two types of multiple mode coupled resonator filters are used as examples to describe the idea of decoupling the guided cavity eigen modes from the Y-parameters of the device. All modes are grouped into two families, the symmetric and the anti-symmetric, but for the sake of simplicity only two modes are described here. However, this principle of decoupling the modes can be extended to all the modes. Fig.1 presents two typical structures of two mode coupled resonator filters, the transversely coupled one shown in (a) and the longitudinally coupled one shown in (b), along with the profile description of each mode.

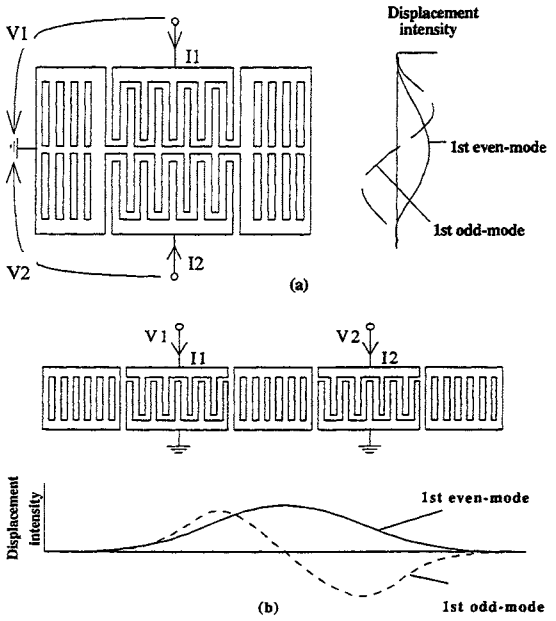


Fig.1 Basic schematic diagrams of two types of two port coupled resonator filters, (a) transversely coupled resonator, (b) longitudinally coupled resonator filters.

The electrical behavior of both structures can be expressed by a 2x2 Y matrix:

$$[I_i] = [Y_{ij}] \times [V_j] \qquad i, j = 1,2 \tag{1}$$

The mode profiles imply the following electrical conditions, which correspond to the mode generation and reception:

$$I_s = I_1 + I_2 \quad \text{and} \quad V_s = V_1 = V_2 \tag{2}$$

$$I_a = I_1 - I_2 \quad \text{and} \quad V_a = V_1 = -V_2 \tag{3}$$

where the subscripts s and a denote the symmetric and the antisymmetric eigenmode quantities respectively. Therefore the admittances of both the symmetric and the anti-symmetric modes can be separately deduced from the Y-parameters of the device.

$$Y_s = Y_{11} + Y_{22} + Y_{12} + Y_{21} \tag{4}$$

$$Y_a = Y_{11} + Y_{22} - Y_{12} - Y_{21} \tag{5}$$

In the case of more than two modes involved, the admittance of Y_s or Y_a should be understood as a summation of the admittances of all the modes of the same family. In this way the eigenmodes are decoupled. Then in principle the resonant behavior of each single mode can be modeled as a series R,L,C arm, and the corresponding motional and electrostatic elements can all be extracted from the decoupled admittances. The detailed procedures to determine these parameters are presented in the next two sections.

In turn, the Y matrix of the whole device can be expressed as:

$$Y_{11} = Y_{22} = \frac{1}{4}(Y_s + Y_a) \tag{6}$$

$$Y_{12} = Y_{21} = \frac{1}{4}(Y_s - Y_a) \tag{7}$$

Eq.(6) and (7) imply an equivalent circuit like the one shown in Fig.2(a), where there are two resonant arms, $y_s = \frac{Y_s}{4}$ and $y_a = \frac{Y_a}{4}$. When all the electrostatic parasitic elements are added, the complete equivalent circuit for the two port coupled resonator filter is shown in Fig.2(b), where R_b is the series resistance at input and output ports, R_g is the common ground resistance, C_b is the bridge capacitance between input and output and C_t is the transducer capacitance.

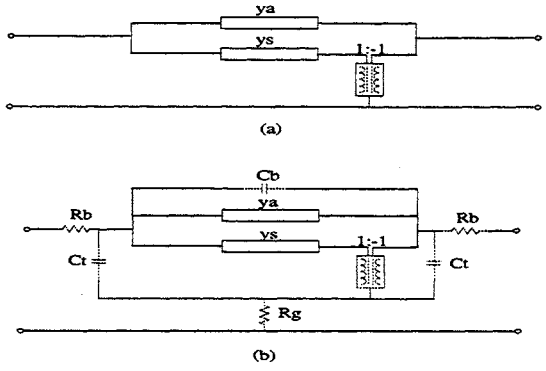


Fig.2: The unique equivalent circuit configuration for multiple mode coupled two port resonator filters, (a) only motional elements, (b) entire circuit.

3. EXPERIMENTAL DETERMINATION OF ELECROSTATIC ELEMENTS

The electrostatic parameters in the equivalent circuit (Fig.2(b)) can be determined based on the decoupled eigenmode admittances, Y_s and Y_a. According to the electrical driving or receiving conditions (2) and (3), in the low frequency range far away from the acoustic resonance the electrical equivalent circuits can be expressed by Fig.3(a) and (b).

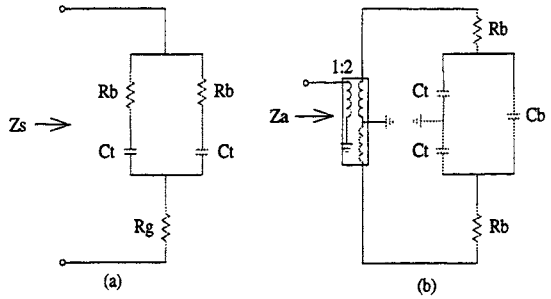


Fig.3: Simplified electrostatic equivalent circuit, (a) Symmetric driving, (b) Anti-symmetric driving conditions.

Their impedances can be deduced as:

$$Z_s = \frac{1}{Y_s} = (R_g + \frac{R_b}{2}) - j\frac{1}{\varpi\,(2C_t)} \tag{8}$$

for symmetric driving condition and

$$Z_a = \frac{1}{Y_a} = (\frac{R_b}{2}) - j\frac{1}{\varpi\,(2C_t + 4C_b)} \tag{9}$$

for anti-symmetric driving condition. Therefore:

$$R_g = real(Z_s) - real(Z_a) \tag{10}$$

$$R_b = 2real(Z_a) \tag{11}$$

$$C_t = \frac{1}{2(-\varpi imag(Z_s))} \tag{12}$$

$$C_b = \frac{1}{4\varpi\,(-imag(Z_a) + imag(Z_s))} \tag{13}$$

There *real* and *imag* denote the real part and the imaginary part respectively.

4. EXPERIMENTAL EXTRACTION OF MOTIONAL PARAMETERS FOR EACH MODE

The equivalent circuits for the decoupled single eigenmodes, which include the electrostatic elements, can be shown in Fig.4(a) and (b) and correspond to the admittances Y_s and Y_a.

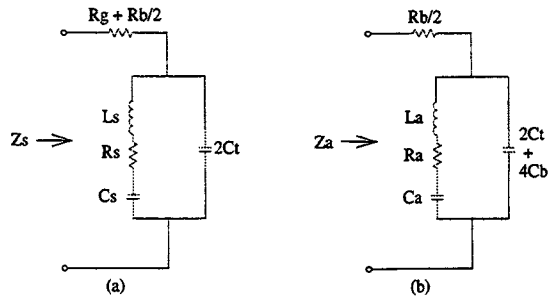


Fig.4: Equivalent circuit for decoupled single eigenmodes, (a) symmetric, (b) anti-symmetric driving conditions.

In order to accurately extract the values of the motional elements in resonant arms, the determined electrostatic elements need to be substracted from the Y_s and Y_a at first to obtain the admittances of the resonant arms $\overline{Y_s}$ and $\overline{Y_a}$. Then the motional elements of the resonant arms can be obtained in the following way. At first the synchroneous frequencies which make the reactances of each mode be zero are found as ϖ_s and ϖ_a. Suppose that $\overline{Z_s} = \frac{1}{\overline{Y_s}}$ and $\overline{Z_a} = \frac{1}{\overline{Y_a}}$, then for the symmetric arm:

$$R_s = real(\overline{Z_s}(\varpi_s)) \tag{14}$$

$$L_s = \frac{1}{2}(\frac{d}{d\varpi}(imag(\overline{Z_s})))_{\varpi=\varpi_s} \tag{15}$$

$$C_s = \frac{1}{L_s \varpi_s^{\,2}} \tag{16}$$

And for the anti-symmetric arm:

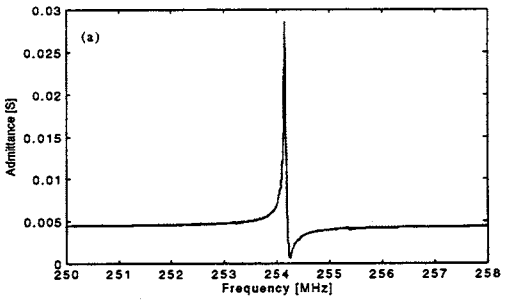
$$R_a = real(\overline{Z_a}(\varpi_a)) \tag{17}$$

$$L_a = \frac{1}{2}(\frac{d}{d\varpi}(imag(\overline{Z_a})))_{\varpi=\varpi_a} \tag{18}$$

$$C_a = \frac{1}{L_a \varpi_a^{\,2}} \tag{19}$$

5. COMPARISON OF THEORY AND EXPERIMENTAL RESULTS

In this section two experimental devices, where one is a doubled mode transversely coupled resonator filter at 250MHz on $37°\,YX-Quartz$ and the other is a two port longitudinally coupled resonator filter at 430MHz on $35°\,YX-Quartz$ are taken as examples to check the validity of the unique equivalent circuit as well as the accuracy of the parameter extraction approach mentioned above. Fig.5(a) and (b) present a comparison of the decoupled eigenmode admittances for measurement and equivalent circuit model for the 250MHz transversely coupled resonator filter. Fig.6 gives the insertion loss comparison.



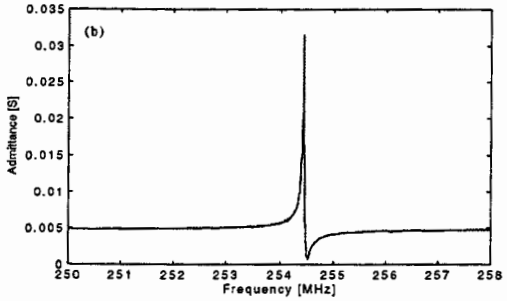


Fig.5: Comparison between measured and equivalent circuit admittances, (a) symmetric mode, (b) anti-symmetric mode.

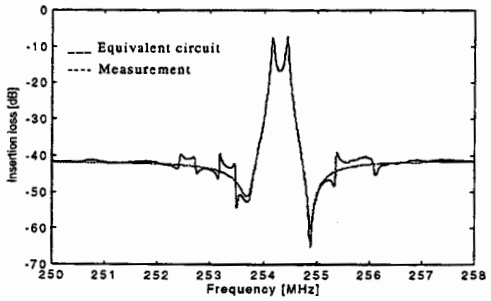


Fig.6: Comparison between measured and equivalent circuit $|S_{12}|$ for 250 MHz filter.

Parasitic elements (a)			
In/out resistance [ohms]	Resistance to ground [ohms]	Bridge capacitance [pF]	Transducer capacitance [pF]
R_b	R_g	C_b	C_t
5.643	3.503	0.05205	1.4108

Motional elements (b)					
Res. [ohms]	Ind. [μH]	Cap. [fF]	Res. [ohms]	Ind. [μH]	Cap. [fF]
R_a	L_a	C_a	R_s	L_s	C_s
116.930	807.905	0.48427	116.125	789.939	0.49638

Table1: Values of Parasitics (a) and Motional (b) elements

The values of the motional and the electrostatic parasitic elements are given in Table 1(a) and (b). The two port longitudinally coupled resonator filter, in fact, is designed to act as a single mode resonator filter. Since the undesired longitudinal modes are not suppressed completely, they apparently show up with reasonable Q value (shown in Fig.7). If it is modeled with single mode resonator equivalent circuit the result greatly deviates from the actual situation. However if it is treated with the equivalent circuit for multiple mode coupled resonator filter mentioned above. a more satisfying result is obtained. In Fig.7 we show the results corresponding to two modes (the fundamental symmetric and anti-symmetric ones), three modes (one more symmetric mode), four modes (two more symmetric modes) and five modes (three more symmetric modes) separately along with the measured result. It can be clearly seen that the more modes involved, the better the agreement with the measured results is. Fig.8(a) and (b) give the comparison of the eigenmode admittances.

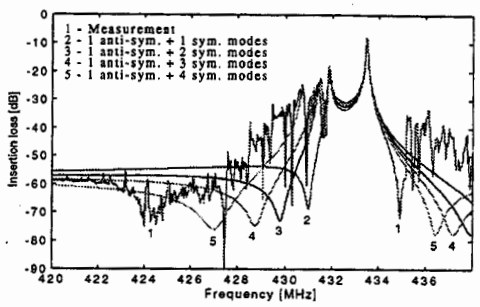


Fig.7: Comparison between measured and equivalent circuit $|S_{12}|$ for 430 MHz filter

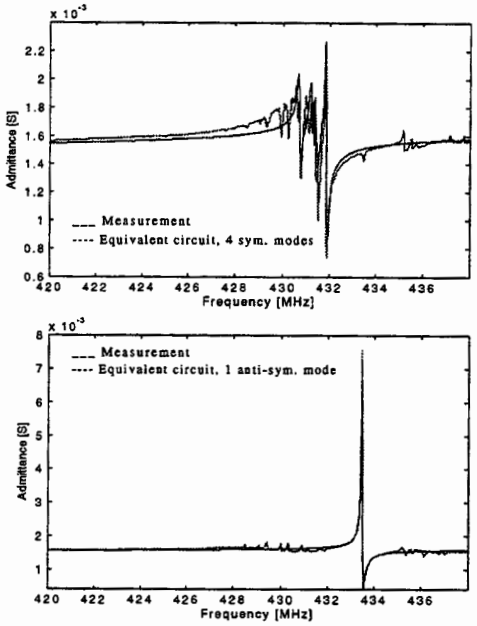


Fig.8: Comparison between measured and equivalent circuit admittances, (a) symmetric mode, (b) anti-symmetric mode.

Parasitic elements (a)			
In/out resistance [ohms]	Resistance to ground [ohms]	Bridge capacitance [pF]	Transducer capacitance [pF]
R_b	R_g	C_b	C_t
7.75743	0.13568	0.00485	1.16148

Motional elements, anti-symmetric mode (b)		
Resistance [ohms]	Inductance	Capacitance [fF]
R_a	L_a	C_a
122.481439	654.668787	0.205906

	Motional elements , symmetric modes (c)		
	Resistance [ohms]	Inductance [μH]	Capacitance [fF]
	R_s	L_s	C_s
1 sym. mode	708.347306	1490.903567	0.091095
2 sym.modes	1470.96774	2443.811094	0.055690
3 sym.modes	2146.70000	3443.528021	0.039555
4 sym.modes	1553.30000	2705.704381	0.050463

Table2: Values of Parasitics (a) and Motional (b), (c) elements

The five mode model leads to a much better description. The values of the motional and the electrostatic parasitic elements for this device are given in Table 2(a), (b) and (c).

The comparisons for both devices are done based on the full Y matrix. In both cases the good agreement between measured and equivalent circuit performances is shown.

6. CONCLUSION

An unique equivalent circuit configuration is applied to two port multiple mode coupled resonator filters, and an accurate parameter extraction procedure is established for both the motional elements and the basic electrostatic parasitic elements in the circuit. The comparison between the theory and the experiment shows a good agreement. This circuit configuration is readily compiled into widely used CAD tools for electronic circuit design and it will be very helpful for system design. From SAW filter analysis and design viewpoints, this approach provides means by which the mode behavior can be identified more clearly.

References

[1] P.S.Cross and R.V.Schmidt, "Coupled Surface-Acoustic_Wave Resonators", The BELL System Technical Journal, Oct. 1977, pp. 1446-1482.

[2] M.Tanaka, T.Morita K.Ono and Y.Nakazawa, "Narrow Bandpass Filter Using Double-Mode SAW Resonators on Quartz", 38th Annual Frequency Control Symposium, 1984, pp.286-293.

[3] P.V.Wright, "A New Generalized Modeling of SAW Transducers and Gratings", 43rd Annu. Symp. Freq. Contr. Proc., (1989), pp. 596-603.

[4] C.S.Hartmann, D.P.Chen and J.Heighway, "Modeling of SAW Transversely Coupled Resonator Filters Using Coupling-Of-Modes Modeling Technique", 1992 Ultrasonics Symp. Proc. pp. 39-43.

[5] G.Tobolka, "Mixed Matrix Representation of SAW Transducers", IEEE Trans. on Sonics and Ultrsonics, Vol. 26, No.3, 1979.

DETERMINATION OF THE EQUIVALENT CIRCUIT PARAMETERS OF TRANSVERSELY COUPLED SAW RESONATOR FILTERS

Jörg Beier* and Bernd Neubig**

*Technical University of Ilmenau, D-98684 Ilmenau, Germany
**TELE QUARZ Group, D-74924 Neckarbischofsheim, Germany

Abstract - Transversely Coupled SAW resonator Filters (TCF's) are usually described by the transmission characteristics (S21) at a given termination impedance. The element values of their equivalent electrical circuit can not be derived directly from these S-parameter-data. But they are important for the applications of TCF's in Voltage Controlled SAW Oscillators (VCSO's). Because of the fundamental similarity of two-pole TCF's to bulk wave Monolithic two-pole Crystal Filters (MCF's), the equivalent circuit of an MCF can be used for describing the TCF. The parameters for this equivalent circuit can be determined in different ways, e.g. from measurements of the complex impedance at the symmetric and the antisymmetric frequency of the SAW device, as shown in [1], or from S11-data.
This paper compares the two measurement procedures and presents experimental results showing a good correlation between the observed filter performance and the results computed from the equivalent circuit parameters obtained with these methods.

INTRODUCTION

In the last few years Transversely Coupled SAW resonator Filters (TCF's) have found a wide field of application. They were first presented by TIERSTEN and SMYTHE in 1975 [2]. This filter type offers a small bandwidth, very low insertion loss across the passband, steep passband skirts, high stopband rejection and a small size. Therefore they are used for Voltage Controlled SAW Oscillators (VCSO's) [3] and in communication systems [4]. The basic configuration of a TCF is shown in Fig.1. This two-pole filter consists of two identical SAW resonators, which are in close longitudinal proximity to each other (see Fig.1). HARTMANN et al. [5] presented a Coupling-Of-Modes theory and an equivalent electrical circuit diagram for TCF units. Fig.2 shows this electrical circuit. Y_A and Y_B are two serial resonant circuits.
A TCF unit is usually described by its transmission characteristic (S21) at a defined termination impedance. The element values of the equivalent electrical circuit can not be calculated from these S21-data directly.
It is possible to use a numerical optimization method to get the element values from the S-parameter data and a good fitting between the measured filter characteristic and the computed characteristic of the equivalent circuit.

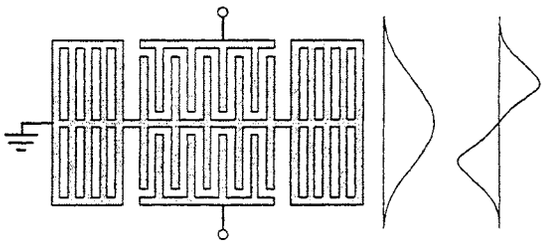


Fig.1: Basic configuration of a Transversely Coupled SAW resonator filter

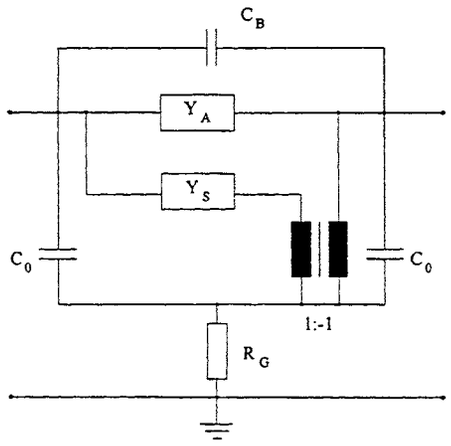


Fig.2: Equivalent electrical circuit diagram of a TCF (HARTMANN et al. [5])

But in the most cases it is only possible to get a very good correlation for one part of the S21-characteristic, either magnitude or phase. The fitting method is complicated, time consuming and has a high risk of ill-conditioned results due to the large number of interdependent variables. Therefore it is a problem to calculate the element values with help of a fitting method.

THEORETICAL BACKGROUND

There is a fundamental similarity between a two-pole TCF and an MCF, also called Dual. The idea is to use the well known equivalent electrical circuit of an MCF for characterizing the TCF.
Fig.3 shows one possible equivalent circuit for MCF's, using an inducting coupling network.

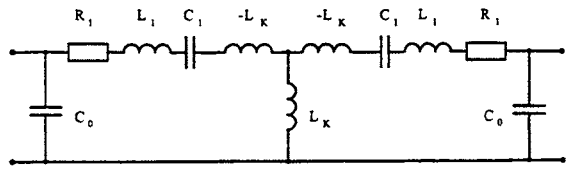


Fig.3: Equivalent electrical circuit of an MCF with inducting coupling network

Usually the equivalent electrical circuit of this passive two-port device has a symmetrical structure. In this case it can be transformed into the equivalent structure of a X-network. This transformation can be done in several steps [1] using the bisection theorem of BARTLETT [6]. The transformation provides an electrical circuit which is identical to the equivalent circuit of HARTMANN et al. [5] but without R_g and C_B. This circuit is shown in Fig.4 .

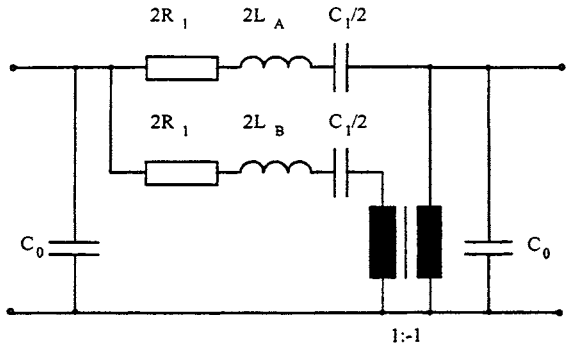


Fig.4: Structure of the equivalent electrical circuit after finishing the transformation (with $L_A=L_1-L_K$ and $L_B=L_1+L_K$)

This illustrates the similarity between the two filter types and the possibility to use an MCF-like equivalent electrical circuit for describing a TCF unit. For most VCSO applications the neglecting of R_G and C_B can be accepted because the VCSO characteristics are mainly determined by the passband parameter.
The quality of the equivalent circuit can be improved by introduction of a bridge-capacitance C_B as shown in Fig.5 (HARTMANN et al. [5]).

MEASUREMENT PROCEDURE

There are different ways to calculate the parameters of the MCF-like equivalent electrical circuit of a TCF.
The two procedures used here, are well known from measurements of quartz crystal units (for more details see [9]). In both cases the static capacitance C_0 has to be determined separately, the admittance G_0 is neglected.
The basic measurement sequence is shown in Fig.6 .

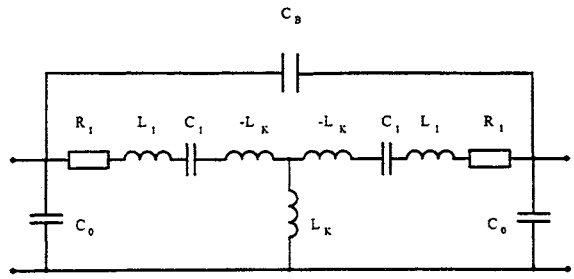


Fig.5: MCF-like equivalent circuit, extended by C_B

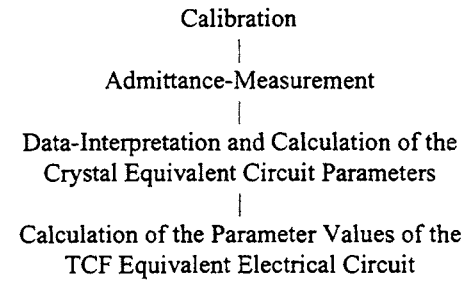


Fig.6: Basic measurement sequence of the two procedures

The admittance measurement was done with the "Direct amplitude/phase transmission method" as the first procedure and the "Single-port reflection method" as the second procedure.
The "Two-point iterative method" was used to calculate the crystal circuit parameters. For further methods see Standard IEC 444-5 [9].

First Procedure

Filter units like MCF's or TCF's with two coupled resonators have two characteristic modes at two typical frequencies: the symmetric frequency f_{sym} and the antisymmetric frequency f_{asym} . These modes can be activated if the two resonators are parallel-connected in the first case and serially-connected in the second case. Fig.7 shows both cases for a TCF unit [7].
The circuit with the connected resonators works as a one-port device. In both modes this one-port device behaves like a single resonator with a resonance frequency f_{sym} and f_{asym} respectively and different values of the parameters of the equivalent electrical circuit, the serial resonant network.

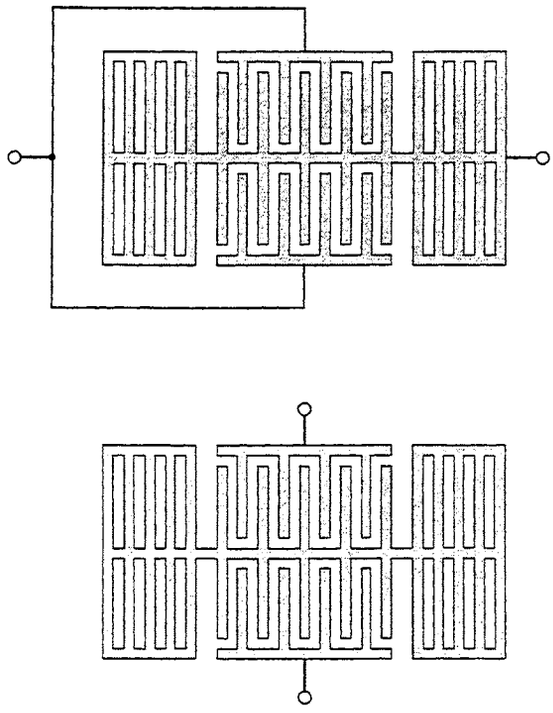


Fig.7: Symmetrical and antisymmetrical measurement mode of a TCF unit

In the symmetrical mode, the frequency and the complex impedance (C^*, R^*) have to be measured, for the antisymmetrical mode only the frequency f_{asym} has to be determined.
For frequencies up to approx. 500 MHz this can be done by conventional amplitude and phase measurements in a transmission system using e.g. the Pi-network system with error correction according to IEC 444-5, as it is commonly used for bulk wave crystal measurements [8], [9].

Second Procedure

In a first step the symmetrical mode is used to determine f_{sym} and the the value of S11. With help of equation (1) the admittance Y can be calculated from the measured S11-value:

$$R_0Y = (1\text{-}S11)/(1+S11),\ R_0 = 50\ \text{OHM} \qquad (1)$$

As a second step the freqency f_{asym} has to be measured. This can be done using the antisymmetrical mode of the TCF or with help of a measurement with a shorted filter output (see Fig.7 or 8).
The crystal circuit parameters are obtained by using the "Two-point iterative method" [8], [9].
The frequencies f_{sym} and f_{asym}, the element values R^*, C^*, and the static capacitance C_0 are the results of both measurement procedures.

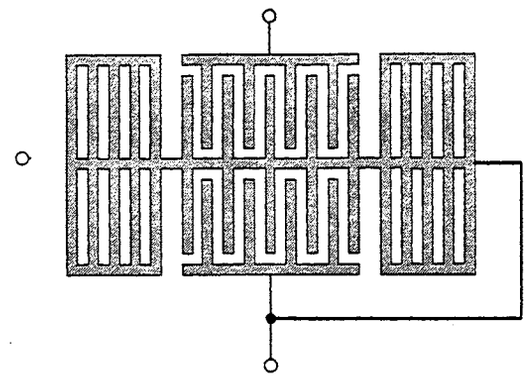


Fig.8: TCF-unit with shorted output

With help of the equations (2)-(7) the element values of the MCF-like equivalent electrical circuit can be calculated as:

$$C_1 = C^*/2 \qquad (2)$$

$$R_1 = 2R^* \qquad (3)$$

$$f_m = (f_{sym}+f_{asym})/2 \qquad (4)$$

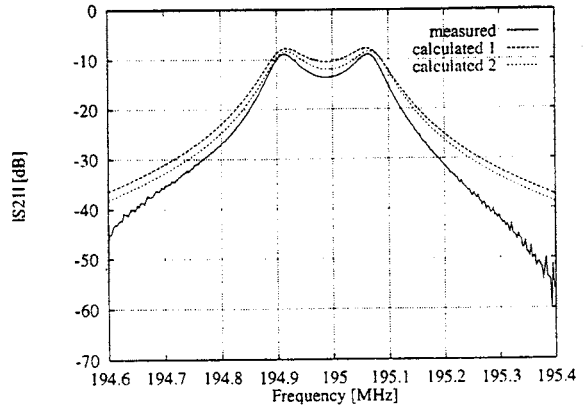
$$k = (f_{asym}-f_{sym})/f_m \qquad (5)$$

$$L_1 = 1/(4\pi^2 C_1 f_m^2) \qquad (6)$$

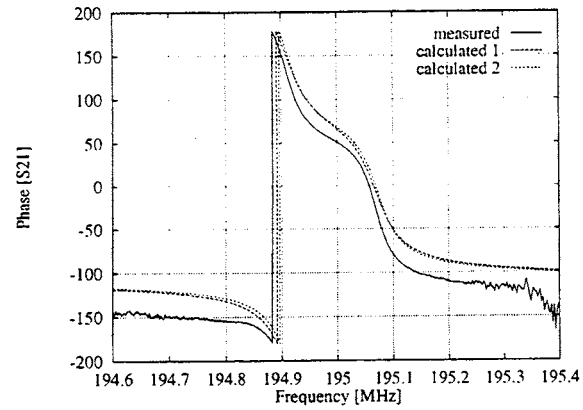
$$L_K = k/(4\pi^2 C_1 f_{sym}^2) \qquad (7)$$

EXPERIMENTAL RESULTS

The measurements using the first procedure were done with help of a network analyzer HP 8751A in combination with a Pi-network-system with error correction according to standard IEC 444-5, normally used for bulk wave crystal measurements. The TCF unit was connected to the Pi-network through an adapter element.
For the measurements using the second procedure a network analyzer system was used, which consists of ananalyzer unit HP 8510B, a synthesized sweep generator HP 83651A and an S-parameter test set HP 8514B. A special test fixture, based on an SMA-interface, was used for adapting the filter units to the S-parameter test set. Test fixtures for very high precision measurements (with N-connector and APC7- or APC3,5-interface) are shown in the standard IEC 444-5 [9].
The center frequency of the used TCF-units were 195 MHz and 220MHz. The following figures show the S21-characteristics computed from the determined element values of the TCF equivalent electrical circuit in comparision to the measured S21 response. All graphs under the title "calculated 1" are based on the first measurement procedure and all graphs under the title "calculated 2" are based on the second measurement procedure.

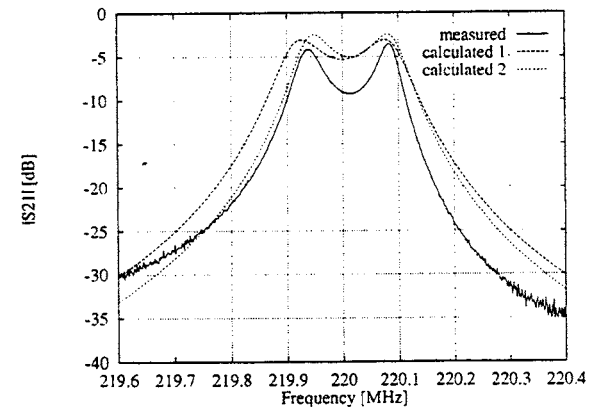


(a)

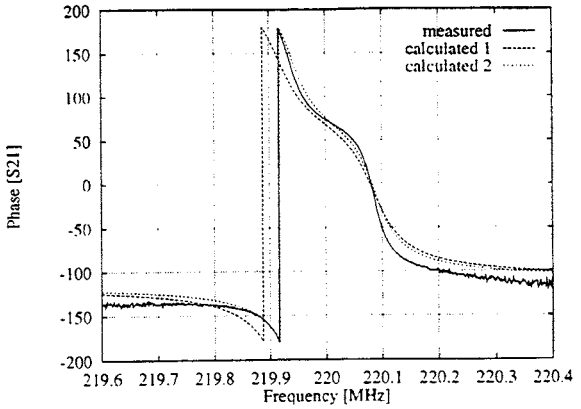


(b)

Fig.9: Comparision of the measured and the computed S21-characteristic of the TCF 195MHz
(a) magnitude
(b) phase

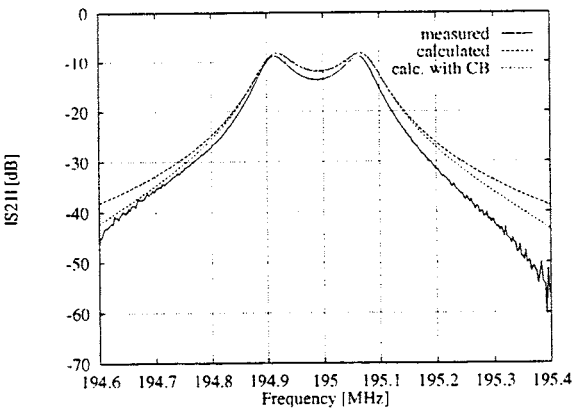


(a)

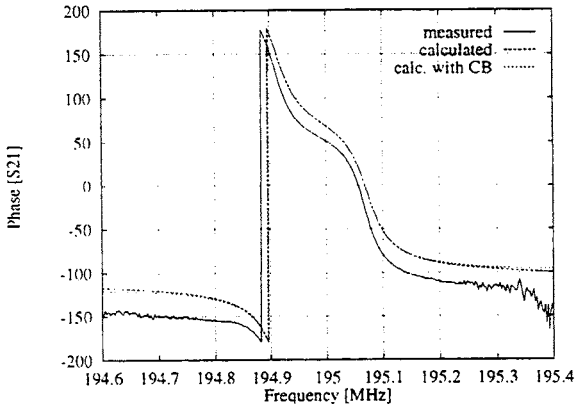


(b)

Fig.10: Comparision of the measured and the computed S21-characteristic of the TCF 220MHz
(a) magnitude
(b) phase



(a)



(b)

Fig.11: S21-characteristic of the TCF 195 MHz measured, calc. without C_B and with C_B=0,04pF
(a) magnitude
(b) phase

The quality of the equivalent circuit can be improved by introduction of a bridge-capacitance C_B (see [5]) as shown in Fig.5 .
Fig.11 shows the S21-characteristic of the TCF 195MHz without the additional C_B and with C_B = 0,04pF . The second measurement procedure was used.

DISCUSSION

A good correlation between the measured and the computed S21-characteristics in magnitude and phase in the passband area is observed. The results, obtained with help of the S-parameter measurement procedure, are more accurate. The quality of the MCF-like equivalent circuit for TCF units can be improved by introduction of a bridge-capacitance C_B. But the determination of the accurate value of C_B is difficult. Differences between the measured and the computed characteristics are mainly due to the non-ideal adapter between the Pi-network reference plane (used for calibration) and the SAW device if the first measurement procedure is used and the non-ideal test fixture if the second measurement procedure is used.
Undesired modes present in the SAW device are also not considered in the MCF-like equivalent circuit. The presented equivalent electrical circuit for two-pole TCF units does not work in the stopband area.

SUMMARY

The presented method allows the fast and simple determination of the element values of the equivalent electrical circuit of two-pole TCF units, which is very useful for circuit simulations, e.g. in oscillator applications (VCSO's) and other applications with main attention in the passband area. The paper describes the theoretical background of this method, compares two measurement procedures and presents experimental results. These results show a good correlation in the passband area between the observed filter performance and the results computed from the equivalent circuit parameters obtained with this method.

ACKNOWLEDGEMENT

The authors are grateful to Mr R. Zimmermann (Tele Quarz Group) and Mr U. Stehr (Technical University of Ilmenau) for their support of this work.

REFERENCES

[1] J. Beier, B. Neubig, "Measurement Of the Equivalent Circuit Parameters Of Transversely Coupled SAW Resonator Filters", *IEEE Ultrasonics Symposium*, Cannes, France, 1994, Paper CC-6

[2] H.F. Tiersten, R.C. Smythe, "Guided Acoustic Surface Wave Filters", in *IEEE Ultrasonics Symposium Proceedings* 1975, pp. 293-294

[3] U. Stehr, J. Beier, "S-Parameter-Simulation eines 195 MHz-VCSO mit TCF", Research Report, Technical University of Ilmenau, 1994

[4] C.C.W. Ruppel et al., "SAW Devices for Consumer Communication Applications", *IEEE Trans. UFFC*, vol.40, No.5, September 1993, pp. 438-452

[5] C.S. Hartmann et al., "Modeling of SAW Transversely Coupled Resonator Filters Using Coupling-Of-Modes Modeling Technique", in *IEEE Ultrasonics Symposium Proceedings* 1992, pp. 39-42

[6] *Akademischer Verein Hütte*, Ingenieur-Taschenbuch Band IV B Fernmeldetechnik. Berlin und München: Verlag Wilhelm Ernst & Sohn, 1962

[7] B. Neubig, "Monolithische Quarzfilter", *FUNKSCHAU*, Heft 10, 1978

[8] B. Neubig, "Fortschritte auf dem Gebiet der Schwingquarz-Meßtechnik bis 500 MHz", in *Proceedings of the ZVEI-Symposium "Schwingquarze"* 1989, pp. 106-126

[9] Standard IEC 444-5, First Edition, 1994

Accurate Analysis of Pseudo-SAW devices*

P. Ventura, J. M. Hodé, M. Solal, and L. Chommeloux

Thomson Microsonics, 399 Route des Crêtes, BP 232, 06904 Sophia-Antipolis Cedex

*This work was partially supported by DRET

Abstract

The analysis of Pseudo-SAW low-loss bandpass filters is much more complex than Rayleigh-SAW filters. In the upper part of the stopband the physical properties of Pseudo-SAW type waves are greatly perturbed by bulk acoustic wave radiation. In this paper, an analysis method based on a rigorous FEM / BEM numerical model [1] is presented. All electric and acoustics mode contributions as well as mass loading effects are taken into account, resulting in very accurate simulations. Comparisons between simulations and measurements for one port SAW resonators on 36° Y rotated lithium tantalate are presented, showing excellent agreement.

I - Introduction

The need of wide bandwidth low-loss filters lead SAW designers to employ resonator filters using Pseudo Surface Acoustic Waves (PSAW) on high electro-mechanically coupled substrates such as the 41°, 64° rotated Y-cut, X propagation LiNbO3 and the 36° rotated Y-cut, X propagation LiTaO3. Furthermore, in order to increase the mechanical part of strip reflectivity high metal thickness is often used [2].
Due to the presence of a close Surface Skimming Bulk Wave (the slow shear), the physical nature of a PSAW is very complex. As was pointed out by Plessky [3], the localization depth of its spatial harmonics in a periodic grating is very sensitive with frequency. Mechanical and electrical boundary conditions also create great perturbations of the electro-mechanical behavior of PSAW. Classical COM or P-matrix models used for Rayleigh-type SAW devices cannot be applied to accurately simulate PSAW devices. From infinite periodic grating analysis, it is possible to develop COM or P-matrix like models which are used to simulate finite grating devices. This approach [4-5] consists of introducing new parameters in the classical COM model (dispersion, attenuation due to bulk wave radiation).

Last year [1], a 2D mixed FEM (Finite Element Method) / BEM (Boundary Element Model) was presented to compute the dispersion curve and the coupling efficiency of PSAW infinite periodic transducer. Even if infinite gratings are impossible to realize physically, they are very close to synchronous resonators if reflective gratings are long enough. In that case, there is no need to derive approximate COM parameters and an exact analysis is possible. In this paper we use a numerical mixed FEM / BEM model ([1] and [6]) to accurately simulate the admittance of synchronous resonators : all the electrical and acoustical contributions will be taken into account as well as the thickness and the shape of the electrodes. This new analysis tool can be directly used to compute the transfer function of ladder type resonator filters based on synchronous PSAW resonators with only the 2D approximation.

II - 2D analysis of infinite periodic transducers

Last year [1], we presented a 2D mixed FEM / BEM model of an infinite periodic array of metallic electrodes based on the harmonic admittance concept [7]. Figure 1 represents the geometry of the problem : the interface of a semi-infinite piezoelectric medium is covered by a p-periodic array of metallic electrodes with infinite length and other dimensions finite.

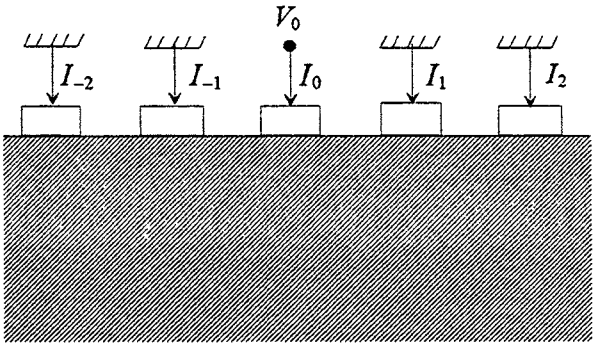


Figure 1 : schematic of the problem

Let I_n be the electrical current per unit length entering in the n^{th} electrode at the potential V_n. The purpose of the analysis is to compute mutual admittances Y_n defined as the ratio I_n/V_0 for the elementary potential excitation of the array (figure 1). As was explained in [1] and [7], the harmonic admittance method consists of

solving the dual problem of the array driven by an harmonic potential excitation (figure 2) :

$$V_n(\gamma) = V_0 \exp(-j2\pi n\gamma) \qquad (1)$$

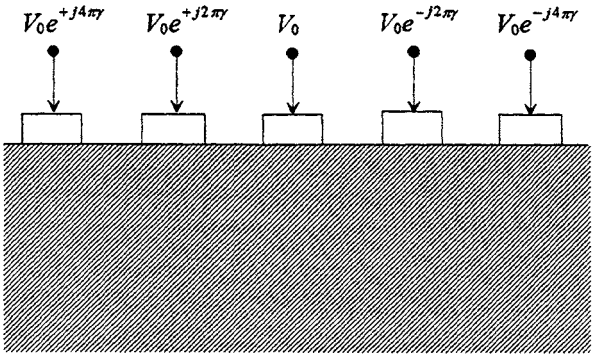


Figure 2 : harmonic excitation of the array

According to Floquet's theorem, the electrical current $I_n(\gamma)$ can also be written :

$$I_n(\gamma) = I_0(\gamma)\exp(-j2\pi n\gamma) \qquad (2)$$

The quotient $I_n(\gamma)/V_n(\gamma)$ doesn't depend on the electrode number and is called the harmonic admittance $Y(\gamma)$ of the array. The mutual admittance Y_m for the elementary excitation of the array is then obtained by Fourier transform of $Y(\gamma)$:

$$Y_n = \int_0^1 Y(\gamma)\exp(-j2\pi n\gamma)d\gamma \qquad (3)$$

By using the periodic harmonic Green's function $\mathbf{G}_\gamma^p(x)$ and a FEM discretization of the metallic electrode, we succeeded [1] in accurately computing $Y(\gamma)$ for any shape of electrode while taking into account all the acoustical (Rayleigh-type SAW, PSAW, Bulk) and electrical contributions. The periodic harmonic Green's function is the response of the piezoelectric substrate to a linear combination of Dirac ($\sum_{n=-\infty}^{+\infty}\delta(x-np)\exp(-j2\pi n\gamma)$) and allows only single period modelization. In the case of a synchronous resonator, it is only necessary to take into account one electrode. It is important to mention that in the case of rectangular isotropic metallic electrodes and of pure transverse shear polarization (which represents a good approximation for PSAW on Y+36° LiTaO3), it is not necessary to use the Finite Element Method : relationships between stresses and displacements at the electrode-substrate interface can be analytically derived from the expansion of the displacement field inside the electrode into SH modes of free isotropic plates.

Figure 3 is an example of a harmonic admittance curve for rectangular aluminum electrodes. In that figure, we can observe four different contributions :

$$Y(\gamma) = Y_{PS}(\gamma) + Y_{RS}(\gamma) + Y_B(\gamma) + Y_D(\gamma) \qquad (4)$$

where $Y_{PS}(\gamma)$, $Y_{RS}(\gamma)$ are the PSAW and RSAW contributions, $Y_B(\gamma)$ contains the acoustical coupling due to bulk modes, and, $Y_D(\gamma)$ corresponds to the electrostatic contribution. $Y_{PS}(\gamma)$ and $Y_{RS}(\gamma)$ give the complex and real pole behaviour of the harmonic admittance resulting in very fast variations in the neighbourhood of $\gamma_{PS} = fps_{PS}$ and $\gamma_{RS} = fps_{RS}$ where s_{PS} and s_{RS} are respectively PSAW and RSAW slownesses for short-circuited electrodes.

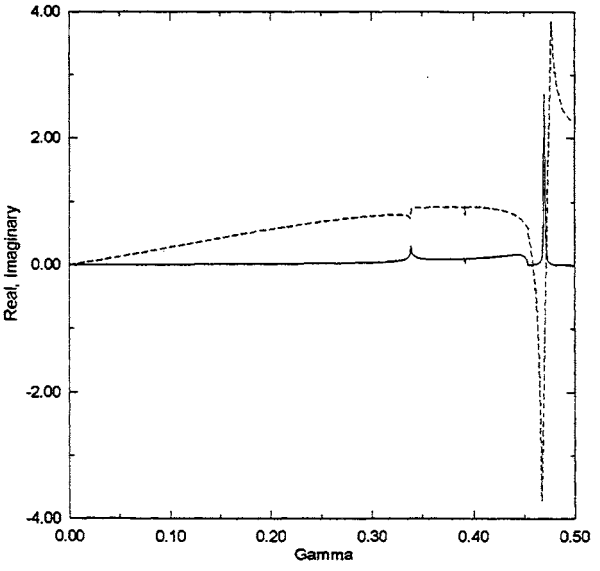


Figure 3 : real (solid line) and imaginary (dashed line) of the harmonic admittance : (width/period : 0.58, height/λ : 6.9% and product frequency period of 1893.14).

When removing PSAW and SAW contributions (see III), the resulting function of γ is much smoother with only derivative discontinuities caused by Surface Skimming Bulk Waves (SSBW) while the electrostatic coupling between electrodes gives the $\sin(\pi\gamma)^2$ shape. Figure 4 is the residual admittance harmonic curve without PSAW contribution.

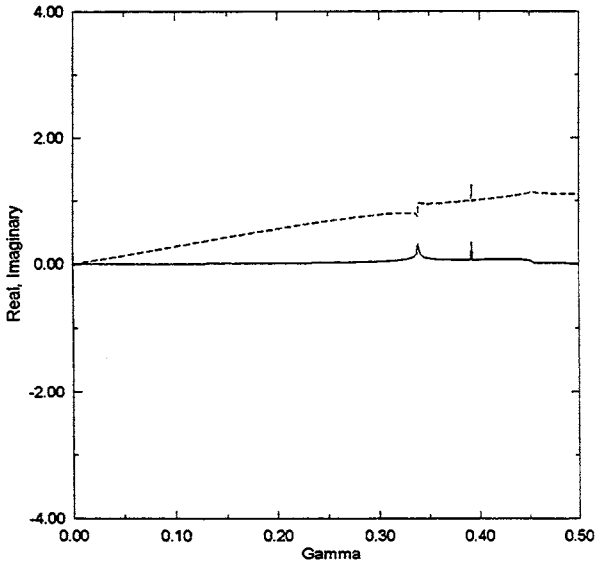


Figure 4 : real (solid line) and imaginary (dashed line) of the harmonic admittance without PSAW contribution

III - PSAW contribution and dispersion curve

Due to the very fast variations of $Y_{PS}(\gamma)$, (3) can not be computed numerically. Therefore, it is necessary to find an analytical expression of that function. We used Zhang et al's model [7] and assumed that the PSAW contribution to Y_n can be written :

$$Y_n^{PS} = \begin{cases} Y_{sc}\exp(-j|m|\varphi_{sc}) & \text{for } n \neq 0 \\ Y_{sc}\left(1+j/\tan(\varphi_{sc}/2)\right) & \text{for } n = 0 \end{cases} \quad (5)$$

By computing the Fourier serie of Y_n^{PS} we obtained an analytical expression for $Y_{PS}(\gamma)$:

$$Y_{PS}(\gamma) = \frac{jY_{sc}}{\tan(\varphi_{sc}/2)} \frac{\sin(\pi\gamma)^2}{\sin(\pi\gamma)^2 - \sin(\varphi_{sc}/2)^2} \quad (6)$$

Therefore, $Y_{PS}(\gamma)$ depends on two parameters φ_{sc} and Y_{sc}. φ_{sc} is the phase propagation of the short circuited eigenmode while Y_{sc} characterizes the amplitude of the SAW coupling between two electrodes.

These two parameters are computed using an iterative algorithm. Figure 5 shows the dispersion curve numerically calculated for shorted electrodes and figure 6 represents the amplitude of the PSAW coupling in the array (real and imaginary part of Y_{sc}). The piezoelectric substrate is 36° Y rotated X propagation LiNbO3. The aluminum electrodes of the array are assumed to be rectangular (width : 1.04 µm, height : 3100 Å) and the period is 2.084 µm. In the 980 MHz to 1000 MHz frequency range corresponding to the end of the stopband, due to the presence of close SSBW perturbation and smooth behaviour of PSAW contribution, it was not possible to extract PSAW parameters with the present algorithm : it corresponds to the hached part of figures 5 and 6. The only way to solve this problem will be to find a model of the SSBW contribution to the harmonic admittance and then to introduce it in the iterative algorithm.

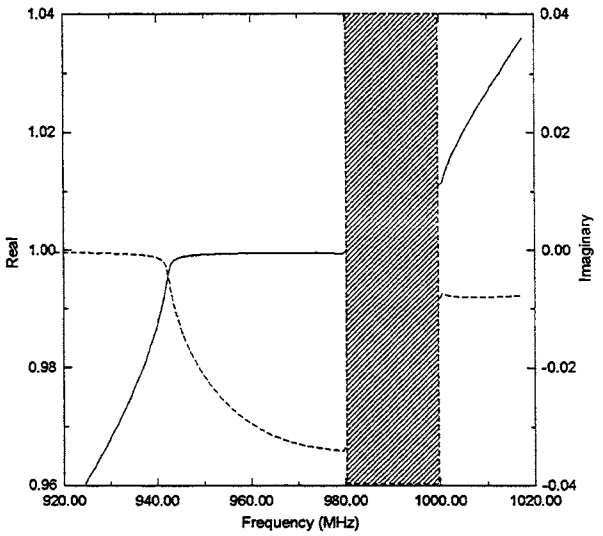


Figure 5 : Real (solid line) and Imaginary part (dashed lines) of φ_{sc}/π

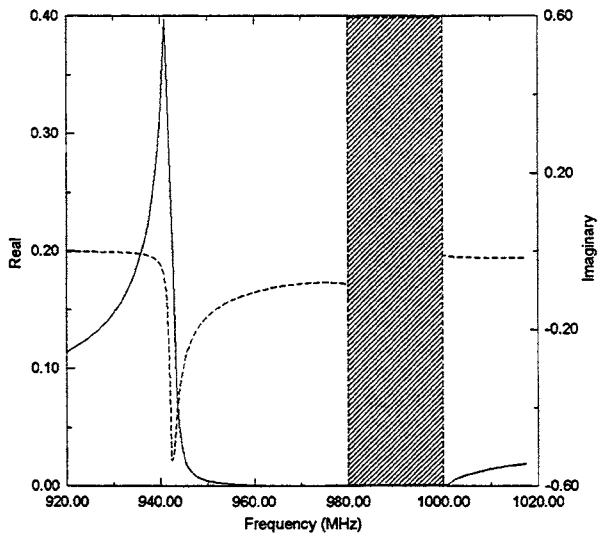


Figure 6 : Real (solid line) and Imaginary part (dashed lines) of Y_{sc}

IV - Application to the analysis of synchronous resonators

After having computed frequency dependence of both parameters φ_{sc} and Y_{sc} for each SAW, we are able to compute the PSAW and RSAW contributions to the mutual admittance Y_n between any couple of electrodes belonging to the infinite periodic array (see equation (5)).

The residual contribution Y_n^r to the mutual admittance Y_n contains acoustic bulk and dielectric effects :

$$Y_n^r = Y_n - Y_n^{PS} - Y_n^{RS} \tag{7}$$

Y_n^r is obtained by computing the Fourier transform of the residual harmonic admittance

$$Y_r(\gamma) = Y(\gamma) - Y_{PS}(\gamma) - Y_{RS}(\gamma) \tag{8}$$

$$Y_n^r = \int_0^1 Y_r(\gamma)\exp(-j2\pi n\gamma)d\gamma \tag{9}$$

Due to the much smoother nature of the residual harmonic admittance, we decided to compute this finite bounded integral by using a classical FFT. In the 980 MHz to 1000 MHz frequency range it was not possible to extract PSAW parameters (see section III). Fortunately, in that frequency range, due to stopband effects PSAW contribution can also be computed by FFT.

We are now able to compute the admittance matrix of any synchronous PSAW resonator. Let us focus on a one electrical port synchronous PSAW resonator (Figure 7).

Figure 7 : one port synchronous resonator

If N_t is the number of electrodes pairs in the active part of the resonator, the admittance Y of the resonator can be written :

$$Y = \sum_{i=1}^{N_t}\sum_{j=1}^{N_t} Y_{2|i-j|} \tag{10}$$

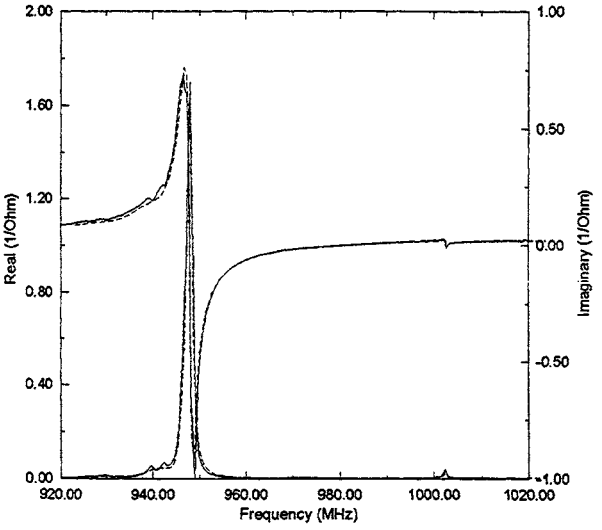


Figure 8 : Real and imaginary part of the admittance (simulation with velocity correction : dashed line, measure : solid line)

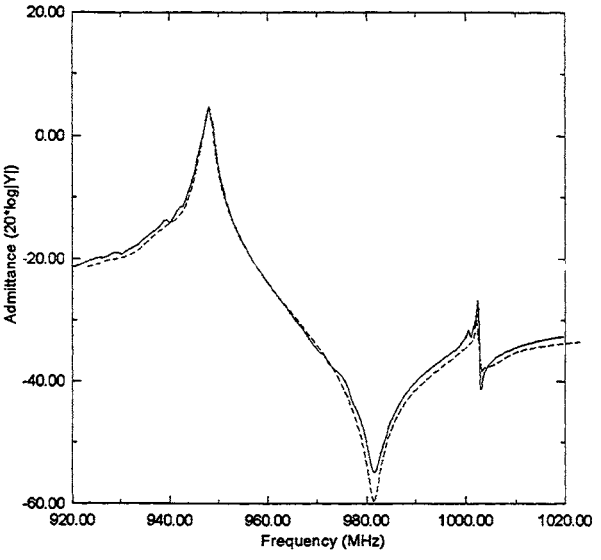


Figure 9 : admittance (20log|Y|) (simulation with velocity correction : dashed line, measure : solid line)

Figure 8 shows the comparison between simulation and experiment for a one port synchronous resonator on Y+36° LiTaO3. The number of electrode pair of the transducer is 60 and each reflective array is 150 shorted electrodes long. The characteristics of the resonator (period, shape of electrodes) are still the same as in figures 5 and 6. In the simulation values, the velocity has been corrected by 0.995, and strip resistivity loss have been taken into account. The comparison with experience is very good especially for the Q factor and the stopband length. Due to the infinite periodic array model, small differences can be observed before the resonance peak.

In order to highlight the end of the stopband simulation, badly taken into account by classical COM models (because SSBW contribution is far from being negligible), the admittance of the resonator is also plotted in logarithmic scale (figure 9). In that frequency range, it exists a small difference between simulation and experience for the loss, but the shape and the discontinuity in the curve are very well modelled.

V - Conclusions

An original method has been presented to compute the admittance of one port synchronous resonator. Based on a mixed FEM / BEM numerical model and on the harmonic admittance concept, this method permits accurate computation of mutual admittances in a periodic array. Furthermore, no assumptions are made on the electrode geometry, and, the model is able to take into account the very complex behaviour of PSAW.
A comparison between simulation and experiment has been realized for a one port synchronous resonator, showing a very good agreement. This powerful numerical tool can also be used to derive parameters of P-matrix and COM models in order to analyse more complex structures.

References

[1] P.Ventura, J. Desbois and L. Boyer, "2D Analysis of SAW and PSAW Periodic Transducers Using FEM and BEM", *Proceeding of the EFTF*, 1994

[2] V. P. Plessky and C. S. Hartmann, "Characteristics of Leaky SAWs On 36-$LiTaO_3$ In Periodic Structures of Heavy Electrodes", *IEEE Ultrason. Symp. Proc.*, pp. 1239-1242, 1993.

[3] V. P. Plessky, "A Two Parameter Coupling-of-Modes Model For Shear Horizontal Type SAW Propagation in Periodic Grating", *IEEE Ultrason. Symp. Proc.*, pp. 195-200, 1993.

[4] M. Solal and P. Ventura, "Analysis and Design of a very Low Loss SAW Resonator Filter on 36° Lithium Tantalate", pp. 614-623, *Proceeding of the EFTF*, 1994

[5] V. P. Plessky, D. P. Chen and C. S. Hartmann, " 'Patch' Improvements to COM Model For Leaky Waves", *to be published in 1994 IEEE Ultrason. Symp. Proc.*

[6] H.P. Reichinger and A.R. Baghai-Wadji, "Dynamic 2D Analysis of SAW-Devices Including Massloading", *IEEE Ultrason. Symp. Proc.*, pp. 7-10, 1992.

[7] Y. Zhang, J. Desbois and L. Boyer, "Characteristic Parameters of Surface Acoustic Waves in a Periodic Metal Grating on a Piezoelectric Substrate", *IEEE Trans. on Ultrason., Ferr. and Freq. Control*, Vol. 40, No 3, p. 183, 1993.

SESSION 10

TELECOM APPLICATIONS

Chairman : K. HILTI

Telecommunication Network Synchronization

Michael Wolf

Alcatel SEL

When the network is not properly synchronized, synchronously operating network elements cause corruption of transported data, or at least corruption of the phase stability of the data signal, which may cause bit errors. The requirements for a suitable network synchronization method are based on economical, reliability, robustness and structure objectives for the network.

The current switched network uses the extended connectivity of the 2 Mbit/s connections between the network nodes for synchronization. With SDH transport systems these connections loose their phase stability and their stability in terms of unchangeable connectivity. This new flexibility of network configuration is a major benefit for an optimized usage of the network resources, but it impacts the network synchronization stability. Therefore, future network synchronization is to be carried on optical line signals instead of 2 Mbit/s signals. The different connectivity at that network level of optical line signals (compared to the 2 Mbit/s level) requires an adaptation of the synchronization methods. Section 3 gives some suitable methods, which are discussed versus the basic requirements for synchronization. For optimizing the network, different methods may be combined.

1. Why Synchronization is Needed

The digital switching and multiplexing techniques for 64 kbit/s (Basic Rate) channels are based on synchronous operation requiring synchronization of the Primary Rate (2 Mbit/s) signals. The evolving systems of the Synchronous Digital Hierarchy (SDH) are defined for synchronous operation at the 155 MHz level.

Any synchronously operating network element employs a buffer for accommodation of phase fluctuations of the received data signal, called wander. But the techniques for controlling the buffer fill are different at the Basic Rate / Primary Rate and at the SDH level.

At the Primary Rate the buffer fill is accommodated by *controlled slips. A slip deletes or repeats an entire frame of 125 µs duration.* Since this corrupts the data, the rate of such processes should be limited to less than a few per day. The phase adaptation step of 125 µs and the implemented hysteresis of more than 18 µs allows for accommodation of a certain amount of phase noise without generation of slips and for accommodation of some frequency offset with a low slip rate.

In SDH the buffer fill is accommodated by *byte stuffing processes. A stuffing process causes transport of additional bytes per frame or less bytes per frame* according to a buffer overflow or underflow condition. Any particular payload is identified by a pointer value within the SDH frame. Because the stuffing process changes this pointer value, it is called a pointer justification. For payload rates above 2 Mbit/s the justification phase step is 160 ns and the hysteresis is 640 ns. The hysteresis is a factor of 28 lower than the hysteresis at the Primary Rate and that results in a much higher sensitivity to phase noise. The phase adaptation step of 160 ns is a factor of 781 lower than that at the Primary Rate which leads to a 781 times higher rate of buffer control actions in the case of a frequency offset.

However, the pointer justifications do not corrupt the data, they do only degrade the payload phase stability. Therefore a certain rate of pointer justifications is tolerable on a connection through the network. But excessive jitter or wander of the signal should be avoided by keeping the probability of multiple "simultaneous" pointer justifications at a very low value. Several pointer justifications appear to be "simultaneous", when they are separated by time intervals, which are below the time constant of some clock filter used for filtering the signal clock. "Simultaneous" pointer justifications therefore cause overlapping phase transients.

Derived from the technical aspect described above, the Primary Rate requires a long term phase stability of

$$\mathrm{MRTIE}(\tau) < 18\,000 \text{ ns with } \tau < 100\,000 \text{ s}$$

and SDH requires a short term stability specified as

$$\mathrm{rmsTIE}(\tau) < 300 \text{ ns with } \tau > 100 \text{ s}.$$

The long term stability can be achieved by using high stability clocks (e. g. Cesium or Rubidium) or by synchronizing the network in a master / slave approach. The short term stability is provided by a suitable clock design using digital PLL technique, phase stepper technique for avoiding phase hits or transients on switching redundant PLL modules and phase build–out for avoiding phase transients on input reference switching.

Further benefit from a good network synchronization comes from the possibility to derive radio frequencies from digital line signals with the required clock accuracy.

2. Requirements for Network Synchronization

Economics: For economical reasons a master / slave synchronization technique is used instead of high stability oscillators at every network node.

Reliability: Redundant synchronization paths are used between the network nodes. The probability, that a network node clock is operating in the holdover mode at some frequency offset is reduced by this.

Robustness: To make the network synchronization robust against any type of catastrophic impact it is better to have a synchronization architecture, which uses only locally available informations for selecting input reference signals. Such an approach does not rely on a network wide software control of the synchronization structure.

Structure: Large networks have internal structures. Such structures may be hierarchical and/or regional. The structure of clock distribution within the network should follow such structures.

Slaving principles: A basic requirement is to slave equal or lower quality clocks on a certain quality master clock. This principle ensures that the best holdover performance is achieved for the network under failure conditions.

However, the SDH technique inserts a chain of low quality clocks between two high quality clocks slaved to each other. In a failure case a high frequency drifting clock (the SDH Equipment Clock, SEC) synchronizes a much more stable clock. This can be avoided by switching off (squelching) the reference clock provided by the SDH system to the stable clock equipment. A timing marker, called Synchronization Status Message, is available in SDH for supporting that function. It is coded in the SDH frame overhead

Synchronization paths: Synchronization should be carried over the most stable signal paths. The requirement for stability is twofold: these paths should be stable in terms of routing and network configuration, and they should provide signals of high phase stability.

Such properties have 2 Mbit/s signals carried over PDH multiplex systems. But not, if they are carried over SDH systems, because these systems degrade both types of stability. Firstly, the configuration stability disappears because of flexible routing and secondly, the phase stability is degraded by phase noise generating pointer justifications.

In SDH systems the clock is transported by the optical line signals.

Radio links, independent from transmission network systems, can also be used for synchronization paths, e. g. GPS.

3. Network Synchronization Methods

Figure 1 shows the hierarchically structured connectivity between the switches at the 2 Mbit/s level. This structure was generated in order to have redundant transport capacity (two way routing). This can also be used for redundant synchronization paths.

The connections shown in figure 1 are carried within multiplex line signals, which are connecting adjacent nodes. The related connectivity of line systems is shown in figure 2(a). When generating a hierarchical synchronization structure from this, the following can be seen:

- The resulting structure is very complex and
- there is no redundant path to some of the nodes.

Therefore, some methods for improving the synchronization structure are required.

(1) The bold lines in figure 3 represent *direct transmission line systems to far distant nodes*, which pass some intermediate nodes without connecting them. With these lines the problem of the path redundancy can be solved. The same method can also be used for making the synchronization hierarchy simpler. This is shown in the example in figure 5. The drawback of this method is, that the network structure is determined very much by requirements coming from the synchronization, while the network operators want to derive their network structure from the traffic capacity demand, which means cost optimization for the network.

(2) In figure 4 a *ring synchronization method* is shown. This method is well known from SDH ring networks. A specific Synchronization Status Message (SSM) is used for reverting the synchronization direction after a link failure without the risk of creating a timing loop. This method, applicable to the SDH equipment clocks, can be upgraded for synchronizing network node clocks. It generates a nice synchronization structure

of the network, but cascading of many node clocks is necessary. Current standards limit such cascading to a number of ten for limiting the phase noise accumulation. However, the phase stability performance of modern node clock equipment exceeds the standard specifications by a factor of ten. This means, that cascading of much more clocks is possible without any impact on the phase stability performance. A network example of this method is shown in figure 6.

Further methods are available:

(3) *The network wide synchronization structure can be controlled by a network management system.* This allows for the most flexibility for managing the synchronization redundancy, but the method requires informations about the current status of all network elements. In large networks, such information handling is increasingly difficult to manage and it causes long delay. The potential need for switching many synchronization paths after a failure may cause high phase transients in the network by an accumulation effect.

(4) Completely in contrast to that is the following method: *Use improved stability oscillators* at the network nodes without using redundant synchronization paths. When the unique synchronization path fails, the clock provides sufficient frequency stability for autonomous operation for a long time. A network management system is only needed for supporting the maintenance by failure reporting. The SDH networks employ means for protection switching and re–routing functions, which are used for making the payload links highly available without the need for rapid maintenance actions after any failure. This is done for reducing the operation cost of the network. However, going without redundant synchronization paths, requires quick maintenance actions at least on synchronization failures. A further drawback of this method is the higher cost of high stability oscillators.

(5) An improvement and upgrade of the above method is the *use of radio systems*, such as GPS, which distribute timing information. This method gives a very high phase stability to the network. But the telecommunication network is no longer self supporting, because a third system is used for a basic part of its functions.

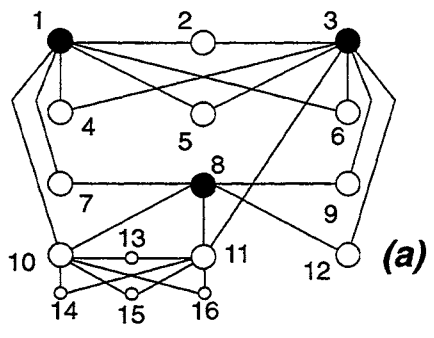
4. Conclusion

Telecommunication network synchronization is needed for minimizing the rate of frame slips and pointer justifications. This is required for maintaining a low error rate of the transmitted digital signals. SDH systems require an improved short term stability compared with the phase stability required for digital switching systems.

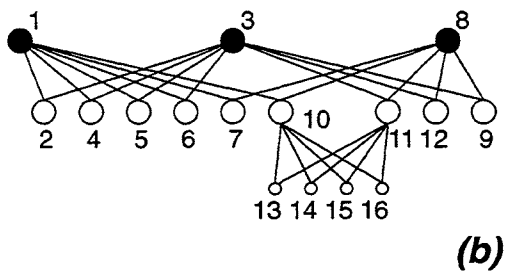
Concepts for network synchronization are based on economic considerations, reliability and robustness requirements and on a target network structure. Some network operators call for a self supporting transport network, which does not need a third system for clock distribution.

The connectivity and the routing principles of the switched 2 Mbit/s network facilitate a redundant synchronization scheme. However, flexible network configuration based on SDH systems destroys the ability for clock distribution at 2 Mbit/s. Using clock distribution over optical line signals cause some structure problems. To overcome that, some methods have been described: the direct line method, the ring synchronization method, the network management method, the improved oscillator method, and the methods using radio systems, e. g. GPS.

The network synchronization architecture may be optimized by combining the various methods according to the underlying network structures.

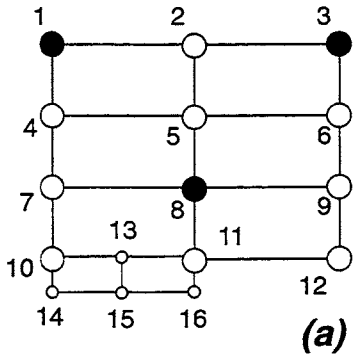


Idealized geographical locations of the network nodes

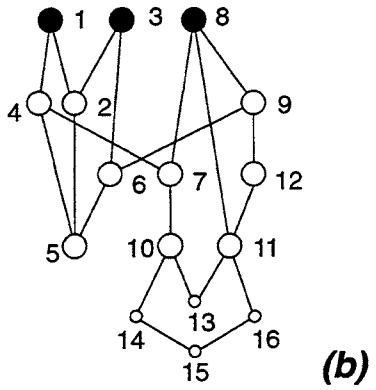


Hierarchical positions of the network nodes

Figure 1: Structure of the switched network with synchronization over 2 Mbit/s.

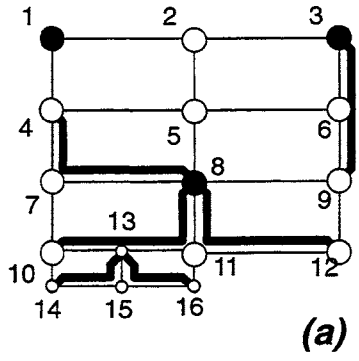


Idealized geographical locations of the network nodes

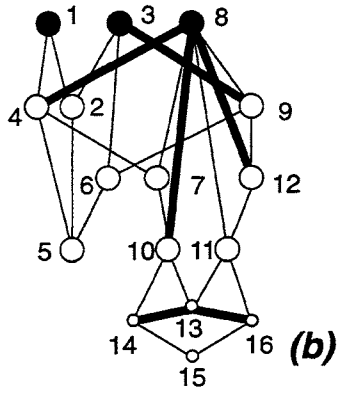


Hierarchical positions of the network nodes (for synchronization)

Figure 2: Structure of the transmission network with synchronization over optical line systems.

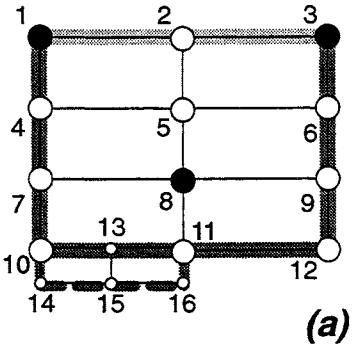


Idealized geographical locations of the network nodes

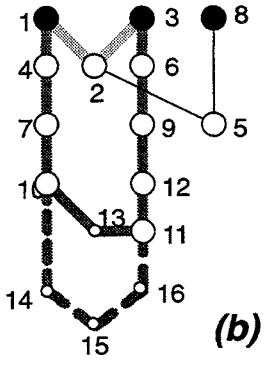


Hierarchical positions of the network nodes (for synchronization)

Figure 3: Direct transmission links for synchronization.



Idealized geographical locations of the network nodes



Hierarchical positions of the network nodes (for synchronization)

Figure 4: Principle of ring network synchronization.

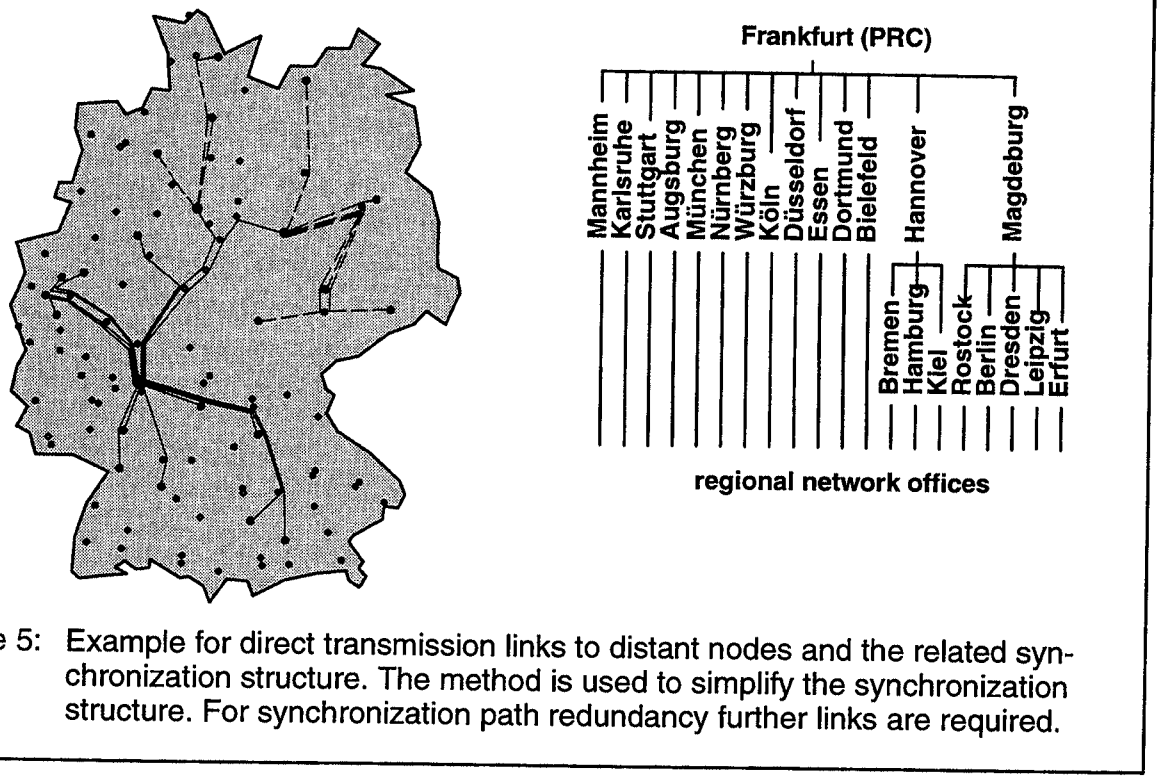


Figure 5: Example for direct transmission links to distant nodes and the related synchronization structure. The method is used to simplify the synchronization structure. For synchronization path redundancy further links are required.

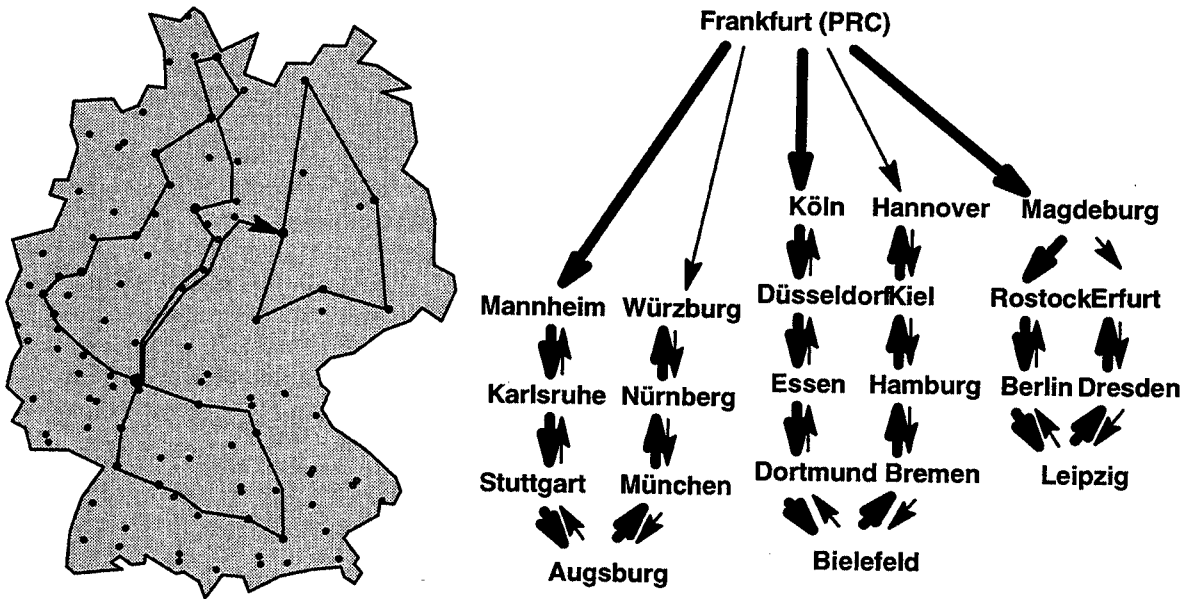


Figure 6: Example for ring synchronization combined with a direct transmission link for synchronizing the north eastern ring via Magdeburg.

Telecommunications Network Synchronisation and Wander Generated by SDH Equipment

Ralph Urbansky

Philips Kommunikations Industrie AG
Nuremberg

Abstract

The Synchronous Digital Hierarchy (SDH) is the new international transmission standard and provides for enhanced flexibility in the network with the introduction of switching capabilities. The introduction of an SDH-based synchronisation concept is one means of reducing the impact on jitter and wander associated with frame structures, which are optimised with respect to switching capabilities.

In this paper the wander introduced by an SDH data link will be discussed. It will be shown that the wander is reduced significantly by application of the Adaptive Threshold Modulation technique for waiting time jitter reduction in conjunction with suitable antialiasing filtering for frame structure interference suppression. Simulation and measurement results demonstrate the feasibility of this approach.

1 Introduction

The Synchronous Digital Hierarchy (SDH) is the new international transmission standard and will replace the existing transmission infrastructure based on the Plesiochronous Digital Hierarchy (PDH). To allow for a smooth transition from PDH to SDH networks, PDH services are also supported. This implies that the SDH has to provide full compatibility with the PDH with respect to the relevant standards.

The main advantage of SDH networks is the enhanced flexibility by the introduction of switching capabilities. The switching capabilities are provided by certain frame formats and rates, which are almost synchronous to the payload signals to be switched. The SDH employs a synchronisation method combining the techniques known from digital exchanges and PDH equipment. Therefore, the SDH network synchronisation quality has an impact on the jitter and wander performance of digital connections, which may result in an increased probability of bit errors or frame slips respectively.

In addition to these inherent mechanisms, the implementation of the relevant components may contribute to a degradation or an improvement of the jitter and wander performance. Section 2 offers some background information on network synchronisation requirements with respect to SDH. Section 3 analyses the components which have a major impact on the phase transfer characteristic. In Section 4 simulation results of certain implementations will be presented and compared with the corresponding measurement results.

2 Network Synchronisation

The telecommunications infrastructure currently installed consists basically of digital exchanges (switches) and transmission links of the Plesiochronous Digital Hierarchy (PDH) providing point-to-point connections between these exchanges.

Digital exchanges commonly switch 64-kbit/s telephone channels. The efficient implementation of a switching fabric requires a synchronous multiplexed frame format, i.e. the channels occupy fixed time slots in the frame of the multiplexed signal. The switch fabric consists of a time slot interchanger, which writes the data into a buffer and reads it in the desired order. Before the signal is fed into the switch fabric, the signal frames are phase aligned in a frame buffer to compensate the different signal delays in a geographically distributed synchronised network. A bit rate offset, e.g. due to a synchronisation reference failure, is accommodated in the frame buffers by generating controlled frame slips, i.e. by repeating or deleting a frame. The objectives for the controlled slip rate performance are given in ITU-T Recommendation G.822 as the basis for the frequency tolerances of slave clocks in unsynchronised or holdover mode.

The wander of a signal refers to the long-term phase movements due to propagation delay variations (resulting from e.g. temperature effects), slave clock noise and low-frequency components of phase adaptation processes. To avoid slip buffer overflows in a synchronised network, the wander has to be limited. The wander reference model is given in an annex to G.823 for European bit rates and is reproduced in Fig. 1.

This basic configuration consists of a central primary reference clock and two switching nodes, which are locked to the clock by synchronisation links. The two switching nodes are connected by a transmission link. The total network wander budget is defined as the phase difference between clock and data of a synchronous equipment and is

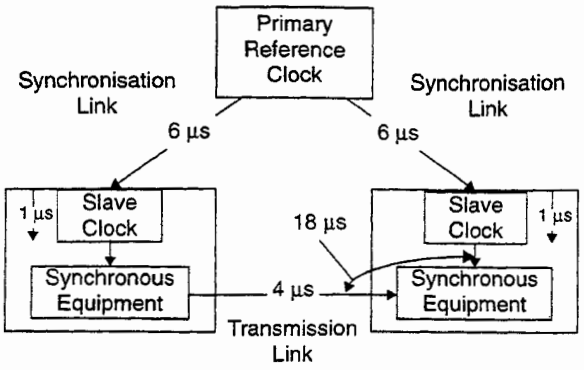


Fig. 1: Wander reference model

limited to 18 µs. It is subdivided into 2 times 6 µs for the synchronisation links including up to two slave clocks each, 2 times 1 µs for the switch internal clocks and 4 µs for the data path link. If these limits are met, no byte slips will occur. Otherwise, the slip performance objectives apply.

Standard SDH network element (NE) implementations require synchronisation but do not provide for timing transparent synchronisation links in contrast to PDH links.

A new timing distribution strategy is recommended in ITU-T Recommendation G.803 by using new SDH equipment clocks (SEC) and the STM-N line signal as the synchronisation distribution channel. It is combined with a synchronisation status messaging concept as a workaround for the clock hierarchy rule violation (low quality SECs within the chain of high quality switch internal slave clocks).

If the data path consists only of SDH NEs, wander considerations do not apply, because the SDH resynchronisation technique, based on pointers indicating the phase offset between transport frame and payload frame, decouples the input timing from the output timing. The VC-12 (2-Mbit/s level) pointer justification threshold spacing corresponds to at least two pointer step sizes of 3.5 µs each. As the synchronisation network wander is kept within 6 µs, VC-12 pointer adjustments are very infrequent events.

Single pointer adjustments impose phase hits onto the payload signal resulting in large jitter amplitudes, but the wander due to slave clock noise is reduced. Problems arise, if the data path is routed through a network consisting of a mixture of both SDH and PDH islands. In this more realistic case, the wander due to synchronisation clock noise feedthrough will not be absorbed at the next node. It will be encoded by the plesiochronous mapping process of SDH or PDH synchronisers. Therefore, it propagates through the network and accumulates.

Network simulations have been carried out considering various network scenarios [1]. In essence, if a path traverses more than two SDH islands, the wander budget cannot be met if standard techniques are employed. For larger networks comprising more than 10 islands, the wander budget can be exceeded by more than an order of magnitude. The simulation results indicate that the wander in the data path depends primarily on the synchronisation performance and on the number of islands rather than on the number of network elements per island.

In [1] the justification technique employing the Adaptive Threshold Modulation [2, 3] (in [1] referred to as phase dithering technique) has been investigated. The results show a superior performance in full accordance with the wander budget if only this type of pointer processor is involved. If some standard pointer processors are included, the performance decreases almost linearly in proportion to the number of pointer processors. The jitter results may deteriorate if a certain type of desynchroniser is employed which subsamples the spectrally shaped high-frequency pointer adjustment sequence. The corresponding aliasing effect translates high-frequency components into a lower-frequency region thus generating jitter.

Therefore, in addition to these wander generation effects resulting from the synchronisation clock noise feedthrough, some other implementation deficiencies may contribute to jitter and wander. Therefore, some implementation aspects of synchronisers, pointer processors and desynchronisers will be considered and their impact on the wander in the data path.

3 Wander Sources

In current PDH based transmission networks wander is generated primarily by long-term propagation delay variations of the signal due to temperature changes or laser frequency drifts. The contribution to wander by the bit rate adaptation process of PDH equipment is negligible. The wander of a digital switch is not defined, as the slip buffer decouples output timing from input timing.

SDH equipment imposes a strong clock noise feedthrough onto the payload signal if the justification thresholds are not exceeded. Therefore, the SDH pointer processor in conjunction with the synchronisation clock noise will be considered as a wander source. In addition, both the synchroniser due to the +/0/- justification technique employed and certain desynchroniser implementations may contribute to wander.

In Fig. 2 a general phase justification model is shown. It applies to all relevant types of equipment if equipment specific blocks are adapted to the required functionality.

The buffer is operated as an elastic store driven by write and read address counters. The input timing block consists of the clock recovery and the frame detector, which controls the write process of payload data into the buffer. The output timing block consists of the clock generator and a frame generator feeding the read counter with the payload timing. The phase detector monitors the buffer fill. It forwards the phase information to the justification circuit, which operates as a controller in a control loop (PLL, Phase Locked Loop). The control loop includes the phase detector, the justification controller, the output timing block and the read counter.

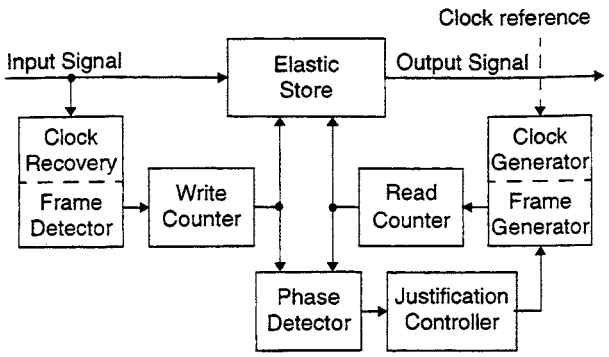


Fig. 2: General phase justification model

For switch applications the justification controller inserts positive or negative frame slips. For transmission equipment bit or byte justifications are generated by shifting the payload frame within the transport frame. In the case of SDH equipment the frame detector/generator blocks include pointer processors, which calculate/generate the offset value between transport frame and payload frame. The PDH interfaces of synchronisers and desynchronisers generally do not need the framing circuitry except for payload supervision purposes.

3.1 Slave Clock

Clocks have a major impact on the jitter and wander characteristics of telecommunications equipment. Standardisation bodies have devoted a considerable amount of attention to specifying slave clocks for SDH equipment. This work has been influenced by 3 new aspects:

- Increase in the number of slave clocks in a synchronisation distribution chain by more than an order of magnitude;
- short term stability requirements to comply with payload jitter requirements by preventing excessive pointer adjustment activity;
- cost considerations.

Fig. 3 shows a simplified block diagram of a slave clock, which represents the basic structure of a PLL. For simplicity, additional supervision and input selection blocks are not shown. The phase detector delivers the phase difference between input signal and oscillator signal to the controller. The controller gain in conjunction with other gain factors determines the cut-off frequency of the PLL. The voltage-controlled oscillator generates the output frequency. The cut-off frequency has to be optimised with respect to the contradictory requirements of both the input phase noise rejection and the oscillator noise suppression. Furthermore, hybrid PLL structures employing digital signal processing include analogue-to-digital (A/D) and digital-to-analogue (D/A) conversion stages as shown in Fig. 3. In principle, these structures are sensitive to limit cycle oscillations due to quantisation errors and aliasing effects resulting from the sampling of the phase difference. These errors may easily be reduced by an appropriate choice of the word lengths and the sampling frequency.

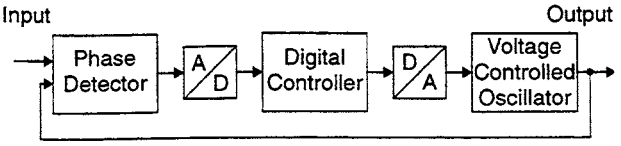


Fig. 3: Simplified slave clock model

The clock noise is generally modelled by a white noise source with Gaussian amplitude distribution followed by a set of filters for spectral shaping purposes. The long-term phase variations, i.e. the synchronisation network wander, is not considered in most simulation models.

3.2 Synchroniser

The synchroniser maps the incoming PDH tributary signal to the SDH-internal frame format. Frequency offsets are accommodated by bit justification techniques commonly used in PDH equipment. The SDH frame format has been chosen to simplify switching capabilities rather than to minimise the justification jitter. In PDH equipment the stuff ratio as the ratio between stuff events and stuff opportunities resulted in a high-frequency jitter spectrum. These jitter components could easily be removed by the desynchroniser internal PLL with low-pass-filter characteristic.

The SDH-frame period corresponds to an integer multiple of the nominal tributary frame period. In the case of nominal frequencies, no justification event will occur, which corresponds to the nominal justification ratio equal to zero. Frequency offsets of both polarities are accommodated by a positive-zero-negative justification scheme. Infrequent stuff events due to small frequency offsets are associated with arbitrary low jitter spectrum components, which cannot be reduced by the desynchroniser-PLL. In the context of wander measurements, no high-pass filter is applied. Therefore, wander of 1 UI_{pp} (Unit interval peak-to-peak, corresponding to one bit period) is generated for a threshold spacing of 1 UI between positive and negative justification. A larger threshold spacing would result in additional wander of equal magnitude.

Different techniques are known to reduce the synchroniser phase error due to the phase quantisation effect. These techniques are based on dithering the phase difference before quantisation. The dithering is performed by shifting the sampling instances in time [4] or by adding a modulation waveform to the phase difference [5]. For the latter method a fine resolution phase difference is required, which may be obtained by employing clock recovery circuits running at multiples of the input bit rate.

The resulting justification sequence consists of high-frequency positive and negative justification events. It may be considered as a pulse-width modulation of the phase difference. The original phase can be recovered easily by the low-pass filter characteristic of the desynchroniser-internal PLL.

3.3 Pointer Processor

The pointer processor provides the adaptation of received signals to the internal frame phase and frequency, which is defined by the local slave clock. It is composed of the pointer interpreter and the pointer generator. The pointer interpreter extracts the payload signal and the associated phase information from the transport frame. The pointer generator calculates a new pointer value indicating the phase offset of the payload frame to the new generated frame, which is synchronised to the network slave clock. This frame alignment technique permits the payload frame to float within the transport frame thereby reducing the buffer size requirements in contrast to the slip buffering technique known from digital exchanges.

The switching process is performed on the basis of transport frames after recalculation of the pointer value and generating a new frame rather than delaying the payload. This mechanism compensates the constant signal delays of different sources while reducing the buffer size requirements for wander accommodation.

A further buffer size reduction is supported by pointer movements in steps of +/- 1 offset values. The introduction of positive and negative justification opportunities within the transport frame corresponds to bit stuffing techniques known from PDH to allow for error free pointer adjustments. Additionally, the cyclical definition range of the pointer supports unsynchronised (plesiochronous) operation of SDH network elements.

Therefore, two techniques are combined within the SDH:

- the resynchronisation technique, which derives the output signal timing from the network clock at the cost of timing transparency
- the justification technique, which reduces the network clock noise feedthrough and is subject to jitter considerations as in PDH networks.

Pointer adjustments impose jitter onto the payload signal, which becomes apparent at the desynchroniser. Therefore, to control network jitter accumulation, only the resynchronisation aspect has been considered by keeping the probability of pointer adjustments low. This has been done by specifying a minimum threshold spacing and limiting the clock noise. However, this approach minimises the network jitter at the cost of increased network wander.

If the wander in the data path is considered, the clock noise feedthrough as a major wander source should be reduced taking into account the jitter accumulation objectives. The following functional blocks shown in Fig. 4 have a major impact on the wander:

- the pointer jitter accumulation filter
- the phase detector resolution
- the sampling circuit
- the decision circuit.

The pointer jitter accumulation filter is most commonly realised as a nonlinear filter by the pointer processor threshold spacing, which introduces the resynchronization

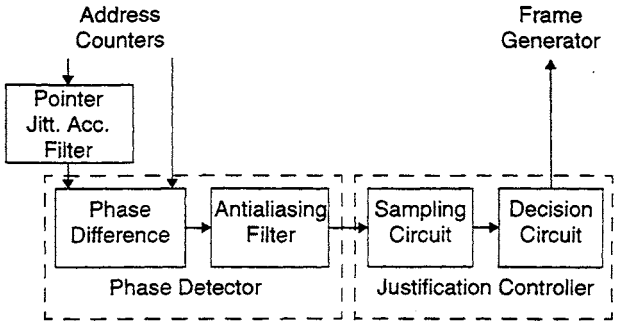


Fig. 4: Phase error relevant components of a pointer processor

effect. The phase error can be measured by applying a suitable random clock noise with respect to the thresholds or by switching the polarity of a frequency offset at the local clock.

This type of filter can be substituted by a linear low-pass filter, which may be considered as the digital counterpart of a desynchronizer in a PDH network. If this filter meets the SDH-PDH desynchronizer requirements, obviously no jitter accumulation will occur. As mentioned before, the introduction of an improved decision circuit will relax the jitter suppression requirements.

The phase detector resolution is an implementation-specific parameter. If hybrid implementations are employed in conjunction with an analogue antialiasing filter and a subsequent analogue-to-digital converter, the resolution depends only on the converter step size. In typical applications, the phase difference is sampled or subsampled with a fixed phase relationship to the output signal. The resulting phase error is associated with the time resolution of the byte clock ($t_b = 1/f_b = 51.44$ ns). At the TU-12 level this effect is negligible.

The sampling circuit subsamples phase detector values at justification opportunities. The phase detector monitors the buffer fill by calculating the difference in the values between the read and write address counters. The difference signal consists of integer values, which represent the instantaneous number of bytes within the elastic store. Assuming constant frame phase offsets, the signal spectrum consists of high-frequency components at multiples of the byte clock frequency (f_b=19.44 MHz). The buffer fill response of a VC-4 buffer (140 Mbit/s level) has already been presented in [6]. Fig. 5 shows an example of the corresponding buffer fill over time of a VC-12 buffer (2-Mbit/s level). The received higher order pointer (AU-4

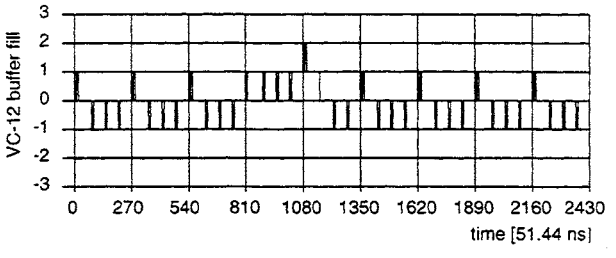


Fig. 5: VC-12 buffer fill

pointer) value is chosen equal to 10 in conjunction with a frame phase offset of 317 bytes. In this example the peak-to-peak buffer fill is 3 bytes.

The minimum time interval between adjacent justification events is 2 ms for lower order pointer processing (TU-12, coresponding to 2 Mbit/s). As a consequence of the subsampling of the justification controller, a maximum phase error could be expected corresponding to the peak-to-peak value of the phase difference due to the aliasing effect. If the sampling point is fixed in relation to the justification opportunity of the transmitted frame, only the data gaps of the received frame contribute to the phase error. The major part of the phase error is caused by a missing byte (V-byte) in an almost regular sequence of VC-12-bytes. It is superposed by a smaller part due to the Section Overhead (SOH) gaps. This effect becomes apparent for constant frequency offsets applied at the pointer processor node, which results in the so-called gapped pointer sequence. The aliasing effect due to subsampling can be avoided by a suitable antialiasing filtering as proposed in [6]. The lowest frequency to be suppressed should be the frame repetition frequency (f_f = 8 kHz).

Fig. 6 depicts decision circuits. The standard decision circuit represented by Fig. 6a is used if the waveform added before quantisation is kept to zero. To improve the clock noise suppression capabilities phase dithering techniques known from PDH can be applied. In [7] different waveforms have been proposed for application at the decision circuit in Fig. 6a. The corresponding phase error spectrum resulted in a reduction of low-frequency components.

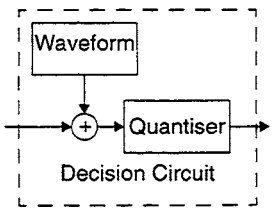


Fig. 6a: Threshold Modulation

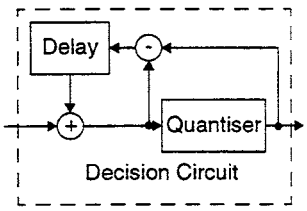


Fig. 6b: Adaptive Threshold Modulation

[2, 3] have proposed and investigated the decision circuit employing an improved phase dithering method as shown in Fig. 6b. This method employs higher-order Sigma-Delta-Modulation techniques and has been called Adaptive Threshold Modulation. The resulting phase error spectrum equals a high-pass filtered white noise characteristic. The absence of low-frequency phase error components simplifies the error suppression filter, which recovers the phase information at the next node (pointer processor or desynchroniser).

3.4 Desynchroniser

The desynchroniser demaps the PDH payload signal from the SDH frame by removing all stuff bits and bytes. As shown in Fig. 2 the data is written into the elastic store with the associated gapped clock. To comply with PDH jitter requirements, the read clock is smoothed by a PLL. It consists of the phase detector monitoring the buffer fill, the justification controller and a voltage controlled oscillator (VCO) as the clock signal generator.

[7, 8] have proposed different implementations to accommodate the large jitter amplitudes which are superposed onto the PDH payload by pointer adjustments. As a rule, nonlinear techniques are employed which are based on leaking out bits rather than bytes, thus reducing the phase steps. The leakage rate may be adapted to the frequency offset. The jitter amplitudes due to single leaked bits may be reduced by phase dithering techniques.

In addition to these implementation-specific output jitter waveforms, similar phase error generation mechanisms are encountered as in pointer processors. The error of the justification controller (decision circuit) (Fig. 4) corresponds to imperfections due to bit leaking or cycle stealing techniques or VCO jitter. The pointer jitter accumulation filter reflects the jitter suppression characteristic of a desynchroniser.

The sampling circuit may cause a similar error as in a pointer processor. If a digital signal processing circuit samples the buffer fill only once per frame or once per leaking period, the data gaps of three consecutive bytes (R-bytes) may result in an approximate phase error of up to 12 µs.

In ITU-T Recommendation G.783 a linear low-pass filter is considered as being the ideal characteristic of a desynchroniser. In [6] a digital implementation of a 140-Mbit/s desynchroniser with linear low-pass characteristic and a cut-off frequency of f_c = 0.35 Hz has been presented. The basic principle includes a numerical controlled oscillator (NCO) followed by an analogue PLL to suppress the high-frequency jitter generated by the NCO. At the 2-Mbit/s level the desynchroniser may be simplified by omitting the analogue PLL thanks to the relaxed jitter requirements.

4 Results

This section deals with simulation and measurement results which demonstrate the phase transfer characteristic of pointer processors. The pointer processor basically covers all effects which have been discussed with respect to synchronisers and desynchronisers. The effect of the synchroniser justification threshold spacing is covered by the corresponding effect at the pointer processor. The phase error at the desynchroniser due to the gapped frame structure (R-byte) is based on the same mechanism as the gapped pointer effect.

The phase error (time interval error) is defined as the low pass filtered phase difference of a VC-12 between input and output of a pointer processor. The low-pass filter reflects the phase transfer characteristic of an ideal desynchroniser with a maximum cut-off frequency of f_c = 10 Hz as per

G.783. The input signal is mapped synchronously into the VC-12-container to avoid additional synchroniser errors. As the input signal phase is kept constant, only the output phase is subjected to a low-pass filter. The phase error of the simulation results is scaled in time units [µs] and plotted against a frame offset value between input and output signal.

Fig. 7 shows a simulation result of a standard pointer processor. The AU-4 pointer values are zero. The phase error scaled in µs is plotted against the offsets between input and output frames scaled in bytes. In this figure the effects of the gapped pointer and the threshold spacing can be identified. The almost regular pointer sequence associated with the sawtooth signal with frequent phase correction steps is superposed by a sawtooth signal of one frame period corresponding to a frame delay of 2430 bytes. This sawtooth error is due to the V-byte gap in the frame structure, whereas the smaller irregularities are caused by the SOH gaps.

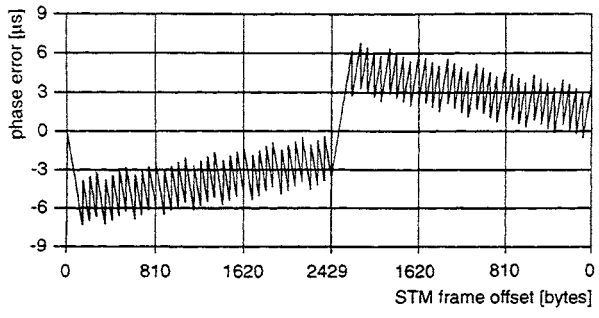


Fig. 7: Gapped pointer effect and threshold spacing

The constant frame movement simulates a constant frequency offset at the pointer processor clock, thereby forcing the buffer control loop to operate towards one threshold. Changing the polarity of the frequency offset, which is simulated by a decrease of the frame offset, enforces the operation towards the opposite threshold. For this simulation the delay of the input frame has been increased up to a value of 2429 bytes before it has been decreased to zero. A pointer adjustment is carried out if the thresholds of +/- 1 byte are exceeded, which results in a phase error of 3 VC-12 bytes (10.7 µs) due to the threshold spacing. This phase error may be even worse, because the threshold spacing as per ITU-T Recommendation G.783 should be at least 2 bytes.

Fig. 8 represents the simulated phase error of a phase detector employing an antialiasing filter. Almost the same result would be obtained if the model included a justification controller employing Adaptive Threshold Modulation (Sigma-Delta Modulation). The high-frequency components generated by the Adaptive Threshold Modulation would be almost fully suppressed by the low-pass filter representing the desynchroniser. The absence of low-frequency phase errors of the justification controller provides timing transparency.

In this simulation the STM-frame offsets are consistently kept to zero, whereas the AU-4 pointer value is increased,

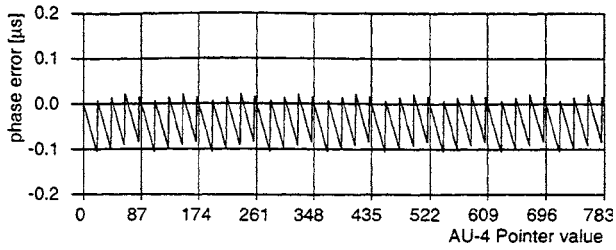


Fig. 8: Phase error simulation of a VC-12 (2-Mbit/s level)

because a variation of the frame delay at a fixed pointer value would lead to an error-free result. The increasing AU-4 pointer value results in a moving VC-4 frame offset. The residual phase error is due to gaps in the frame structure, which modify their position according to the AU-4 pointer value. The AU-4 pointer step size equals a phase shift of 125/783 µs =159.642 ns. A VC-12 byte, however, is shifted by only 154.321 ns. This phase error accumulates until a VC-12 byte is shifted over the SOH gap, thereby correcting the average shift value with a periodicity of 36 phase steps per frame. The same principle applies to the missing VC-12 byte at the V-byte gap. This phase error has the opposite polarity and one phase correction step per frame.

In Fig. 9 the simulation result of a pointer processor employing an antialiasing filter and a standard justification controller is shown with the same parameters as in the previous simulation. This demonstrates that different implementations of phase error affecting functional blocks can be combined. However, the associated effects can be distinguished. The observed pointer sequence is almost regular except for the additional residual phase error of the phase detector. The irregularities may be considered as interference between the residual phase detector error and the sawtooth error associated with the pointer sequence.

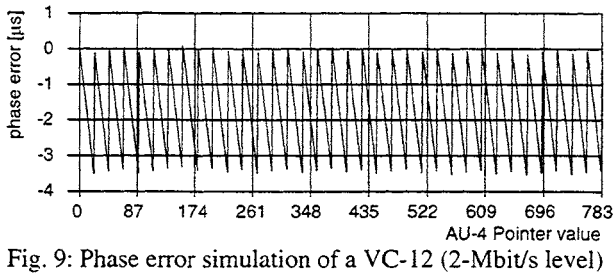


Fig. 9: Phase error simulation of a VC-12 (2-Mbit/s level)

Fig. 10 shows the measurement configuration for monitoring the phase error. The measurement setup consists of a 2-Mbit/s test signal generator, a synchroniser operated at nominal frequency to avoid bit justification errors, an Add/Drop Multiplexer (ADM) including the AU-4 pointer processor, a second Add/Drop Multiplexer (ADM) including the TU-12 pointer processor, and a desynchroniser with a linear low-pass filter characteristic (f_c = 0.4 Hz). The phase difference between test signal and the desynchronised output signal is calculated in UI (1 $UI_{2Mbit/s}$ = 488 ns) and recorded over time. The network elements are part of the PHASE (PHILIPS Advanced SDH Equipment) system family.

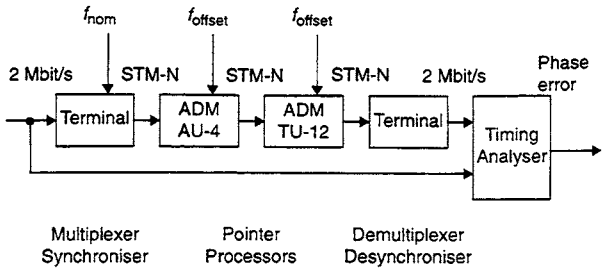


Fig. 10: Measurement configuration

The clock of the first ADM (AU-4) is operated with a frequency offset of 0.2 ppm (parts per million) to generate continuously moving pointer values. The phase error of this pointer processor is less than 15 ns (no gapped pointer effect). The second ADM (TU-12, device under test) is switched to ITU-T mode corresponding to the implementation of the simulation result in Fig. 9.

The measurement result in Fig. 11 shows a comparable characteristic with respect to the corresponding simulation result in Fig. 9. The pointer sequence is almost regular thereby demonstrating the suppression of the gapped pointer effect. The interference between the SOH and the V-byte is not visible due to measurement noise.

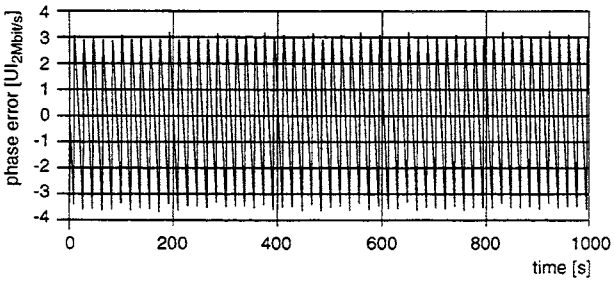


Fig. 11: Measurement result, ITU-T mode

Fig. 12 shows a measurement result of the same setup as in Fig. 10, where the justification controller of the second ADM (TU-12) is switched to Adaptive Threshold Modulation. As expected from the corresponding simulation in Fig. 8 the high-frequency phase error components are suppressed. The residual observable phase error is due to the frame restructuring mechanism resulting from AU-4 pointer movements. It matches well with the simulation results, thus demonstrating both the applicability of the simulation models and the superior performance.

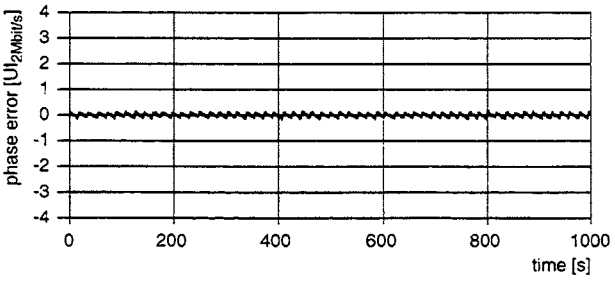


Fig. 12: Measurement result, Adaptive Threshold Mod.

5 Conclusion

The introduction of SDH-based transmission networks provides significantly more flexibility thanks to the inherent switching capabilities. Network synchronisation and timing aspects have gained considerable attention in the context of compatibility to PDH equipment with respect to jitter limits and to the wander budget of digital exchanges in conjunction with the frame slip performance.

In this paper, the major sources affecting the phase transfer characteristic of SDH equipment are discussed. It is shown that the implementation has an impact on the jitter and wander performance. The so-called gapped pointer effect is demonstrated to be caused by a subsampling of the incoming gapped frame structure. The application of a suitable antialiasing filter significantly reduces the phase error from more than 3.5 µs to values of less than 150 ns. These results are further improved by the introduction of Adaptive Threshold Modulation, which provides timing transparency. Measurement results are confirmed by simulation results obtained on the basis of simplified models.

This paper demonstrates that small implementation modifications enable SDH equipment to comply with the network wander budget. This is considered to be a far more cost-efficient solution than a significant tightening of the network synchronisation requirements.

6 References

[1] G. Garner: "Total Phase Accumulation and Jitter in a Network of VT Islands Containing Pointer Processors that Employ Phase Dithering", T1X1.3/04-095, September 1994.

[2] M. Klein, R. Urbansky: "Network Synchronization - A Challenge for SDH/SONET?", IEEE Comunications Magazine, September 1993.

[3] R. Urbansky: "Telecommunications Network Synchronization Using Improved SDH Pointer Processing Techniques", Proc. 8th EFTF, March 1994.

[4] P. E. K. Chow: "Jitter Due to Pulse Stuffing Synchronization", IEEE Trans. Comm., July 1973.

[5] W. D. Grover, T. E. Moore, J. A. McEachern: "Waiting Time Jitter Reduction by Synchronizer Stuff Threshold Modulation", Proc. IEEE Globecom '87.

[6] R. Urbansky: "Simulation Results and Field Trial Experience of Justification Jitter", Proc. 6th World ITU Forum, October 1991.

[7] R. G. Kusyk, W. A. Krzymien, T. E. Moore: "Analysis of Techniques for the Reduction of Jitter Caused by SONET Pointer Adjustments", IEEE Trans. Comm., April 1994.

[8] H. Sari, G. Karam: Cancelation of Pointer Adjustment Jitter in SDH Networks", IEEE Trans. Comm., December 1994.

SOME ASPECTS OF TIMING IN TELECOMMUNICATION AND POWER SYSTEMS

Andrzej DOBROGOWSKI, Mieczysław JESSA

Institute of Electronics and Telecommunication
Poznań University of Technology, Poland

Abstract

Modern telecommunication network requires synchronisation of all nodal clocks of the network. Different possibilities of creating the synchronisation layer are considered. The influence of the "state of technology" is emphasised. Timing source synchronised to GPS is presented. The system monitoring timing signal parameters is discussed. The role of time signals for dynamic tracking of the state vector of power systems is considered.

The rapid changes of the property relations and successful decentralisation of the economy in the countries in transition to the free market economy have caused a significant increase of requirements on timing signals. The countries can not use directly the experiences of the western societies because of the lack of suitable technical structure of telecommunication networks and power systems. Our considerations are referred to the recommendations we have formulated for telecommunication networks and power systems in Poland in terms of time and frequency signals.

I. Telecommunication networks

The dissemination of the synchronisation signals from Primary Reference Clock (PRC) to the arbitrary nodal clock is based on the hierarchical structure shown in Fig.1. In practice there are two basic methods of distributing synchronisation signals in each branch of the "tree" (Fig.1).

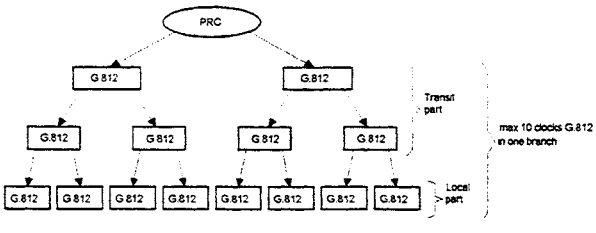


Fig. 1. Timing signal distribution in telecommunication network.

First of them consists in creating a direct track for synchronisation signals to all nodes (Fig.2), completely independent of the SDH network. In the second approach (Fig.3) the nodal G.812 clocks are included into the SDH network structure and they are synchronised directly from STM-N lines. In western countries PDH networks synchronised to PRC already exist. Therefore, the simplest and natural method of the SDH network synchronisation is the usage of the synchronised PDH nodes for delivering the synchronisation signals. The transition from the direct track clock synchronisation to the synchronisation from the STM-N lines can be delayed in time to the moment of the getting to complete knowledge about topology and principles of possible reconfiguration of SDH networks on large areas.
In some countries, e.g. in Poland, the synchronised PDH network does not exist and it will not be created. It seems that for the SDH network synchronisation purposes we should apply in this case STM-N lines synchronising G.812 clocks together with the SDH

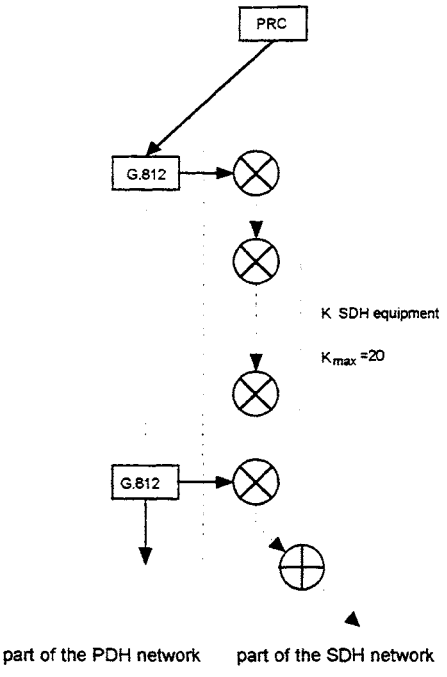


Fig.2. Direct track for synchronising signal.

equipment. The question is if this idea is sufficiently motivated for today.

The synchronisation of the G.812 clocks from the STM-N lines enables to combine creating the synchronisation layer of the network with installation of the SDH network itself, i.e. the integration of methods and equipment. The whole undertaking may be performed by one company and
carefully watched by the customer. The customer needs only the proper knowledge, measurement equipment and the right procedures. This possibility seems to be very attractive from the decision point of view but it does not have sufficient motivations in countries like Poland.

We think the direct application of the STM-N lines for the G.812 clock synchronisation in the relatively large countries may include many mistakes. First, there is nowhere a large SDH network synchronised only from the STM-N lines nowadays.

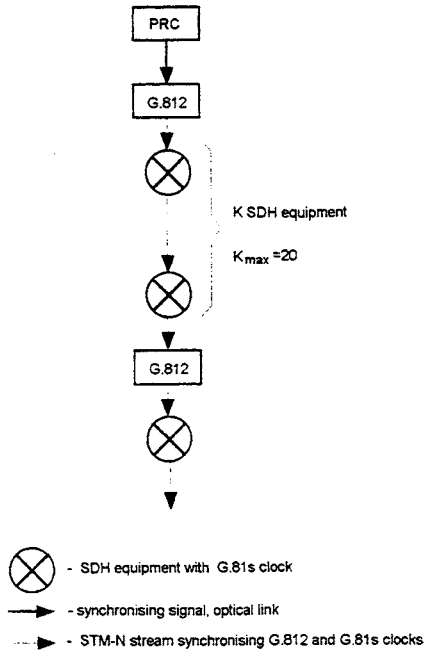


Fig. 3. Network synchronisation from recoverd timing signal of the STM-N line.

Second, the synchronisation layer of the network could change dynamically due to the reconfiguration process of the SDH network. Thus, a timing loop can arise which may make the transmission in a large part of the network impossible. Additionally, due to the lack of knowledge about the final structure of the SDH network it is difficult to perform reliable simulations of the synchronisation layer behaviour. The partial solution of this problem is a division of the whole country into smaller sub-areas, each of which would be controlled by its own PRC. This solution is rather expensive and requires numerous and well-trained staff. Some companies, however, have noticed the problem and have been offering low-cost PRC with rubidium oscillators synchronised to GPS for example. Another disadvantage is currently rather small ability of the Synchronisation Supply Units (SSU) containing G.812 clock to the direct usage of the Synchronisation Status Messages (SSM) generated in the SDH network. Finally, a single error in a message for a properly operating SSU may effectively block the proper work of some important components of the network.

To overcome the above mentioned problems we recommend an application of at least two Primary Reference Clock for medium size countries and the distribution of the synchronising signals to all SSUs via spare optical fibres or windows of the optical cables (the number of the spare optical fibres is about 100 % and optical windows about 400 %). We think that this kind of network synchronisation should be only temporary and slowly eliminated, proportionally to the increase of the knowledge about the behaviour of large SDH networks and proportionally to the advance in integration of SSU (in terms of control and transmission) with the SDH equipment.

In Fig.4 we can see the basic scheme of the single SSU for transit nodes. The primary and the secondary synchronisation lines are completed with the third line connected to the GPS receiver.

The GPS receiver delivers the UTC scale to the computers and measurement equipment. It generates also the backup for SSU synchronisation signal 2048 kHz with the long term accuracy offered by the GPS system, i.e. about 10^{-13}. We have called this device a Synchronised Time Standard (STS). The precise time scale in each telecommunication node enables marking of the events and correct determination of the state of the telecommunication system. The priority of the input of the SSU at which the STS is installed depends on a specific situation. We think that STS supplied with high stability VCXO can replace much more expensive (about 4 times) SSU at the lowest levels of the telecommunication network hierarchy.

The rapid increase of the importance of time and frequency signals in telecommunication and the real possibility of selling these signals for different services, e.g. mobile communication, satellite communication, TV, power systems, etc. introduces new

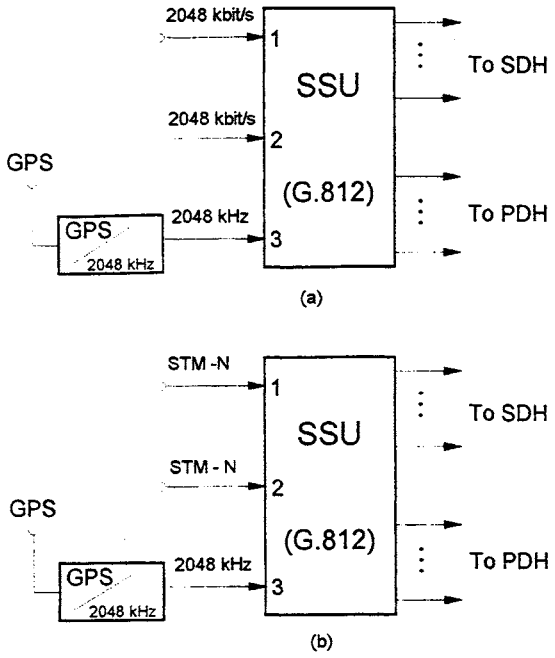


Fig. 4. Synchronisation Supply Unit (SSU); a) direct synchronisation by 2048 kbit/s data stream, b) synchronisation from STM-N lines.

requirements concerning monitoring of the distribution of timing signals in the network and their parameters in many nodes and in many points inside the node. The next reason for monitoring the timing signal parameters is a significant correlation (30-90%) between bit-error-rates observed in existing networks and the quality of the timing signals (only a few of them are used for network synchronisation). It seems the GPS system can offer a chip and easy access to the UTC scale, a very important factor for the proper interpretation of the results of measurements in different, geographically separated nodes.

The above mentioned reasons had caused that we have designed and built a low-cost (less than 15.000$) measurement system SP-2000. The system consists of the following devices (contained in one package):

1. *8 channel (R)TIE, TVAR, slips, etc. meter.*
2. *8 channel phase comparator with the output signal bandwidth 0-100 kHz (for f_{in} >=1MHz) and 0-10 Hz.*
3. *Fast frequency comparator with input frequency ranging from 10 kHz to 10 MHz.*
4. *High stability VCXO with nominal frequency 2048 kHz and frequency drift < $2x10^{-10}$ per day.*
5. *A circuit for rough and precise correction of frequency; sinusoidal and saw-tooth modulation of the output signal frequency.*
6. *Two independent circuits for direct frequency conversion from 5 MHz to 2048 kHz.*

The dedicated software contains two packages: a measurement program and an analysing program. The program "Analyse" reconstructs the measurement process and enables an inclusion of

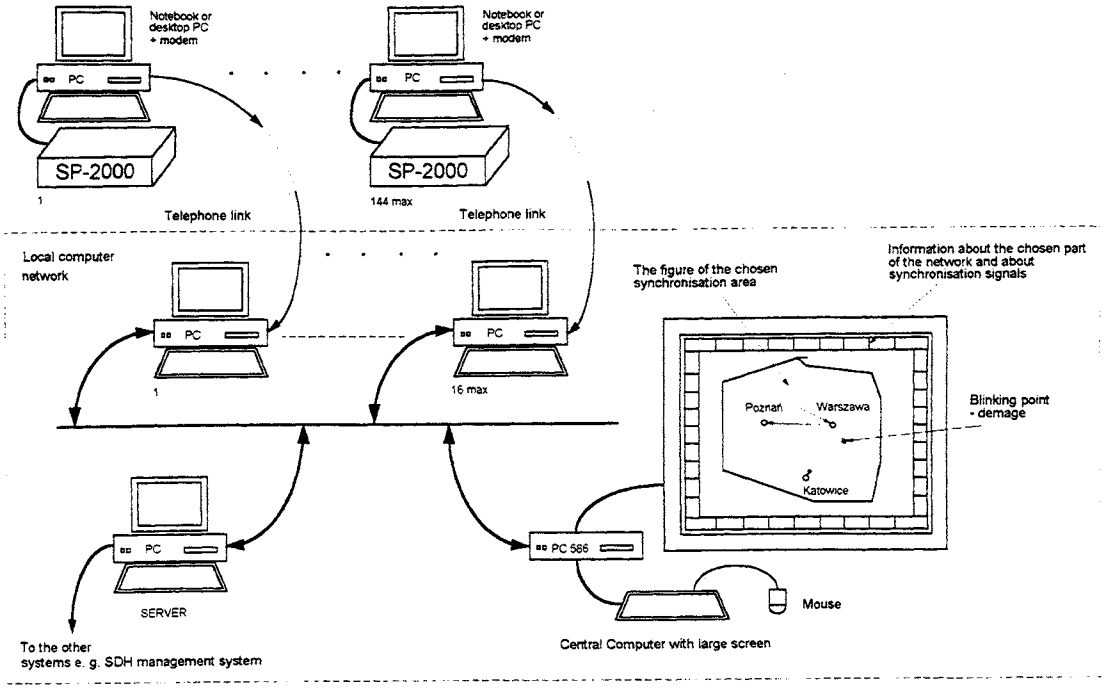


Fig. 5. Remote monitoring system of the timing signals.

user's procedures.

The SP-2000 system is controlled from a PC computer (notebooks are preferable) and communicates via modem with the external world. The whole system is supplied with separate PDH probes extracting the timing signals from 2048 kbit/s (HDB-3) data stream and SDH probes extracting the timing signal from the STM-1 line (NRZ) and converting directly the frequency 155.52 MHz into 2048 kHz signal.

In Fig.5 we present the basic idea of the remote monitoring system. The single SP-2000 measures the timing parameters of the chosen lines of the nodal equipment and sends the results to the computer network. The local computer network analyses the data, creates reports and shows the route of the timing signals and their parameters on a screen. The network server transmits the necessary information to the SDH management system.

II. Power systems

In the power systems the effective usage of the accuracy of the time signals is limited by the speed of the power sensors, contactors, enclosures and it does not exceed now the value 1 ms. The information about precise time is necessary in two main places of the power system. The first is the central control system, mainly the decision system implemented in the computer network. The other contains remote circuits and actuators organised into geographically separated groups. In the first case we simply supply the UTC scale, e.g. from STS to the computers. The time scale is put in the computer with desired accuracy (relative to the UTC), e.g. 5, 10, 100 μs or 1 ms. The easy access to the precise time scale for the geographically separated control systems enables the determination of the state vector and the time sequence of the events in the power system. It means we can mark the events and show what are the results. It leads to the significant savings of money and time.

A single group of the remote control circuits and actuators requires the same time scale but the time scales in different groups may be different. The main problem consists in the time scale distribution from the central clock to the "slave" clocks of the single group in the presence of large disturbances caused by switching of the large powers, transient or permanent faults. One of the effective method of distributing time signals in the presence of low-pass disturbances is the application of the GPS receivers. The reason is the relatively large distance between the GPS carrier and the peak of the spectrum of disturbances. A drawback of this method is a large price of the equipment and hard environment conditions (e.g. temperatures between -40 and +60°C, high humidity) in comparison to the typical transmission systems.

III Conclusions

Modern telecommunication network requires synchronisation of all nodal clocks of the network. Many countries (e.g. Poland) are going to build the Synchronous Digital Hierarchy (SDH) without the stage of the Plesiochronous Digital Hierarchy (PDH), despite the existing PDH islands. In western countries synchronisation signals for SDH parts of the network can be delivered from synchronised PDH parts. In many countries such possibility does not exist because they have at most PDH islands and they do not intend to develop plesiochronous hierarchy. We suggest for these countries an application of the G.812 clock synchronised to systems in the chosen nodes and further distribution of timing signals by the means of optical fibres which are not used for other purposes in the cables. The presence of many suppliers of the PDH and SDH equipment and the lack of the complete international technical norms requires the precise measuring of the timing signals parameters. In the paper we described the automatic monitoring system and suitable equipment devoted for measuring of the timing signals parameters. Both the system and the specific devices can be used in on-line and off-line monitoring of many parameters of timing signals in many different applications.

In a power system we require mainly an information about the accurate date and timing signals having several standard frequencies. The error between the local date and the UTC time should not exceed presently the value 1 ms. We expect this limit will be decreased in the future to about 1 μs. The information about the precise time in geographically separated power stations is used to determine the state vector of the power system and the sequence of events. We suggest the usage of the GPS receivers synchronising local time scale generated by the high-stability VCO oscillators (the day frequency drift less then 1×10^{-9}). The precise, stable and relatively cheap local "time generator" can provide the data with sufficiently small long-time error during temporary loss of the GPS signals (e.g. because of military operations) or their improper quality. We hope the results of our work will be applied in other, small and large, systems in many cities.

Stability characterization of Stand Alone Synchronisation Equipment

R. Bonello*, A. Manzalini*, A. Godone**, C. Novero**, P. Tavella**

*CSELT, Centro Studi E Laboratori Telecomunicazioni, Via G. Reiss Romoli, 274, Torino, Italy

**IENGF, Istituto Elettrotecnico Nazionale G. Ferraris, Strada delle Cacce, 91, Torino, Italy

Stability characterisation of clock units intended for operation in telecommunications synchronisation networks is performed in terms of typical TIE, ADEV and TDEV parameters. Stability performance in a commercially available clock unit is measured and compared with relevant requirements currently agreed in telecommunications standardisation bodies. Phase noise, frequency offset and drift components effects on measured stability are illustrated as well.

Introduction

Network synchronisation plays a fundamental role in the evolution towards telecommunications digital networks based on new transmission techniques, such as Synchronous Digital Hierarchy (SDH), which strongly relies on synchronous equipment to provide easier digital signals multiplexing and cross-connecting facilities, and improve network robustness and efficiency [1]. In fact, network synchronisation implies that all the transmission equipment is actually timed by a common reference signal, which is distributed throughout the network by means of synchronisation links interconnecting all the geographical nodes where equipment is located; synchronisation links and nodes can thus be considered as basic elements of a *synchronisation network*, explicitly devoted to distribute the timing reference over the whole telecommunications network [2],[3].

Simple hierarchical topologies are currently preferred in synchronisation networks, with synchronisation links from a high quality master clock (*primary reference clock*) directly synchronising few higher level nodes, which, in turn distribute timing to lower level nodes; multiple links to each node are provided for network robustness against link failures (Fig. 1).

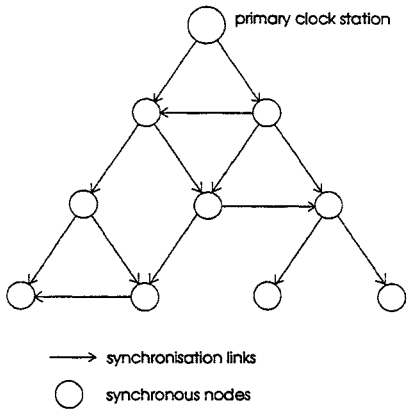


Figure 1: typical synchronisation network topology

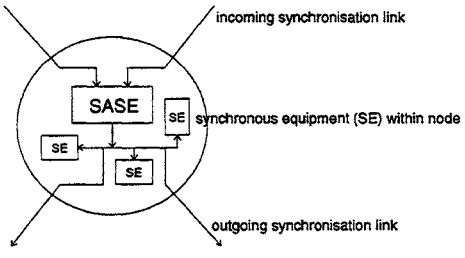


Figure 2: synchronous node; timing selection and distribution via SASE

Within each node, a special timing unit known as Stand-Alone Synchronisation Equipment (SASE) is typically required [3], in order to monitor and filter all timing signals carried by synchronisation links incoming from higher level nodes, to provide node

synchronisation and to feed synchronisation links towards lower level nodes (Fig. 2).

In detail, SASE units should check the quality level of available incoming timing signals and choose and track the best one (locked operating mode) to distribute a high quality timing reference signal to all the synchronous equipment (SE) in the node, and to the outgoing synchronisation links as well.

Furthermore, SASE units should cope with temporary unavailability of every external reference signal from incoming synchronisation links, providing a stable internally-generated timing signal, whose frequency should reproduce as close as possible the lost reference one (holdover operating mode), to keep the transmission equipment in the node properly working in spite of the intervening network synchronisation failures.

Due to their major role in networks synchronisation, operational modes and performance requirements for SASE units are currently being under study in several international standardisation bodies (ITU-T, International Telecommunication Union, Telecommunication standardisation sector; ETSI, European Telecommunications Standards Institute), in order to ensure proper operation of the whole synchronisation network. As a matter of fact, ETSI is completing a Technical Standard document [3] on synchronisation networks which will thoroughly specify timing characteristics of SASEs and other slave clocks intended for timing supply to synchronous equipment. In particular, since timing stability affects the overall synchronisation network quality [4],[5], detailed clock stability performance requirements will be included, considering both the locked and holdover operating modes. At the same time, as the first timing units designed as SASE become commercially available, measuring their actual performance pursues a twofold goal: on the one hand, it gives a rough evaluation of their compliance to well established basic requirements, on the other hand, it provides a clear picture of clock state-of-the-art technology, suggesting possible refinements in technical specifications under study.

In this paper, following a brief description of measurement set-up, typical stability characteristics measured in a commercial SASE are presented in terms of Time Interval Error (Maximum, MTIE, and root mean square, TIE rms), Allan Deviation (ADEV) and Time Deviation (TDEV), as currently addressed in international standards for telecommunications. Furthermore, measurement results will be shortly discussed in order to point out how oscillator phase noise and other environmental effects respectively impact on the overall clock performance, thus giving further information on physical effects which should be kept in account when specifying clock requirements.

Measurement set-up

Clock stability characterisation for telecommunications purposes typically assumes a computer-assisted measurement set-up in order to collect long sequences of samples of measured Time Error (TE, namely, phase difference expressed in time units); based on mathematical elaboration of the acquired samples, the corresponding quantities of interest can subsequently be calculated and compared with prescribed maximum limits in order to check clock compliance with existing requirements.

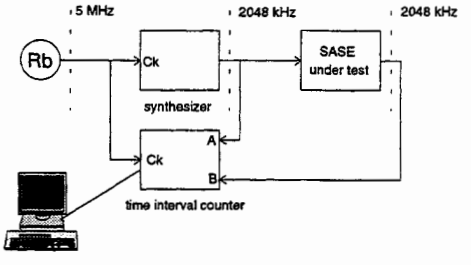


Figure 3: test-bed for stability measurements with clock in locked mode

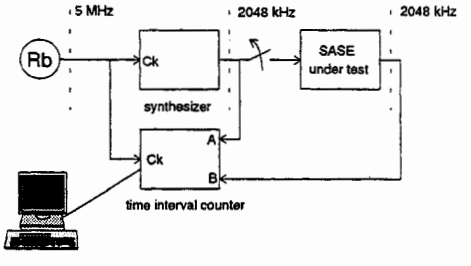


Figure 4: test-bed for stability measurements with clock in holdover mode

Figs. 3 and 4 report the measurements test-beds we set up to characterise clock

stability in the locked and holdover modes, respectively: a frequency synthesiser, timed by a high quality rubidium oscillator, provides a stable reference signal, which is used to slave the clock (SASE) under test in the locked mode of operation; a computer-assisted counter is finally employed to repeatedly measure and acquire TE samples of the clock output signal (channel B) against the synthesised reference signal (channel A).

When characterising clock performance, a test-bed background noise evaluation should be carried out as a preliminary step; Fig. 5 depicts how this was performed in our set-up: the same reference signal provided by the frequency synthesiser was fed to both measurement channels A and B, thus evaluating the trigger relative error between the two.

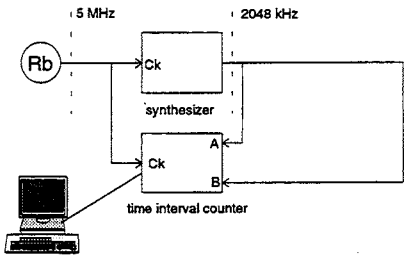


Figure 5: test-bed background noise measurement

Measurement results

A measurement session covering a 5 days observation period (with sampling time τ_0=1 s) was performed for each measurement set-up, and post-processing of the corresponding TE samples sequences finally led to the stability plots presented in this section.

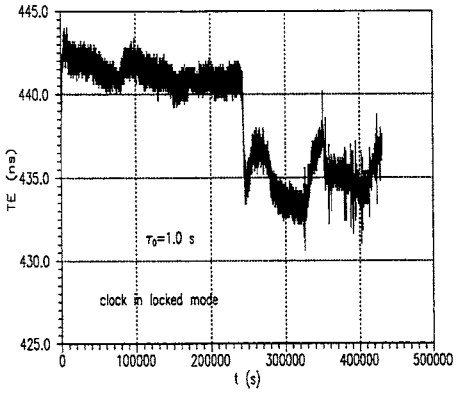


Figure 6: TE samples sequence measured on the clock in locked mode

Figs. 6 and 7 show the measured TE sequences in the locked and holdover mode of operation, respectively.

In Fig. 6 the clock under test is tracking the stable reference signal against which its output stability is measured: a high frequency noise component is evident, spanning a range of few nanoseconds. On the contrary, in Fig. 7 the SASE is not locked to any external signal and frequency offset and drift effects dominate the measured TE behaviour.

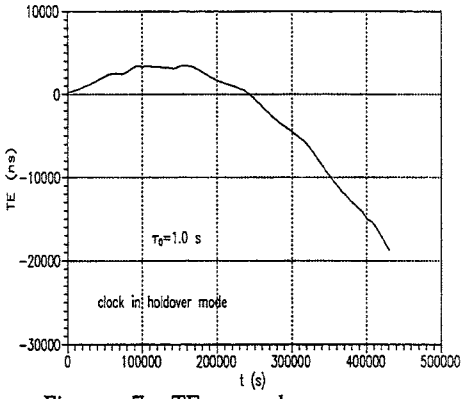


Figure 7: TE samples sequence measured on the clock in holdover mode

Based on statistical processing of the TE sample sequences in Figs. 6 and 7, various recommended stability quantities can finally be computed and plotted. A typical parameter to specify clock stability in communications networks is MTIE: MTIE is computed as a function of observation time $\tau = n\tau_0$ (an integer multiple of sampling time τ_0), according to the following definition [3]:

$$\mathrm{MTIE}(\tau) = \max_{1 \le k \le N-n} \left(\max_{k \le i \le k+n} (x_i) - \min_{k \le i \le k+n} (x_i) \right)$$

where x_i refers to the i-th TE sample and the number N of samples is about 430000 in our case. From the above definition, MTIE(τ) corresponds to the maximum TE variation occurring over any τ-interval within the whole TE sequence, giving a sort of worst-case bound for TE fluctuations.

Fig. 8 shows a plot of MTIE values evaluated from the acquired TE sequences, compared with relevant requirements in current ITU-T standard [6]: *ideal* mode, in the picture, stands for locked mode (with clock tracking an 'ideally' stable reference

signal), while holdover requirements differ according to the clock level in the synchronisation network (*transit*, for higher levels, *local*, for lower ones).

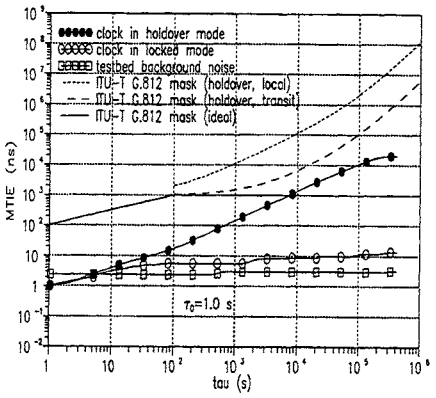


Figure 8: Maximum Time Interval Error (MTIE) results compared with relevant telecommunications standards [6]

It should be noted that the SASE under test is fully compliant with current specifications for the (higher quality) transit clock; in particular, locked mode performance appears to be much better than required (20 ns against 1000 ns), which suggests that current clock technology would actually allow tighter requirements on locked clock stability; in turn, the ability to impose more severe specifications on clock performance would accordingly increase the degree of flexibility in designing the synchronisation network and its expected quality, thus improving the overall communication network efficiency.

Unlike MTIE, TIE rms gives an averaged value of TE variations over the observation interval τ, according to the following definition [3]:

$$\mathrm{TIE}_{\mathrm{rms}}(\tau) = \sqrt{\frac{1}{N-n}\sum_{i=1}^{N-n}(x_{i+n}-x_i)^2}\,,$$

Estimates of TIE rms from the collected TE sequences are shown in Fig. 9.
From its definition, TIE rms is clearly related to the classical standard deviation of frequency fluctuations (assuming such fluctuations as a zero mean process),

$$\sigma(\tau) = \sqrt{\frac{1}{N-n}\sum_{i=1}^{N-n} y_i^{\,2}}\,,$$

where $y_i = (x_{i+n} - x_i)/\tau$ can actually be considered as the i-th normalised frequency sample, with frequency measured over the observation interval τ.

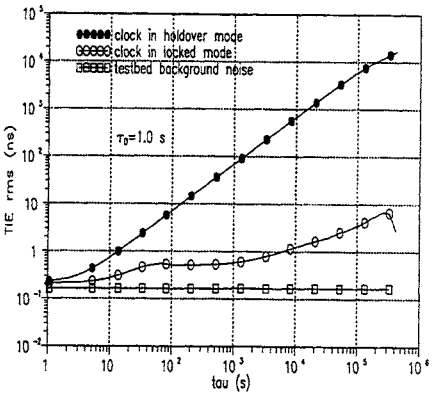


Figure 9: Time Interval Error, root mean square (TIE rms)

Unfortunately, classical standard deviation can suffer of a well-known [7] convergence problem in case of some low frequency noises which may affect clock behavior, especially when considering long-term performance in holdover mode.

To overcome such a convergence problem, the two sample Allan Deviation (ADEV) and its modified version Time Deviation (TDEV) were defined [3] as standard deviations of filtered data, according to:

$$\mathrm{ADEV}(\tau) = \sqrt{\frac{1}{2\tau^2(N-2n)}\sum_{i=1}^{N-2n}(x_{i+2n}-2x_{i+n}+x_i)^2}$$

$$\mathrm{TDEV}(\tau) = \sqrt{\frac{1}{6n^2(N-3n+1)}\sum_{j=1}^{N-3n+1}\left[\sum_{i=j}^{n+j-1}(x_{i+2n}-2x_{i+n}+x_i)\right]^2}$$

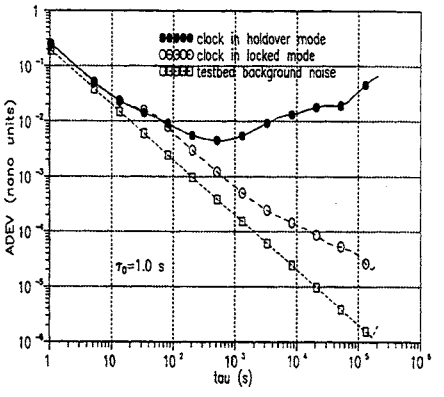


Figure 10: Allan Deviation (ADEV)

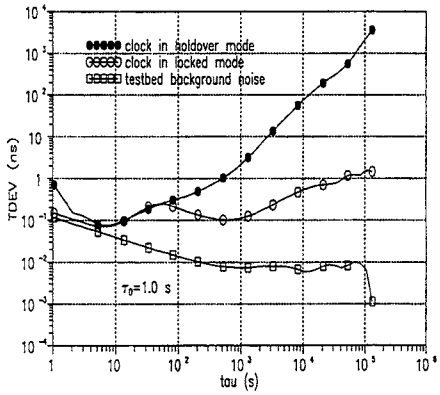


Figure 11: Time Deviation (TDEV)

Figs. 10 and 11 report ADEV and TDEV evaluated from the TE measured sequences. ADEV and TDEV filter out low frequency noise components and any frequency bias, give stable results in any typical practical situation and depict more clearly the relevant characteristics of the underlying phase noise processes.

Discussion

Some comments arise from the examination of the reported results.

i) *MTIE:* according to its current definition based on the detection of the peak value, MTIE strictly depends on the particular sequence and can lead to significantly different results when two or more sequences measured in the same clock operating mode are considered; in the limit case of a sequence with an infinite number of samples, MTIE could reach an infinitely large value. As a consequence, this implicitly suggests that the specification of clock requirements in terms of a peak-value parameter as MTIE should possibly be reconsidered adding, for example, a percentage information, to account for large TE fluctuations occasionally occurring when arbitrarily long measured sequences are considered.

Moreover, in the processing of the same sequence of data for different τ values, MTIE is a monotonically increasing function, because once a peak value is detected, it dictates all the following MTIE values till a larger peak is encountered. It seems therefore difficult to find an immediate relationship between the MTIE behavior and the spectral characteristics of the underlying noise.

ii) *Effect of frequency drift:* from the measurement performed while the SASE is in holdover state, the frequency drift effect is clearly visible (Fig. 7). Such a dominating effect comes out also from the trends of the evaluated statistical parameters: in fact, MTIE and TIE rms show a similar behavior, here proportional to τ [3]. The frequency drift effect is also visible in the last part of the ADEV and TDEV graphs (longer observation times), being the plots respectively proportional to τ and τ^2 as expected [3],[7].

iii) Effect of sinusoidal FM modulation: the behavior of the statistical parameters when the clock is in locked mode suggests the presence of a sinusoidal FM modulation. In fact, it may be worth mentioning that in presence of a sinusoidal frequency modulation written as

$$y = \frac{dx}{dt} = a\cos(2\pi f_m t)$$

where y is the normalised frequency deviation, a is its peak value, and f_m is the modulating frequency, the behaviors of TIE rms, ADEV and TDEV are expressible as:

$$\mathrm{TIE}_{rms}(\tau) = a\tau\frac{\left|\sin(\pi f_m \tau)\right|}{\pi f_m \tau}$$

$$\mathrm{ADEV}(\tau) = a\frac{\sin^2(\pi f_m \tau)}{\pi f_m \tau}$$

$$\mathrm{TDEV}(\tau) = \frac{a\tau}{\sqrt{3}}\left|\frac{\sin^3(\pi f_m \tau)}{(\pi f_m \tau)\cdot n\sin(\pi f_m \tau / n)}\right|$$

with a characteristic 'bump' when τ equals approximately $T_m/2$, where $T_m = 1/f_m$ is the modulation period.

Such a 'bump' can clearly be evinced in Fig. 9, 10, and 11, leading to the estimation of a periodic modulation with T_m almost equal to 140 s. Assuming the SASE behavior can be modeled as a simple second order phase lock loop (PLL), this sinusoidal modulation could well be due to the overshoot of the second order phase-lock loop peaked near its open-loop unity gain Fourier frequency, which is strongly related to the loop time constant (externally selected as 100 s in the SASE under test).

It follows that the characterisation of a clock in terms of TIE rms or ADEV or TDEV may also be interpreted as a tool to detect an unwanted sinusoidal modulation or to measure the extent of transfer of such a modulation through the clock itself [3].

iv) *Effect of PLL:* following the treatment reported in [8] it can be proved that precise relationships exist between typical noises in an autonomous clock and its output in locked mode. For example, independently of the noise type at the input, a clock locked to an ideal reference should exhibit an ADEV plot always proportional to τ^{-1}. The detection of such a behavior indicates the time constant of the PLL, if not known, and is a clear indication that the loop is working well. In the same situation TIE rms should always be flat ($\propto \tau^0$). By inspection of Figs. 9 and 10, it is clearly confirmed the 100 s value for the time constant in the clock under test; this also reflects the interpretation given in the previous paragraph for the sinusoidal modulation affecting the observed TE behavior.

Conclusions

Stability performance in a commercially available SASE unit has been characterised in terms of the MTIE, TIE rms, ADEV and TDEV parameters, both in the locked and holdover modes of operation.

The clock under test appears to be fully compliant with current MTIE requirements in international telecommunications standards [6], especially with respect to the locked mode of operation, being the actual clock performance much better than required (20 ns measured MTIE over 5 days, against 1000 ns required).

The impact of frequency drift and spurious sinusoidal modulations on phase variation has also been discussed, directly showing their effect on measured stability parameters as an example of common physical effects affecting actual clock performance in a typical SASE unit.

Acknowledgments

The work of Roberto Bonello and Antonio Manzalini was carried out in the framework of the national study group on synchronisation (GSS) composed by CSELT, Fondazione Ugo Bordoni, Sirti and Telecom Italia. The help of the other partners is kindly acknowledged.

References

[1] M. Sexton, A. Reid, *Transmission networking: SONET and the Synchronous Digital Hierarchy*. Artech House, Boston, USA, 1992.

[2] ITU-T Recommendation G.803, "Architectures of transport networks based on the synchronous digital hierarchy", June 1991.

[3] ETSI TM3 - Draft ETS DE/TM-3017, "Generic requirements for synchronisation network", Oct. 1994.

[4] R. Bonello, S. Bregni, M. Carbonelli, A. Manzalini, A. Mariconda, D. De Seta, D. Perucchini, and L. Valtriani, "Preliminary experimental results for the development of the new italian synchronization network", *Proc. ICC '95*, to be held in Seattle, USA, June 1995.

[5] R. Bonello, M. Carbonelli, D. De Seta, A. Manzalini, and D. Perucchini, "Impact of measurement configurations on the behaviour of time stability quantities", *Proc. EFTF '95*, Besançon, France, Mar. 1995.

[6] ITU-T Recommendation G.812, "Timing requirements at the outputs of slave clocks suitable for plesiochronous operation of international digital links", Nov. 1988.

[7] J. Rutman, F. L. Walls, "Characterization of frequency stability in precision frequency sources", *Proc. IEEE*, vol. 79, n. 7, June 1991, pp. 952-960.

[8] A. Godone, E. Bava, and C. Novero, "Phase lock of submillimetric backward-wave-oscillators", *IEEE Trans. on Instr. and Meas.*, vol. 38, n.3, June 1989, pp 794-798.

AWARD SESSION

Fred listening to Award Presentation

Monsieur le Préfet de la Région Franche-Comté,
Mesdames, Messieurs, Chers Collègues,
Cher Dr Walls, ... Cher Fred

Mister Prefet, Prefet of Franche-Comté Region
Ladies and Gentlemen, Dear Colleagues,
Dear Dr Walls, ... Dear Fred

Le Comité Exécutif du Forum Européen nommément MM. AUDOIN, BESSON, ECABERT, GAGNEPAIN, KARTASCHOFF et SCHLUETER dans sa séance du 23 décembre 1994 a décerné le Prix Européen Temps-Fréquence au Dr Fred L. WALLS du National Institute for Standards and Technology of Boulder, Colorado.
Le Prix lui sera remis dans quelques instants par Monsieur le Préfet Jean-Louis DUFEIGNEUX, Préfet de la Région Franche-Comté en sa qualité de représentant de l'Etat français.

The Executive Committee of the EFTF namely MM. AUDOIN, BESSON, ECABERT, GAGNEPAIN, KARTASCHOFF and SCHLUETER has decided on December 23rd 1994 to select as the recipient of 1995 European "Time and Frequency" Award Dr Fred L. WALLS of National Institute for Standards and Technology in Boulder, Colorado.
The Award will be presented to him in a few moments by Mr Préfet Jean-Louis DUFEIGNEUX, Préfet de la Région Franche-Comté, acting as a representative of French Government.

Presentation of Dr Walls will be, for me today, a very difficult task. Basically, two reasons for that :

- Even, if I know Fred WALLS very well presentation of such a great scientist is almost an impossible job because of the enormous amount of work he has done and the exceptionnally numerous results he has obtained.

- Because I happen to know him very well I am deeply moved, when delivering to him the recognition message of our Scientific Community.

Let me briefly present Fred to you. Please notice that, even if my English is poor, I did not say introduce Fred to you because he is worldwide known and recognized as a great scientist.

Fred L. WALLS was born in Portland, OR, on October 29, 1940. He received the B.S., M.S., and Ph. D. degrees in physics from the University of Washington, Seattle, respectively in 1962, 1964, and 1970. His Ph. D. thesis was on the development of long-term storage and nondestructive detection techniques for electrons stored in Penning traps and the first measurements of the anomalous magnetic (g-2) moment of low energy electrons. From 1970 to 1973, he was a Postdoctoral Fellow at the Joint Institute for Laboratory Astrophysics in Boulder, CO. This work focused on developing techniques for long-term storage and nondestructive detection of fragile atomic ions. Since 1973, he has been a Staff Member of the Time and Frequency Division of the National Institute of Standards and Technology, formerly the National Bureau of Standards, in Boulder. He is presently Group Leader of the Phase Noise Group and is engaged in research and development of hydrogen maser devices and cesium beam standards for use as ultra-stable clocks, oscillators, frequency synthesis from RF to infrared, low-noise frequency stability measurement systems, and accurate phase and amplitude noise metrology. He has published more than 100 scientific papers and holds 5 patents. He has received two silver medals from the US Department of Commerce for his scientific work.
Dr Walls is a member of th American Physical Society and a senior member of the IEEE. He was the 1988 distinguished IEEE lecturer for the Instrumentation and Measurement society.

I shall not describe here the subject of more than 110 papers and 5 patents but rather point out the extraordinary diversity of Dr Fred WALLS scientific work ranging from quartz resonators and oscillators to atomic frequency standards physics.

Before going to reading of citation I would like to recall here that Fred is also a loving and caring husband and father. So I have a special thought for his spouse Fanda and his 3 kids including Warren whos is presently here, Jeremy and Corina.

Fred has also some more or less unknown qualities. He is the best friend you can imagine in the world, a really dependable friend. He has done a lot for scientific cooperation between Europe and America. He has been a student of H.G. DEHMELT recent Nobel Prize winner, and who knows? He might receive the Rabi Award one day, like another Nobel price winner did. As you can see he is on the right track. But he is also a wonderful fisherman tentatively even trying to teach me... and believe me that is difficult.

Back to our purpose I like now to read the citation before Mr Préfet DUFEIGNEUX presents the award to Fred. Citation says :

"For outstanding contributions to pioneer work in ions storage physics, design and development of passive hydrogen masers, measurements of phase noise in passive resonators, very low noise electronics and phase noise metrology".

Professeur Raymond J. BESSON
Président de la Société Française des
Microtechniques et de Chronométrie

THE QUEST TO UNDERSTAND AND REDUCE 1/f NOISE IN AMPLIFIERS AND BAW QUARTZ OSCILLATORS

Fred L. Walls
National Institute of Standards and Technology
Boulder, CO 80303 USA

ABSTRACT

This paper attempts to trace the highlights in the development of precision oscillators from a historical point of view and speculates on further improvements in their 1/f performance. Using Leeson's model the closed loop phase modulation (PM) noise and frequency stability can be estimated from the stability of the quartz resonator, the resonator quality factor, and the noise of the sustaining amplifier coupled with the stability of the oven and various environmental parameters. The inherent frequency stability of the oscillator is primarily limited by the 1/f PM noise of the BAW quartz resonator and the sustaining amplifier. This paper, therefore, focuses on describing the quest to characterize, understand, and reduce the 1/f PM noise of amplifiers and BAW resonators.

INTRODUCTION

Modern high-stability bulk acoustic wave (BAW) quartz oscillators have evolved through a long and tortuous path. In this paper I trace many of the highlights in the development of these oscillators and speculate on further improvements in their 1/f performance. It is inevitable that some of the important discoveries that have occurred during the past 40 years and some of the people involved, will be left out. For these inadvertent omissions, I ask your indulgence.

Figure 1 shows the basic elements of a quartz oscillator. Using Leeson's model, the closed loop phase modulation (PM) noise and frequency stability can be estimated from the stability of the quartz resonator, the resonator quality factor, and the noise of the sustaining amplifier coupled with the stability of the oven and various environmental parameters [1-6]. The inherent frequency stability of the oscillator is limited primarily by the 1/f PM noise of the BAW quartz resonator and the sustaining amplifier. At 5 MHz, for example, these processes usually dominate the frequency stability for Fourier frequencies from approximately 0.001 Hz to 30 Hz or for measurement times from roughly 0.01 s to 100 s. In several cases it has been shown that the 1/f frequency modulation (FM) noise of the resonator extends out to at least 10^6 s [4,7]. This paper, therefore, concentrates on the quest to characterize, understand (model), and reduce the 1/f PM noise of BAW resonators and amplifiers. To a large extent the descriptions outlined in this paper will follow the historical order of the developments.

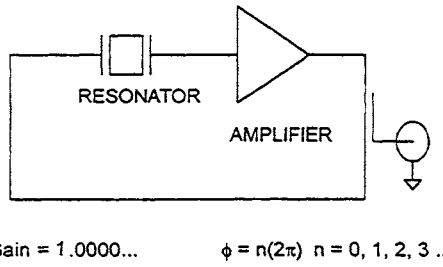


Fig. 1 *Conceptual diagram of an oscillator with the conditions for oscillation.*

LEESON'S MODEL OF AN OSCILLATOR

Leeson's model is used to estimate the random PM noise of an oscillator from the PM noise of the sustaining amplifier and the resonator. The agreement with experimental results is often better than 3 dB [2,3]. The basis for this model is that the phase around the oscillating loop, ϕ, is given by $\phi = 2N\pi$ with N = 0, 1, 2, 3, ... and the gain around the loop, G, is given by G = 1.00.... The consequence of this is that a phase shift, $d\phi$, anywhere within the oscillation loop causes a frequency change, dν, from resonance given by

$$\frac{d\nu}{\nu_o} = \frac{1}{2Q_L} d\phi \qquad (1)$$

where ν_0 is the resonance frequency of the resonator and Q_L is the loaded Q-factor of the

Contribution of the U.S. Government, not subject to U.S. copyright.

resonator. Squaring both sides and taking the Fourier transform of both sides yields

$$S_y^{osc}(f) = \frac{1}{4Q_L^2} S_\phi^{loop}(f) \qquad (2)$$

where $S_y^{osc}(f)$ is the PSD of fractional frequency fluctuations of the oscillator due to PM noise of $S_\phi^{loop}(f)$, within the loop. $S_y(f)$ can be expressed in terms of the PM noise of the oscillator as

$$S_y(f) = \frac{f^2}{\nu_o^2} S_\phi(f), \quad S_\phi(f) = 2\mathcal{L}(f). \qquad (3)$$

Equations 1-3 allows us to finally obtain the PM noise of the oscillator within the bandwidth of the resonator as

$$S_\phi^{osc}(f) = \frac{\nu_o^2}{4f^2 Q_L^2} S_\phi^{loop}(f). \qquad (4)$$

The (PM) noise in the amplifier can be modeled by

$$S_\phi^{amp}(f) = \frac{\alpha_E}{f} + \frac{2kFGT}{P}, \qquad (5)$$

where α_E is a phenomenological constant that characterizes the flicker PM noise coefficient for the amplifier, k is Boltzmann's constant, F is the noise figure, G is the gain, T is the temperature in Kelvin, and P the output power of the amplifier.

The PM noise in the resonator can be modeled by

$$S_\phi^R(f) = \frac{\alpha_R}{f} + \frac{2kT}{P_b}, \qquad (6)$$

where α_R is the flicker coefficient for the resonator, and P_b is the total power dissipated in the resonator and load. Combining these yields

$$S_\phi^{osc}(f) = \left(\frac{\nu_o}{2Q_L}\right)^2 \frac{1}{f^2}\left[\frac{\alpha_E}{f} + \frac{2kFGT}{P} + \frac{\alpha_R}{f} + \frac{2kT}{P_b}\right]$$

$$f<BW, \qquad (7)$$

and

$$S_\phi^{osc}(f) = \frac{\alpha_E}{f} + \frac{2kFGT}{P} \quad f > BW,$$

where the bandwidth of the sustaining loop, BW, is approximately $\nu_o/2Q_L$.

The time-domain frequency stability can be calculated from the phase noise using

$$\sigma_y^2 = \frac{2}{\pi\nu_o\tau^2} \int_0^{fh} S_\phi^{osc}(f) \sin^4(\pi f\tau)\, df, \qquad (8)$$

where f_h is the equivalent bandwidth of the measurement system [8]. Flicker PM noise in the resonator leads to a flicker FM in the oscillator and a fractional frequency stability that is independent of averaging time. Flicker PM noise in the amplifier at Fourier frequencies inside the loop bandwidth also leads to flicker FM in the oscillator. Flicker PM noise in the amplifier at Fourier frequencies larger than the loop bandwidth leads to flicker PM in the oscillator and a time domain stability that varies roughly as $1/\tau$, where τ is the measurement time. For flicker FM noise

$$\sigma_y^2 = \left(\frac{2\ln 2}{\nu_o^2}\right) S_\phi^{osc}(1Hz). \qquad (9)$$

DEVELOPMENT OF BAW RESONATORS AND OSCILLATORS WITH LOW 1/f NOISE

The first modern resonators with low noise and repeatable temperature performance were the 2.5 and 5 MHz 5th overtone AT-cut resonators developed by Art Warner of AT&T Bell Labs in the early 1950's [9]. Figure 2 shows that the Q-factors of these resonators were virtually as good as present day results. The time-domain frequency stability achieved in 1960 at 2.5 MHz is shown in Figs. 3 and 4 [9-11]. Measurements of frequency stability before the work of David Allan and Jim Barnes in 1966 [12-13] probably were analyzed using the standard deviation; moreover, the bandwidth of the measurement equipment was not often stated.

Allan and Barnes' work was pivotal in providing a clear understanding of the statistics of flicker FM noise and providing clear guidelines for time-domain frequency stability measurements which ensured reproducibility of the measurements. The frequency stability of the oscillators reported in [9-11] was dominated in the short-term by flicker PM noise in the electronics. At very long times temperature fluctuations were probably the dominant drivers.

At the time of this work it probably was not realized that the dynamic temperature coefficient of the AT-cut resonator was very large, typically

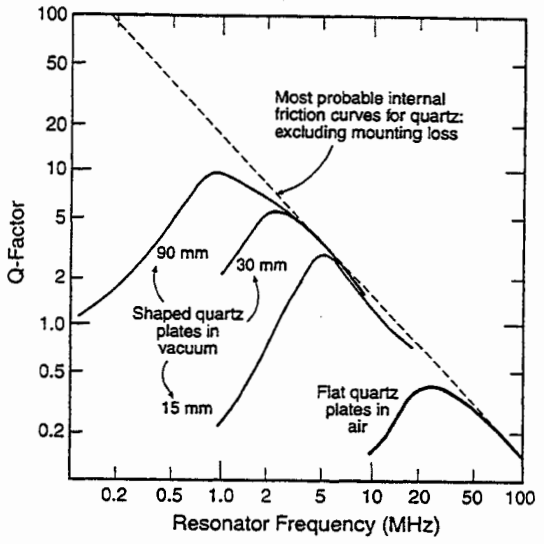


Fig. 2 *Resonator Q factor versus frequency for 15, 30, and 90 mm quartz plates in vacuum. The dashed line shows the product of Q V where V is the resonance frequency. Adapted from Warner, 1960 [9].*

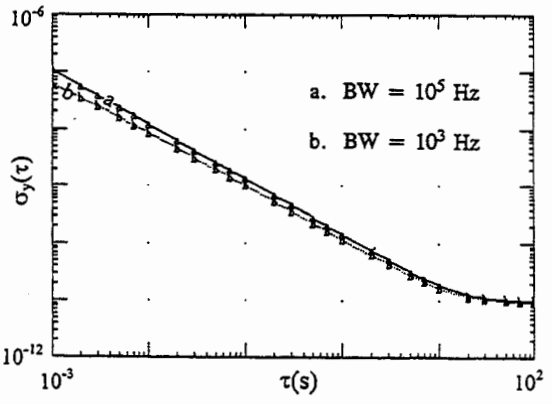


Fig. 3 *Fractional frequency stability versus measurement time for a 2.5 MHz 5th overtone AT-cut resonator for a measurement bandwidth of both 10^3 and 10^5 Hz. From Warner, 1960 [9].*

100 times larger than the static coefficients near turnover. Nevertheless, the long-term frequency stability was very good, as demonstrated by the results in Fig. 4 [11].

The next few years saw a steady improvement of about a factor of 20 dB in the phase noise and a factor of nearly 10 in the time-domain frequency stability as shown in Fig. 5 [14-17]. Virtually all of these oscillators used the new AT-cut resonators designed by Art Warner.

Another major milestone was the development of the Sulzer oscillators first using 2.5 and then 5 MHz 5th overtone resonators around 1968. It is

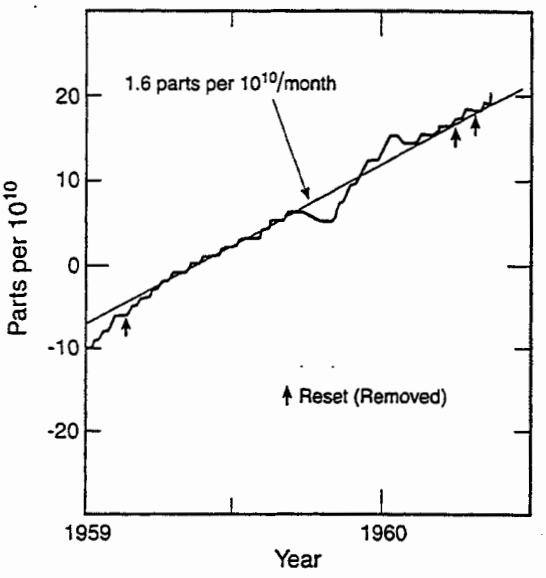


Fig. 4 *Fractional frequency stability of a 2.5 MHz oscillator vs. a cesium standard. Adapted from Anderson and Merrill, 1960 [11].*

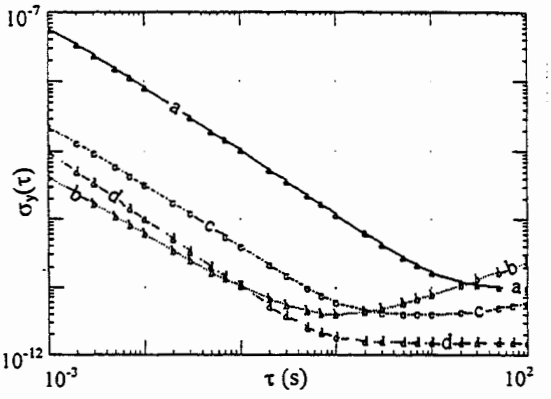


Fig. 5 *Fractional frequency stability of quartz oscillators in a 1 kHz bandwidth. Curve a from Warner, 1960 [9] is for a 2.5 MHz oscillator. Curve b from Stratemeyer, 1964 [15], curve c from Cutler and Searle, 1966 [16], and curve d from Pustarfi, 1966 [17] are 5 MHz oscillators.*

now clear from the circuit diagrams that Peter Sulzer understood much more about the PM noise in the electronics and the role of the dynamic temperature coefficient than any of his predecessors. His ovens had very high thermal gains, good insulation, and high thermal mass around the resonator. The thermal stability achieved at the resonator was approximately 20 nK/s. The result was another factor of 10 improvement in the short-term frequency stability as illustrated in Fig. 6 [18]. The drift of these oscillators after running for many months was typically parts in 10^{12}/day. The primary differences between these extraordinary oscillators

and many present day oscillators are their size, power, and higher PM flicker noise.

The pioneering work of Don Halford and Art Wainwright [19], who introduced an unbypassed resistor in the emitter leg of a transistor to reduce the phase modulation, was incorporated into an oscillator by Helmut Brandenberger [20]. The major advantage over the Sulzer oscillator was that the phase noise from approximately 0.5 Hz to 1 kHz was much lower as shown in Fig. 7. This yielded much better short-term frequency stability than the Sulzer oscillator for measurement times from about 0.001 to 3 s as shown in Fig. 8.

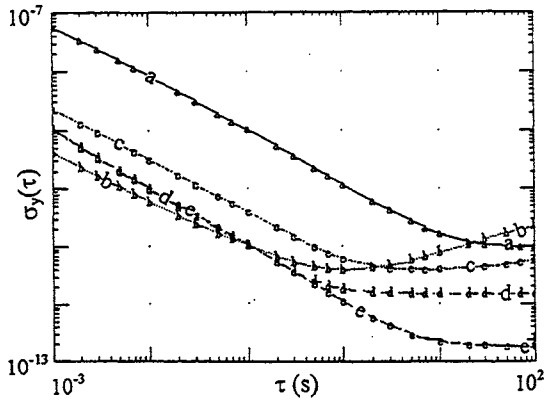


Fig. 6 *Fractional frequency stability of quartz oscillators in a 1 kHz bandwidth. Curve a from Warner, 1960 [9] is for a 2.5 MHz oscillator. Curve b from Stratemeyer, 1964 [15], curve c from Cutler and Searle, 1966 [16], curve d from Pustarfi, 1966 [17], and curve e from Sulzer, 1968 [18] are 5 MHz oscillators*

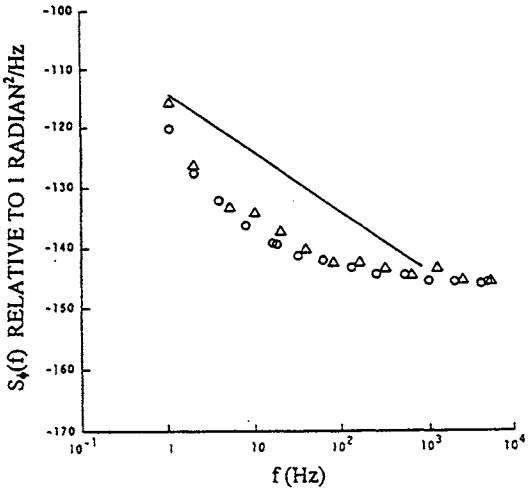


Fig. 7 *The data labeled Δ and O are the PM noise of the 5 MHz Brandenberger oscillator, 1971 [20]. The solid line is the estimated PM noise of a 5 MHz Sulzer oscillator [18].*

The wide band PM noise of approximately $S_\phi(f)$ = -142 dB rel to 1 rad^2/Hz and the long-term stability of 2-5 x 10^{-13} remained about the same.

I joined the Time and Frequency Division of the National Bureau of Standards in the fall of 1973 after the departure of Don Halford. Art Wainwright showed me many of the tricks of the trade. He and I collaborated on the development of the phase bridge method, schematically shown in Fig. 9, for measuring the noise of a resonator pair without the use of the oscillator electronics [21]. Our results showed that the 1/f PM noise of the resonators extended to much higher Fourier frequencies than was observed in quartz oscillators of the day. (See Figs. 10 and 11.) These results suggested that the wide band PM noise of oscillators could be reduced by at least 20 dB without increasing the power dissipated in the resonator.

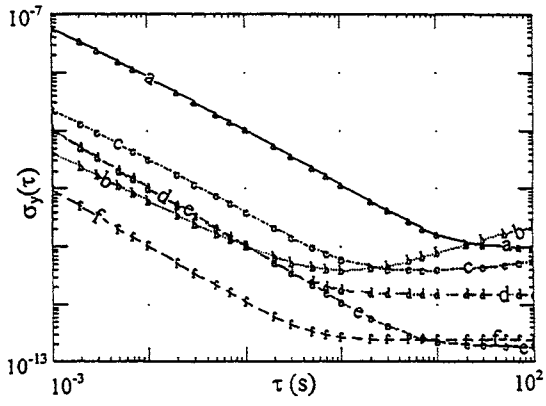


Fig. 8 *Fractional frequency stability of quartz oscillators in a 1 kHz bandwidth. Curve a from Warner, 1960 [9] is for a 2.5 MHz oscillator. Curve b from Stratemeyer, 1964 [15], curve c from Cutler and Searle, 1966 [16], curve d from Pustarfi, 1966 [17], curve e from Sulzer, 1968 [18], and curve f from Brandenberger, 1971 [20] are 5 MHz oscillators.*

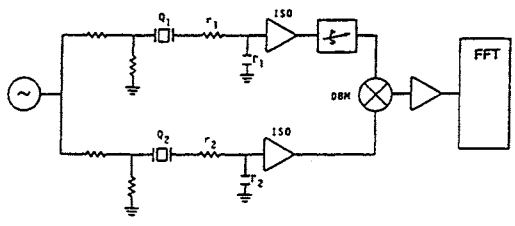


Fig. 9 *Block diagram of phase bridge for measuring the PM noise in resonator pairs. By equalizing the resonator frequencies and Q-factors, the contribution of PM noise in the oscillator to the measurement can be reduced by ~20 dB. From Walls and Wainwright, 1975 [21].*

This challenge of producing oscillators with reduced wide band PM noise was first taken up by Charles Stone around 1975. Over the course of the next several years he produced a series of oscillators from 5 MHz to 100 MHz using AT-cut resonators in which the flicker FM noise of the resonator dominated to much higher frequencies. Typically these oscillators had PM noise floors of less than -175 dBc/Hz . This was about 30 dB lower than previous oscillators and actually much better than the projections as shown in Fig. 12. The time-domain performance of a modern version of this design using SC-cut resonators is shown in Fig. 13 [22]. Oscillators with PM noise floors below -163 dBc/Hz are now commonly available.

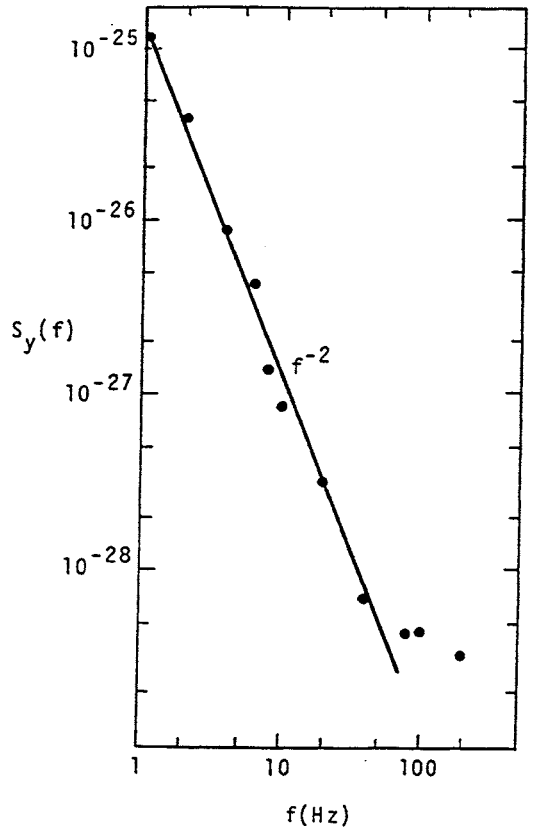


Fig. 10 *Direct phase bridge measurements of fractional frequency noise $S_y(f)$ in a pair of 5 MHz resonators with a Q-factor of 2.5 million. No correction for resonator bandwidth has been applied.*

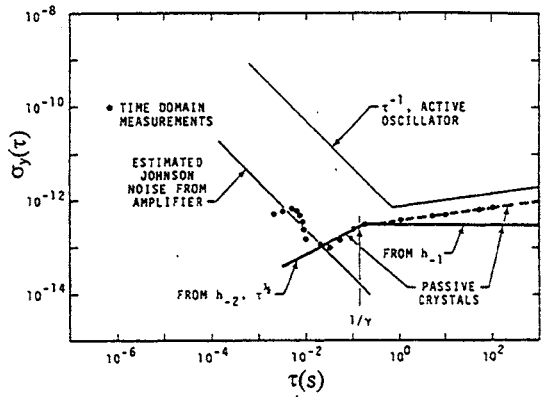


Fig. 11 *Comparison of measured $\sigma_y(\tau)$ vs. the value calculated from the measurement of $S_y(f)$ of the resonators from Fig. 10. From Walls and Wainwright, 1975 [21].*

A major innovation in the design of quartz resonators was the development of the BVA resonator by Raymond Besson [23]. In this new resonator, the electrodes were separated from the quartz by a few micrometers, and the oscillating portion was isolated from the mounting structure by very thin quartz bridges. The goal was to reduce the perturbations from the electrodes and the mounting structure on the frequency stability of the resonator. The idea of separating the electrodes from the resonator had been pursued earlier by Don Hammond but was not followed through to commercial production [24].

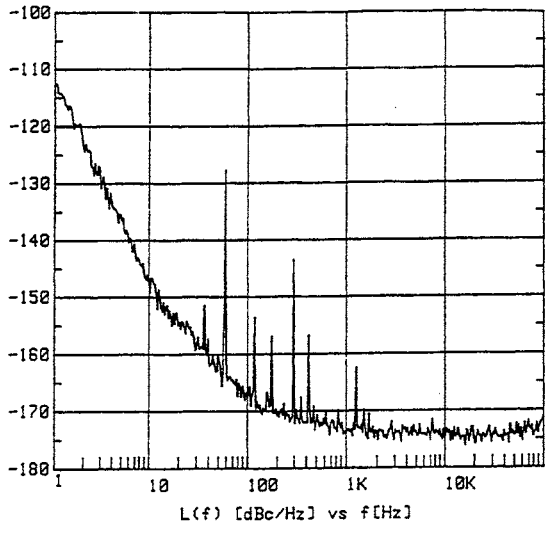


Fig. 12 *Phase noise of a high drive level 5 MHz oscillator. From Stone, ~1977.*

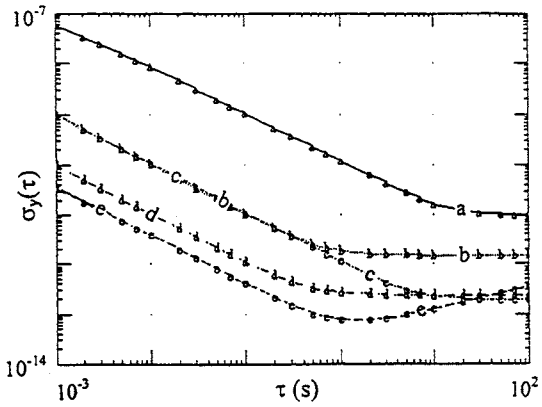


Fig. 13 *Fractional frequency stability of quartz oscillators in a 1 kHz bandwidth. Curve a from Warner, 1960 [9] is for a 2.5 MHz oscillator. Curve b from Pustarfi, 1966 [17], curve c from Sulzer, 1968 [18], curve d from Brandenberger, 1971 [20], and curve e from Stone, 1989 [22] are 5 MHz oscillators.*

In 1978 we were successful at NBS in demonstrating a frequency stability of approximately 8×10^{-14} using a first generation BVA resonator [25]. The approach, shown in Fig. 14, used the resonator in a reflectometer mode with the short-term stability provided by a Charles Stone oscillator. The BVA resonators of this vintage were troubled by mechanical hysteresis due to thermally induced stress between the three plates comprising the resonator. These results hinted at the potential for significant advances but were not practical due to resonator hysteresis and system complexity. Figure 15 shows a modern 10 MHz high-performance BVA resonator from Besson where the points of contact of the three plates are polished so flat that they become bonded by molecular forces [26]. This design eliminated the mechanical problems of the previous models. Unfortunately this resonator was not available at the time the reflectometer experiments were done.

The first oscillators to routinely achieve a fractional frequency stability below 1×10^{-13} were

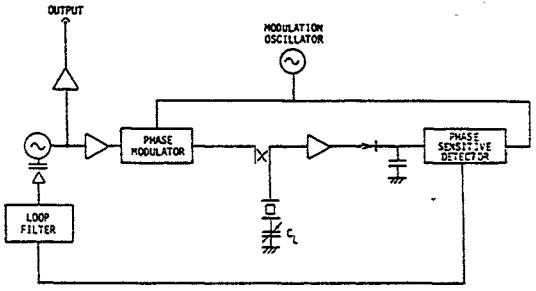


Fig. 14 *Phase modulation system for locking an oscillator to a passive quartz resonator. Adapted from Stein et. al., 1978 [25].*

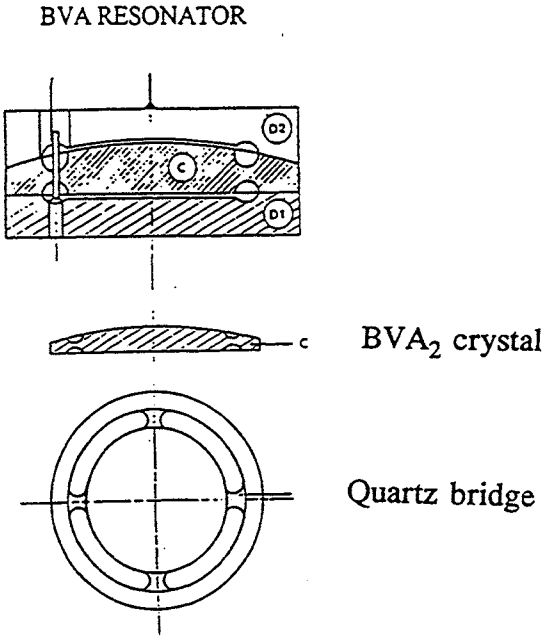


Fig. 15 *Schematic diagram of BVA2 quartz resonator. From Besson 1990.*

those designed by Jerry Norton for various space flights [27-29]. His designs were in many ways similar to those of Peter Sulzer. The oscillator circuits used the parallel mode resonance and very special attention was paid to the thermal design. The oven design of the Norton oscillators and their vacuum environment lead to a thermal performance that significantly outperformed previous oscillators. Table I shows the flicker FM performance of a number of 5 MHz AT-cut and BVA resonators in Norton oscillators. Note that oscillators using resonators with the same Q-factor differ significantly in their flicker FM noise. Still, quite a few of the resonators achieve a fractional frequency stability below 1×10^{-13}. In my opinion, these unique results are the direct result of a better temperature stability than that which was available in other oscillator designs. Obviously, the electronic design was also carefully done. Figure 16 shows the performance of 3 oscillators using BVA resonators [28,29].

The approximate thermal behavior of a traditional 5th overtone AT-cut resonator is given by

$$\frac{d\nu}{\nu_0} \cong 10^{-9}\ \Delta T^2\ 10^{-5}\ dT/dt, \qquad (10)$$

where ΔT is the temperature difference in K from turnover and dT/dt is the rate of change of temperature in K/s. To achieve a frequency stability of 3×10^{-14} requires that the temperature of the resonator change no faster than 3 nK/s.

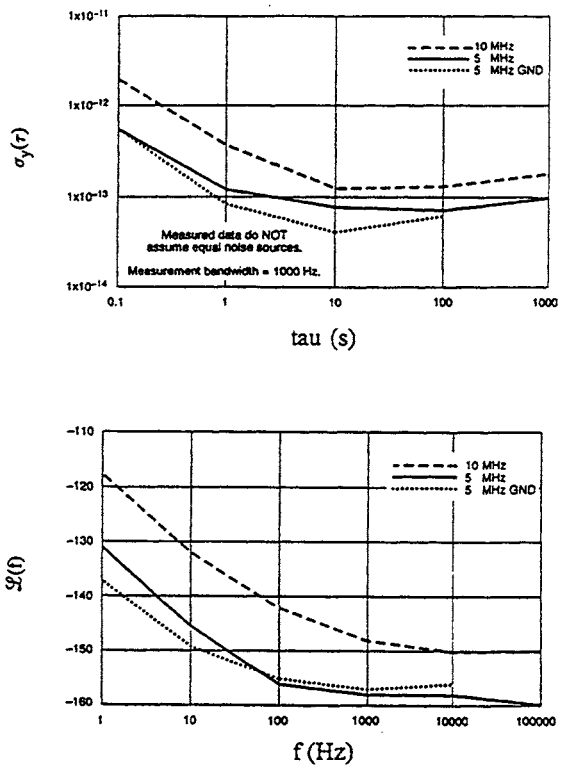


Fig. 16 *Frequency stability of BVA quartz oscillators. From Norton, 1994 [28].*

Table II shows the approximate frequency offsets because of temperature cycling of the oven due to the static thermal sensitivity of AT-cut resonators versus temperature offset of the oven from turnover. Unfortunately the dynamic temperature coefficient dominates these static temperature effects in most ovens, and in my opinion, explains the inability of many oscillator designs to achieve sub 1 x 10^{-13} frequency stability. It is not an accident that the advances of Sulzer and Norton both incorporated significant improvements in thermal design. Analysis and experiments show that the thermal stability of crystal ovens can be greatly improved [30-32]. The fundamental limit to thermal stability with present technology is of the order of a few nK/s and mK/year [33]. In benign thermal environments thermal fluctuations can be eliminated as a driver of instability in quartz oscillators.

One of the most important advances in resonator design was the prediction by Errol EerNisse in 1975 of the existence of a resonator cut that was much less sensitive to in-plane stresses and yet still maintained good temperature turnover characteristics [34]. The new stress compensated (SC)-cut resonator turned out to be the same as the thermal transient compensated (TTC)-cut of Jack Kusters. It was much less sensitive to stress and had a dynamic temperature sensitivity that was about a factor of 100 times smaller than the traditional AT-cut, when the angles were properly chosen [35,36]. Table III shows the approximate thermal performance of a 5 MHz SC-cut resonator. There was another very important advantage of this new design: it was much less sensitive to "activity dips." An activity dip occurs when harmonics of the primary mode couple with higher order modes. These interactions reduce the Q and disturb the phase of the primary mode [37]. The development of the SC-cut meant that the temperature fluctuations of traditional ovens had a much smaller effect on the frequency stability of the oscillator. Suddenly, many manufacturers could achieve a frequency stability of the order of parts in 10^{-13}.

Table IV shows the flicker FM noise achieved with SC-cut resonators by Stone [22], Norton [28], Besson [33], and Wenzel [38]. These results are, to my knowledge, the state-of-the-art for quartz crystal oscillators between 5 and 10 MHz.

How much further can the flicker FM noise of oscillators be reduced? To attempt an answer it is necessary to formulate a model for 1/f noise in resonators and the amplifiers used in the sustaining stage.

ANALYSIS OF 1/f NOISE IN RESONATORS

Jean-Jacques Gagnepain et al., [39] used the resonator phase bridge to study the intrinsic 1/f noise in resonators [21]. Their 1981 paper was a major milestone in attempting to develop theoretical models of 1/f noise in resonators. By varying temperature to change the Q-factor, they were able to eliminate the effects of many other parameters and show that the 1/f noise was somehow associated with the unloaded Q-factor of the resonator. (See Figs. 17-18.) They proposed that

$$S_y(f) = K_\nu/Q^4, \qquad (11)$$

where K_ν is a phenomenological constant that would need to be explained by a theory of 1/f noise. Since the product of intrinsic Q-factor and carrier frequency is a constant (see Fig. 2) [8], this would predict that $S_y(f)$ should scale as ν^4, where ν is the oscillator frequency.

Tom Parker compared the 1/f noise in BAW and surface acoustic wave (SAW) resonators and

suggested that the 1/f noise in all quartz resonators could be explained by

$$S_\phi(f) = K_\phi / Q^4, \qquad (12)$$

where K_ϕ is a phenomenological constant that would need to be explained by a theory of 1/f noise. See Fig. 19 [40]. In this case, $S_y(f)$ should scale as ν^2.

Our work with Peter Handel in 1992 suggested that the intrinsic noise in the resonator might be due to the loss of phonons out of the main mode due to scattering [41]. This quantum process leads naturally to a dependence on the volume of the resonating mode

$$S_y(f) = \beta \mathrm{Vol} / Q^4, \qquad (13)$$

where β is of order 1 if the volume is expressed in cm^3. This would imply $S_y(f)$ scaling as υ^3 for constant diameter electrodes, and as υ for the diameter of the electrodes scaling as $1/\upsilon$.

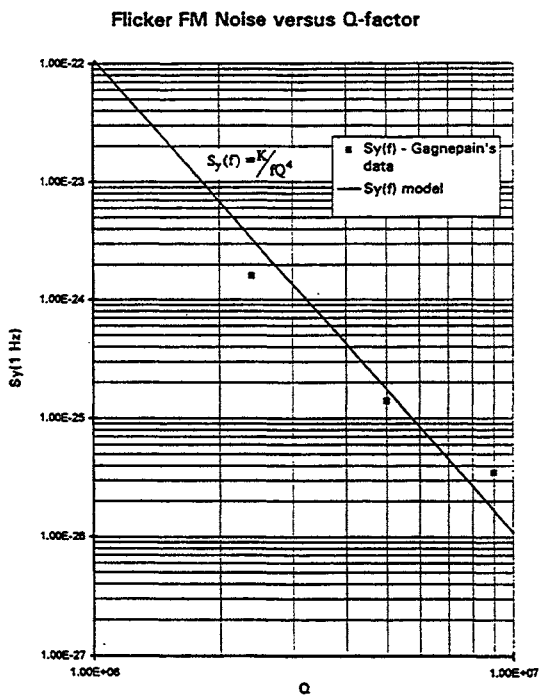


Fig. 17 *$S_y(f)$ versus Q-factor for a particular resonator. The Q was varied by changing temperature. From Gagnepain, et al., 1981 [39].*

Figure 20 shows the attempt to fit the data for a number of AT- and SC-cut resonators to the three models expressed by Eqs. (11-13). The data for each resonator are the best that have been reported for the particular frequency, manufacturer, and

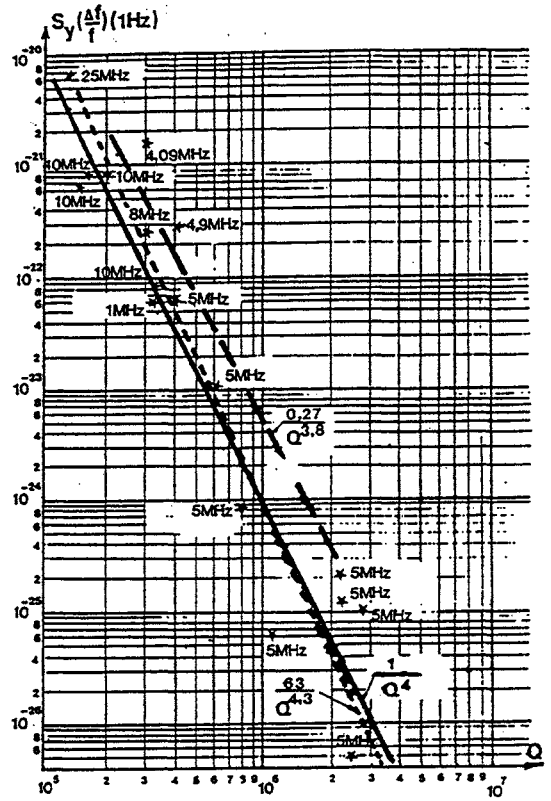


Fig. 18 *1/f frequency noise, 1 Hz from the carrier vs Q-factor. From Gagnepain, et al., 1981 [39].*

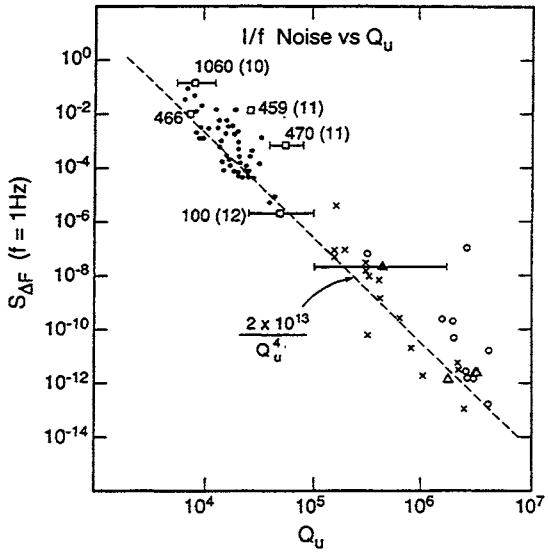


Fig. 19 *1/f noise levels of quartz Rayleigh wave and bulk acoustic wave resonators as a function of unloaded Q. Adapted from Parker, 1985 [40].*

resonator design. Only resonators where the unloaded Q-factor was within 20% of the material limit were considered. The data of Fig. 20 cover the frequency range from 2.5 MHz to 100 MHz for

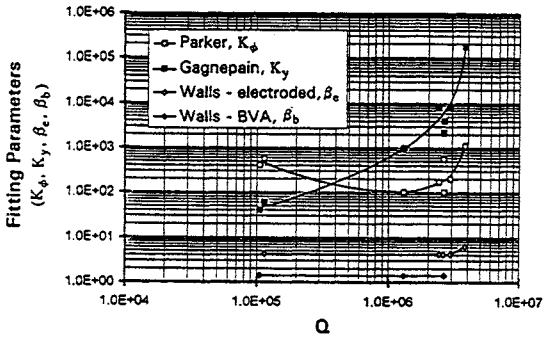


Fig. 20 *Fitting parameters for three proposed models of BAW quartz resonator noise versus unloaded Q-factor. From Walls, et. al [41].*

AT- and SC-cuts. It is very interesting that the data seems to fit the volume dependence model much better than any of the other models. This is in direct contrast to the results of Parker who found that the 1/f noise in SAW devices scaled as 1/vol [42] . The noise in the two types of resonators appears to originate from different mechanisms.

An interesting question is exactly how to calculate the volume. Is it the volume between the diameter of the electrodes or is it the weighted volume of the oscillating mode? Mike Driscoll and William Hansen attempted to address this problem in their 1993 paper by looking at the 1/f noise in several AT-cut, BT-cut, and SC-cut resonators in the frequency range from 40 to 160 MHz [43]. Their results on AT and SC-cut resonators seem to indicate that the acoustic volume is a better choice for use with Eq. (12), although even then the agreement was not very satisfactory. (See Fig. 21 and Tables V and VI.) Driscoll and Hansen's data fits Eq. (12) better than Eq. (13). The 1/f noise they measured on a BT-cut resonator was much higher than predicted by using either the electrode or the acoustic volume.

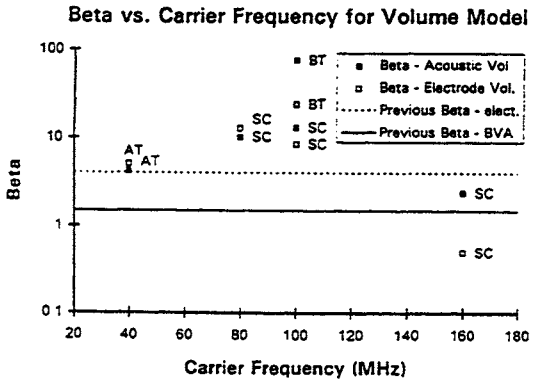


Fig. 21 *Fitting parameter for AT, SC, and BT-cut resonators. Adapted from Driscoll and Hansen, 1993 [43].*

Eva Ferre-Pikal measured the 1/f noise of 5th overtone 100 MHz SC-cut resonators. These resonators were made from the same quartz material by the same manufacturer for electrode areas differing by a factor of 2 above and below the standard size. Some of her unpublished 1993 results are shown in Fig. 22. The two resonators with the lowest 1/f noise have the smallest size electrodes, which is consistent with the volume dependence, or at least a dependence on electrode diameter. The wide dispersion between resonators with the same value of Q, however, clearly indicates that the 1/f noise must also depend on at least one other parameter besides Q-factor. The data of Table I lead to similar conclusions. This other factor must have a very large contribution to the 1/f noise and little effect on the Q-factor. Perhaps this other factor is non-linear coupling to the multitude of higher order modes mediated by stress due to crystal imperfections, the electrodes, and the mounting structure. In the resonators included in Fig. 20, there is a

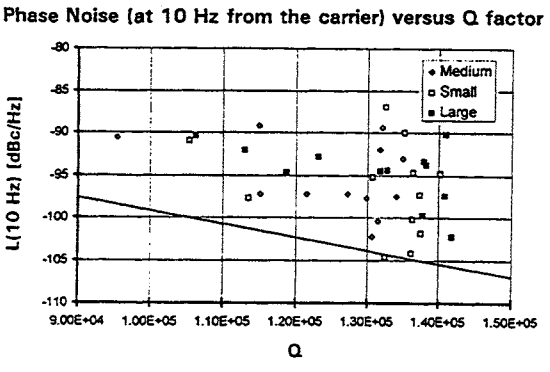


Fig. 22 *L(10 Hz) for 5th overtone 100 MHz resonators with small, medium, or large electrodes as a function of Q-factor. Unpublished data from Ferre-Pikal.*

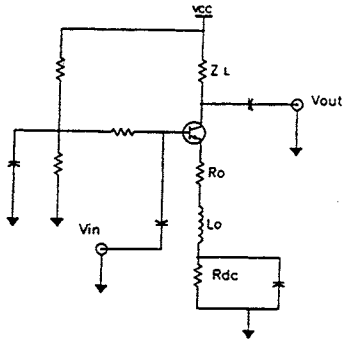
substantial difference between the 1/f noise in the electroded resonators and the BVA designs, which have no electrode on the resonator and greatly reduced mounting stress. Is this just an accident due to the small number of resonators studied and our inability to know the details of manufacturing, or is it an indication that the surface interface and/or the mounting stress contribute to the 1/f noise? We have additional work in progress to compare the 1/f noise in electroded resonators with the 1/f noise in BVA resonators with similar electrode areas and with resonators made with different manufacturing techniques to attempt to explore these additional noise effects further.

The bottom line at this point is that although there is some uncertainty, we can still assert that in general the 1/f noise improves as Q^4 and that the 1/f noise is better at lower oscillation frequencies than at higher oscillation frequencies. We do not understand the origin of 1/f noise in quartz crystal resonators, nor do we even understand how the 1/f noise should scale with frequency. We do, however, have a number of new questions, and we expect that the pursuit of answers to these questions will shed additional light on the problem.

ANALYSIS OF 1/f NOISE IN AMPLIFIERS

The first work that I am aware of, describing the 1/f PM noise in amplifiers was that of Don Halford and colleagues in 1968 [19]. They showed that the 1/f PM noise in amplifiers was due to an intrinsic modulation of the signal that could be reduced by up to 40 dB when negative feedback was used to reduce the rf gain. The most common method was to use an unbypassed resistor in the emitter of a BJT or source lead of a FET. (See Fig. 23.) The work of Svein Andresen showed the dependence of PM noise in doublers on emitter resister [44]. This work was, and still is, the basis for virtually all of the low PM noise designs in use today.

Common Emitter Amplifier



$$\text{Gain} = \frac{Z_L}{r_e' + R_o + \omega L_o}; \qquad r_e' = \frac{1}{g_m}$$

Fig. 23 *Schematic diagram of a common emitter BJT amplifier. Both PM and AM noise depend on the base current noise, base to collector voltage noise, bias, DC gain, and RF gain. In general, changes in gain lead to AM noise while impedance changes lead to PM noise.*

In 1992 we started a program at NIST to understand the relationship, if any, between AM and PM noise in amplifiers and in oscillators. Figures 24-26 show the results of some of our measurements [45,46]. These data show that the AM noise in amplifiers is not just the thermal noise that was previously assumed, but it also had a 1/f character that was remarkably similar to the 1/f PM noise. The AM and PM noise in many oscillators is similar outside the bandwidth of the resonator. Could both the AM and PM noise be due to the same process? (See Figs. 25-26.) We then embarked on a systematic study of the up conversion of the baseband noise of bipolar junction transistors (BJTs) to produce PM and AM noise in linear amplifiers. We have found that the AM noise can be higher than, equal to, or lower than the PM noise, depending on the circuit configuration. Our studies, undertaken primarily by Eva Ferre-Pikal, measured the 1/f components of current and voltage in the transistors and the resulting 1/f PM

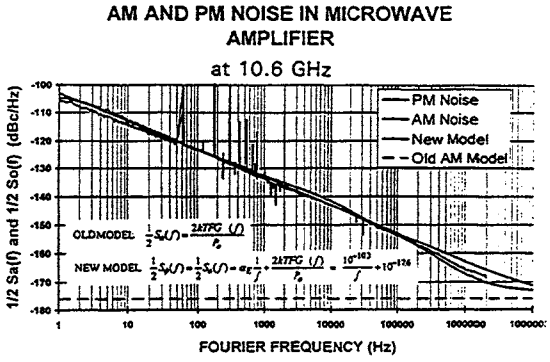


Fig. 24 *AM and PM noise of an x-band amplifier. From Ascarrunz, et. al., 1993 [45].*

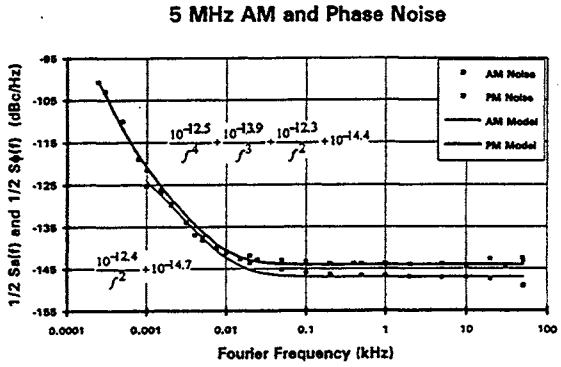


Fig. 25 *AM and PM noise of 5 MHz quartz oscillator. From Nelson, et al., 1994 [46].*

and AM noise for different circuit configurations [47,48]. From these data we have constructed a model for the dependence of both PM and AM noise in linear BJT amplifiers on baseband current and voltage noise and circuit configuration. Our results show that baseband noise modulation of the gain generates the AM noise, while baseband noise

modulation of the impedance at various terminals generates the PM noise. We find that the PM noise has a dependence on carrier frequency that is often higher than that of the AM noise.

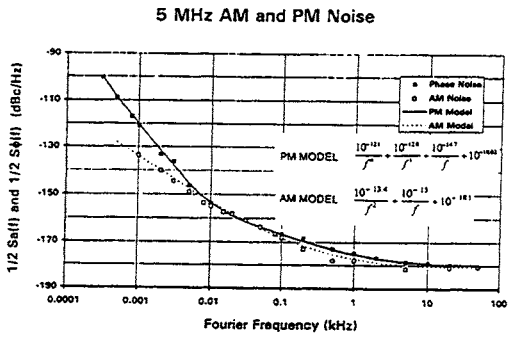


Fig. 26 *AM and PM noise of a 5 MHz oscillator. From Nelson, et. al., 1994 [46].*

Using this new model for PM and AM noise, we have modified a 5 channel distribution amplifier and achieved PM noise of L(10 Hz) = - 169 dBc/Hz ,which is within 2 dB of thermal noise [48,49]. (See Fig. 27.) The AM noise in the distribution amplifier is below our measurement noise. Ferre-Pikal has also made 5 MHz amplifiers with a gain of about 24 dB and PM noise of L(10 Hz) = -158 dBc/Hz. As in the previous example, the PM noise is within 1-2 dB of thermal noise, and the AM noise is below our measurement noise. Additional details are found in [47,48]. A low noise 100 MHz is discussed in [50].

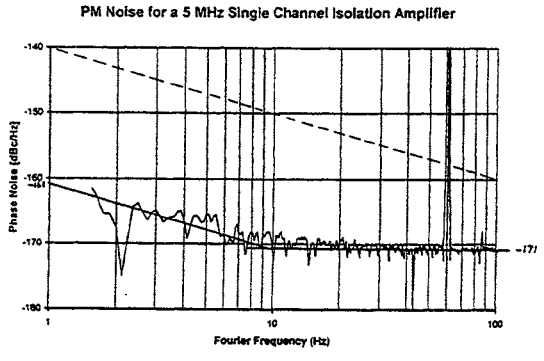


Fig. 27 *L(f) of a 5 MHz distribution amplifier [46,47]. The basic design of Nelson, et. al., 1994 [49] has been modified slightly to reduce the 1/f portion of the PM noise.*

What impact, if any, will amplifiers with lower 1/f noise have on quartz oscillators? Going back to Eq. (7), we see that as the 1/f noise in the resonator improves, the 1/f noise in the amplifier becomes more important.

$$\sigma_y^{osc2} = \frac{2\ln 2}{4Q_L^2} S_\phi^{amp}(1\text{Hz}), \qquad (14)$$

when flicker PM noise in the amplifier is the dominant noise process.

For example consider the phase noise data for the 5 MHz ground based Norton oscillator in Fig. 16. The bandwidth of this oscillator can be estimated to be approximately ± 0.75 to 1 Hz from the data on the resonator. The PM noise outside this range appears to be flicker PM. Using a value of L(10 Hz) = -149 dBc/Hz as a measure of the flicker PM noise in the amplifier, Eq. (14) indicates an oscillator flicker FM of 4.6×10^{-14}, assuming all of the 1/f PM noise originates in the sustaining stage. This result is within 1-2 dB of the observed value and indicates that flicker PM in the sustaining amplifier probably dominates the flicker FM of the oscillator. To achieve a 1/f FM noise floor in the oscillator of 1×10^{-14} requires that the 1/f PM noise in the sustaining amplifier be less than L(1 Hz) = -153 dBc/Hz for a 5 MHz carrier. Using our new model of AM and PM noise in BJT amplifiers it should be possible to reduce the contribution of the electronics well below this value, perhaps as low as L(1) = - 159 dBc/Hz which would be sufficient to reduce the amplifier contribution to approximately 5×10^{-15} at 5 MHz.

It is likely that the performance observed for the 10 MHz SC-cut tactical BVA based oscillators is also limited by the PM noise of the sustaining stage. If this is correct, reducing the PM noise in the sustaining stage should immediately make possible 10 MHz oscillators with a frequency stability floor of less than 1×10^{-13}.

DISCUSSION

Figure 28 shows that the short-term frequency stability of BAW crystal oscillators has improved dramatically over the past 35 years. These improvements have been due to significant reductions in the 1/f PM noise in the sustaining amplifiers, the introduction of the SC-cut resonator and the BVA resonators, and the development of better ovens. These improvements now make possible BAW crystal oscillators with a flicker FM noise of at least 4×10^{-14}. Further improvements in frequency stability should be possible using sustaining amplifiers with still lower 1/f noise. Amplifiers with much lower 1/f PM noise than previously available have been demonstrated within just the past few months [47,51].

A breakthrough in understanding 1/f noise in resonators will, in my opinion, require a reexamination of the effects of stress and nonlinear coupling to the various modes of the resonator. Thermal effects limit the frequency stability of many oscillators; however, the technology to remove temperature effects on quartz oscillators in benign environments is presently available [6, 30-32].

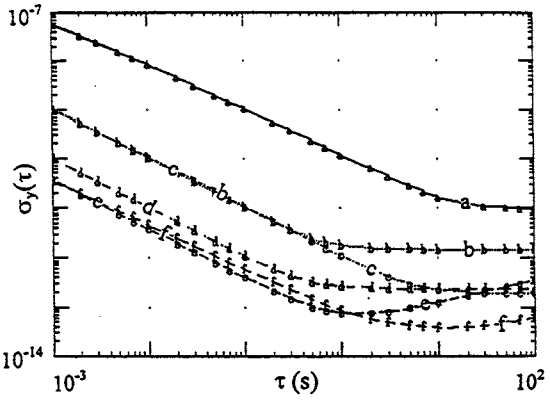


Fig. 28 *Summary of advances in 1/f frequency stability of BAW quartz oscillators from 1960 to 1995. Fractional frequency stabilities shown are for a bandwidth of 1 kHz bandwidth. Curve a from Warner, 1960 [9] is for a 2.5 MHz oscillator. Curve b from Pustarfi, 1966 [17], curve c from Sulzer, 1968 [18], curve d from Brandenberger, 1971 [20], curve e from Stone, 1989 [22], and curve f from and Norton, 1994, [28] are 5 MHz oscillators.*

ACKNOWLEDGMENTS

Over the past 22 years I have benefited greatly from interactions with a large number of colleagues. Discussions with Jean-Jacques Gagnepain, Raymond Besson, Jean Ubersfeld, Charles Stone, Art Wainwright, Andrea DeMarchi, Tom Parker, John Vig, David Allan, Jerry Norton, Eva Ferre-Pikal, Steve Jefferts, and Peter Handel have been especially helpful in my personal quest to understand 1/f noise. I am also grateful to Eva Ferre-Pikal for the use of her unpublished work and Gwen Bennett who helped prepare the manuscript.

REFERENCES

[1] D.B. Leeson, "A simple model of feedback oscillator noise spectrum," Proc. IEEE, vol. 54, 1966, pp. 329-330.

[2] G.S. Curtis, "The relationship between resonator and oscillator noise and resonator noise measurement techniques," Proc. 41st Ann. Symp. Freq. Contr., 1987, pp. 420-428.

[3] T.E. Parker, "Characteristics and sources of phase noise in stable oscillators," Proc. 41st Ann. Symp. Freq. Contr., 1987, pp. 99-110.

[4] F.L. Walls, "The influence of pressure and humidity on the medium and long-term frequency stability of quartz oscillators," Proc. 42nd Ann. Symp. Freq. Contr., 1988, pp. 420-428.

[5] F. L. Walls and J.-J. Gagnepain, "Environmental sensitivities of quartz oscillators, " IEEE Trans. Ultrason,. Ferroelec., and Freq. Contr., vol. 39, pp. 241-249, 1992.

[6] F.L. Walls and J. Vig, "Fundamental limit on the frequency instabilities of quartz crystal oscillators," to be published in Proc. IEEE Trans. UFFC, 1995.

[7] W.R. Attkinson, L. Fey, and J. Newman, "Spectrum analysis of extremely low frequency variations of quartz oscillators," Proc. IEEE, 1963, pp. 379.

[8] D.B. Sullivan, D.W. Allan, D.A. Howe, and F.L. Walls, "Characteristics of clocks and oscillators," NIST Technical Note 1337, pp. TN1-TN342, 1990.

[9] A.W. Warner, "Ultra precise quartz crystal frequency standards," IRE Tran. Instrum., vol. I-7, pp. 185-188,1958; A.W. Warner, "High-Frequency Crystal Units for Primary Frequency Standards," Proc. IRE, vol. 40, pp. 1030-1033, September 1952.

[10] A.W. Warner, "Design and performance of an ultra precise 2.5 megacycle quartz unit," Bell Sys. Tech. J., vol. 39, pp. 1193-1217, 1960.

[11] T.C. Anderson and F.G. Merrill, "Crystal-controlled primary frequency standards: Latest advances for long-term stability," IRE Trans. Inst., vol. I-9, pp. 136-141, 1960.

[12] D.W Allan, "Statistics of atomic frequency standards," Proc. IEEE, vol. 54,1966, pp. 221-230.

[13] J.A. Barnes and D.W. Allan, "A statistical model of flicker noise," Proc. IEEE, vol. 54, 1966, pp. 176-178.

[14] W. L. Smith, "Miniature transistorized crystal-controlled precision oscillators," IRE Trans. Inst., vol. I-9, pp. 141-148, 1960

[15] H.P. Stratemeyer, "The stability of Standard Frequency Oscillators," The General Radio Experimenter, vol. 38, pp. 1-6, 1964.

[16] L.Cutler and C. Searle, "Some Aspects of atomic frequency standards," Proc. IEEE vol. 54, 1966, pp. 136-154.

[17] H.S. Pustarfi, "An improved 5 MHz reference oscillator for time and frequency applications," IEEE Trans. Instrum. Meas., vol. IM-15, pp. 196-202, 1966.

[18] D.J. Glaze, "Improvements in atomic cesium beam frequency standards at the National Bureau of Standards," IEEE Trans. Instrum. Meas., vol. IM-19, pp. 156-160, 1970.
[19] D. Halford, A.E. Wainwright, and J.A. Barnes, "Flicker noise of the phase in rf amplifiers and frequency multipliers: characterization, cause, and cure," Proc. 22nd Ann. Symp. Freq. Contr, 1968, pp. 340-341.
[20] H. Brandenberger, F. Hadorn, D. Halford, and J.H. Shoaf, "High quality quartz crystal oscillators: frequency domain and time domain stability," Proc. 25th Ann. Symp. Freq. Contr, 1971, pp. 226-227.
[21] F.L. Walls, and A.E. Wainwright, "Measurement of the short-term stability of quartz crystal resonators and the implication for crystal oscillator design and applications," IEEE Trans. Instrum. Meas., vol. IM-24, pp. 15-20, 1975.
[22] M. B. Block, J.C. Ho, C.S. Stone, A. Syed, and F.L. Walls, "Stability of high quality quartz crystal oscillators: an update," Proc. 43rd Ann. Symp. Freq. Contr.,1989, pp. 80-84.
[23] R. Besson. "A new piezoelectric resonator design, " Proc. 30th Ann. Symp. Freq. Contr., 1976, pp. 78-83.
[24] D.L. Hammond, and L. S. Cutler, "Crystal resonators," U.S. Pat. No. 3,339,091, Aug. 29, 1967, and Reissue 26, 707, Nov. 4, 1969.
[25] S.R. Stein, C.M. Manney Jr., F.L. Walls, James E. Gray, and Raymond J. Besson, "A frequency-lock system for improved quartz crystal oscillator performance," IEEE Trans. Instrum. Meas., vol. IM-27, pp. 249-252, 1978.
[26] J.R. Norton, R.J. Besson, "Tactical BVA quartz resonator performance," Proc. 47th Ann. Freq. Contr. Symp., 1993, pp. 609-613.
[27] J.R. Norton, "Ultrastable quartz oscillator for spacecraft," Proc. 21st Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 1989, pp. 509-518.
[28] J.R. Norton, "Performance of ultrastable quartz oscillators using BVA resonators," Proc. 8th European Frequency and Time Forum, Munich, 1994, pp. 457-465.
[29] J.R. Norton, "BVA type quartz oscillator for spacecraft," Proc. 45th Ann. Freq. Contr. Symp., 1991, pp. 426-430.
[30] G. Marianneau and J.-J. Gagnepain, "Digital temperature control for ultrastable quartz oscillators," Proc. 34th Ann. Symp. Freq. Contr., 1980, pp. 52-57.
[31] F.L. Walls, "Analysis of high performance compensated thermal enclosures," Proc. 41st Annual Symp. Freq. Contr., 1987, pp. 439-443.
[32] R. Brendel, G. Marianneau, F. Djian, and E. Robert, "Improved OCXO's oven using active thermal insulation," Proc. IEEE UFFC, vol. 41,1994, pp. 269-274.
[33] R. Besson, "Recent advances in the field of ultrastable resonators and oscillators," Proc. IEEE International Symposium on Surface Waves in Solid and Layered Structures and National Conference on Acoustoelectronics, 1994.
[34] E.P. EerNisse, "Quartz resonator frequency shifts arising from the electrode stress," Proc. 29th Ann. Symp. Freq. Contr., 1975, pp. 1-4.
[35] E.P. EerNisse, "Calculation of the stress compensated (SC-Cut) quartz resonator," Proc. 30th Ann. Symp. Freq. Contr., 1976, pp. 8-11.
[36] J.A. Kusters, "Transient thermal compensation for quartz resonators, "IEEE Trans. Sonics Ultrasonics, vol. SU-23, pp. 273-276, 1976. See also J.A. Kusters and C.A. Adams, "Production statistics of SC (or TTC) crystals," Proc. 34th Ann. Symp. Freq. Contr., 1980, pp. 167-174.
[37] F.L. Walls and J.-J. Gagnepain, "Environmental sensitivities of quartz oscillators," Proc. IEEE Trans. UFFC, vol. 9, 1992, pp. 241-249.
[38] C. Wenzel, "New oscillators advance the art of low noise performance," RF Design, pp. 51-52, Jan. 1993.
[39] J.-J. Gagnepain, J. Ubersfeld, G. Goujon, and P. Handel, "Relationship between 1/f noise and Q factor in quartz resonators at room and low temperatures, first theoretical interpretations," Proc. 35th Ann. Symp. Freq. Contr., 1981, pp. 476-483.
[40] T.E. Parker, "1/f frequency fluctuations in quartz acoustic resonators," Appl .Phys. Lett., vol. 46, pp. 246-248, 1985.
[41] F.L. Walls, P.H. Handel, R. Besson, and J.-J. Gagnepain, "A new model of 1/f noise in BAW quartz resonators," Proc. 1992 IEEE Freq. Contr. Symp., 1992, pp. 327-333.
[42] T.E. Parker and D. Andres, "Dependence of SAW resonator 1/f noise of device size," Proc. IEEE 47th Ann. Freq. Contr. Symp., 1993, pp. 178-185.
[43] M.M. Driscoll and W.P. Hansen, "Measured VS volume model-predictor-flicker-of-frequency instability in VHF quartz crystal resonators," Proc. IEEE 47th Ann. Freq. Contr. Symp., 1993, pp. 186-192.
[44] S.G. Andresen and J.K. Nesheim, "Phase noise of various frequency doublers," Proc. IEEE, Trans. Instrum. Meas., 1973, pp. 339-342.
[45] F.G. Ascarrunz, E.S. Ferre, and F.L. Walls, "Investigations of AM and PM noise in X-band devices," Proc. IEEE 47th Ann. Freq. Contr. Symp., Salt Lake City, Utah, 1993, pp. 303-311.

[46] L.M. Nelson, C.W. Nelson, and F.L. Walls, "Relationship of AM and PM noise in selected rf oscillators, Proc. IEEE Trans. UFFC, vol. 41, 1994, pp. 680-684.
[47] F.L. Walls, E.S. Ferre-Pikal and S.R. Jeffers, "The origin of 1/f PM and AM noise in bipolar junction transistor amplifiers," Proc. 1995 IEEE Freq. Contr. Symp., 1995, to be published.
[48] E.S. Ferre-Pikal, F.L. Walls and C.W. Nelson, "Design criteria for amplifiers with low 1/f AM and PM noise," Proc. IEEE 1995 Freq. Contr. Symp., 1995, to be published.
[49] C.W. Nelson, F.L. Walls, M. Siccardi, and A. De Marchi, "A new 5 and 10 MHz high isolation distribution amplifier," Proc. 1994 IEEE Freq. Contr. Symp., 1994, pp. 567-571.
[50] M. Siccardi, S. Romisch, F.L. Walls, and A. De Marchi, "A low noise 100 MHz sidtribution amplifier for precision metrology," Proc. 9th European Frequency and Time Forum, Besancon, 1995.

Table. I. 1/f frequency stability (column 10) along with frequency, Q-factors, and other parameters of resonators tested in oscillators, from Norton [26-29]. Beta (column 11) is from the model of Walls, et. al., Eq 13.

Nominal Freq (MHz)	Q (x10^6)	Type	Overtone	Diam (cm)	Elec Diam (cm)	Thick (cm)	Volume (cm3)	SQR(V/Q^4)(2ln2)	stab (1/f)	Beta
5	2.0	elec	3rd	1.5	0.79	0.114	0.055	6.90E-14	7.00E-14	1.03
5	1.9	elec	3rd	1.5	0.79	0.114	0.055	7.65E-14	2.80E-13	13.40
5	2.35	elec	3rd	1.5	0.79	0.114	0.055	5.00E-14	1.70E-13	11.56
5	2.22	elec	3rd	1.5	0.79	0.114	0.055	5.60E-14	1.40E-13	6.24
5	2.92	elec	3rd	1.5	0.79	0.114	0.055	3.24E-14	6.60E-14	4.15
5	2.89	elec	3rd	1.5	0.79	0.114	0.055	3.31E-14	2.20E-13	44.28
5	3.09	elec	3rd	1.5	0.79	0.114	0.055	2.89E-14	1.00E-13	11.96
5	2.93	elec	3rd	1.5	0.79	0.114	0.055	3.22E-14	1.70E-13	27.94
5	2.86	elec	3rd	1.5	0.79	0.114	0.055	3.38E-14	1.20E-13	12.64
5	2.78	elec	3rd	1.5	0.79	0.114	0.055	3.57E-14	1.50E-13	17.63
5	2.75	elec	3rd	1.5	0.79	0.114	0.055	3.65E-14	1.20E-13	10.80
5	2.88	elec	3rd	1.5	0.79	0.114	0.055	3.33E-14	8.90E-14	7.15
5	2.7	elec	3rd	1.5	0.79	0.114	0.055	3.79E-14	2.00E-13	27.88
5	2.7	elec	3rd	1.5	0.79	0.114	0.055	3.79E-14	1.10E-13	8.43
5	2.7	elec	3rd	1.5	0.79	0.114	0.055	3.79E-14	1.40E-13	13.66
5	2.7	elec	3rd	1.5	0.79	0.114	0.055	3.79E-14	8.40E-14	4.92
5	2.81	elec	3rd	1.5	0.79	0.114	0.055	3.50E-14	2.20E-13	39.58
5	2.81	elec	3rd	1.5	0.79	0.114	0.055	3.50E-14	1.40E-13	16.03
5	2.81	elec	3rd	1.5	0.79	0.114	0.055	3.50E-14	1.20E-13	11.78
5	2.6	BVA	3rd	1.5	1.00	0.1089	0.055	5.08E-14	1.20E-13	5.58
5	2.58	BVA	3rd	1.5	1.00	0.1089	0.085	5.16E-14	7.10E-14	1.90
5	2.63	BVA	3rd	1.5	1.00	0.1089	0.085	4.96E-14	5.90E-14	1.41
5	2.6	BVA	3rd	1.5	1.00	0.1089	0.085	5.08E-14	6.90E-14	1.85
5	2.47	BVA	3rd	1.5	1.00	0.1089	0.085	5.63E-14	1.90E-13	11.40
5	2.61	BVA	3rd	1.5	1.00	0.1089	0.085	5.04E-14	1.50E-13	8.86
10	1.31	BVA	3rd				0.021	9.94E-14	1.20E-13	1.46
10	1.37	BVA	3rd				0.021	9.09E-14	1.40E-13	2.37

Table II. Typical frequency versus temperature coefficients for an AT-cut resonator with a turnover temperature of 85 ° C as a function of oven parameters. Frequency stability of 1×10^{-13} requires $\Delta T/dt < 10$ nK/s.

$$\Delta v/v \cong 3 \times 10^{-8} \Delta T^2 + 10^{-5} dT/dt\times$$

Oven Offset mK	±100 +mK	Oven Change ± 10 mK	± 1 mK	±0.1 mK
0	3×10^{-10}	3×10^{-12}	3×10^{-14}	3×10^{-15}
1	3×10^{-10}	5×10^{-12}	2×10^{-13}	6×10^{-15}
10	5×10^{-10}	2×10^{-11}	6×10^{-13}	6×10^{-14}
100	2×10^{-9}	6×10^{-11}	6×10^{-12}	6×10^{-13}

Table III. Typical frequency versus temperature coefficients for a SC-cut resonator with a turnover temperature of 85°C as a function of oven parameters. Frequency stability of 1×10^{-13} requires $\Delta T/dt < 500$ nK/s.

$$\Delta v/V \cong \times 10^{-9} + 2 \times 10^{-7} dT/dt$$

Oven Offset mK	±100 mK	Oven Change ± 10 mK	± 1 mK	± 0.1 mK
0	4×10^{-11}	4×10^{-13}	2×10^{-15}	4×10^{-17}
1	4×10^{-11}	6×10^{-13}	2×10^{-14}	8×10^{-16}
10	6×10^{-11}	2×10^{-12}	8×10^{-14}	8×10^{-15}
100	2×10^{-10}	8×10^{-12}	8×10^{-13}	8×10^{-14}

Table IV. Frequency stability of selected quartz oscillators. The oscillators reported in Besson [33] and Norton [28] use BVA resonators while those reported in Stone [22] and Wenzel [30] use electroded resonators.

	Resonator Type	Frequency	$\sigma_y(\tau)$ τ 0.1 s	$\sigma_y(\tau)$ 0.1s	$\sigma_y(\tau)$ 10 s	$\sigma_y(\tau)$ 100 s
Besson [33]	BVA 3rd SC	10 MHz			9×10^{-13}	3×10^{-13}
Norton [28]	BVA 3rd SC	10 MHz	2×10^{-12}	3.8×10^{-13}	1.3×10^{-13}	1.8×10^{-13}
Norton [28]	BVA 3rd SC	5 MHz	5.7×10^{-13}	8.4×10^{-14}	3.7×10^{-14}	6×10^{-14}
Stone [22]	Electroded 5th SC	5 MHz	4×10^{-13}	7×10^{-14}	1.2×10^{-13}	3×10^{-13}
Wenzel [38]	Electroded 3rd SC	5 MHz	2×10^{-13}	3×10^{-13}	6×10^{-13}	

Table V. VHF Crystal Flicker-of-Frequency Noise: Best Measured Results in Comparison to Walls' Volume Model Predictions Using Electroded Volume

CRYSTAL CUT	SC	SC	SC	BT	AT
FREQUENCY	80 MHz	100 MHz	160 MHz	100 MHz	40 MHz
OVERTONE	3rd	3rd	5th	5th	5th
UNLOADED Q	125K	119K	75K	300K	250K
BLANK THICKNESS (cm)	0.00678	0.00543	0.00566	0.013	0.021
ELECTRODE DIAMETER (cm)	0.162	0.140	0.254	0.317	0.366
ELECTRODED VOLUME (cm^3)	.00014	.000083	.00028	.001	.0022
BEST MEASURED Sy(f=100 Hz)*	$7.4x10^{-26}$	$3.6x10^{-26}$	$4.1x10^{-26}$	$3.0x10^{-26}$	$1.25x10^{-26}$
RESONATOR Sy(f=100 Hz)**	$6.4x10^{-26}$	$2.6x10^{-26}$	$2.5x10^{-26}$	$2.0x10^{-26}$	$1.0x10^{-26}$
MODEL PREDICTED Sy(f= 100 Hz) (using β_e=4 om eq. 1)	$2.3x10^{-26}$	$1.7x10^{-26}$	$3.3x10^{-26}$	$5.0x10^{-27}$	$2.2x10^{-26}$

* Including contribution of 4-crystal reference oscillator
** Removing contribution of 4-crystal reference oscillator

Table VI. VHF Crystal Flicker-of-Frequency Noise: Best Measured Results in Comparison to Walls' Volume Model Predictions Using Acoustic Energy Volume

CRYSTAL CUT	SC	SC	SC	BT	AT
FREQUENCY	80 MHz	100 MHz	160 MHz	100 MHz	40 MHz
OVERTONE	3rd	3rd	5TH	5TH	5TH
UNLOADED Q	125K	119K	75K	300K	250K
BLANK THICKNESS (cm)	0.00678	0.00543	0.00566	0.013	0.021
ELECTRODE DIAMETER (cm)	0.162	0.140	0.254	0.317	0.366
ELECTRODED VOLUME (cm^3)	.00018	.000056	.00054	.00034	.0012
BEST MEASURED Sy(f=100 Hz)*	$7.4x10^{-26}$	$3.6x10^{-26}$	$4.1x10^{-26}$	$3.0x10^{-26}$	$1.25x10^{-26}$
RESONATOR Sy(f=100 Hz)**	$6.4x10^{-26}$	$2.6x10^{-26}$	$2.5x10^{-26}$	$2.0x10^{-26}$	$1.0x10^{-26}$
MODEL PREDICTED Sy(f= 100 Hz) (using β_e=4 om eq. 1)	$2.9x10^{-26}$	$1.1x10^{-26}$	$6.8x10^{-26}$	$1.6x10^{-27}$	$1.2x10^{-26}$

SESSION 11

MATERIALS II

Chairman : E. PHILIPPOT

"A LOOK INTO THE CRYSTAL BALL"

SOME THOUGHTS ON THE PREDICTION OF THE AGEING OF BAW AND SAW RESONATORS

D.BOWER, N.CHURCHILL, J.DOWSETT, D.SALT

C-MAC QUARTZ CRYSTALS LTD
EDINBURGH WAY, HARLOW, ESSEX CM20 2DE, ENGLAND

ABSTRACT

This paper presents some recent ageing results for a range of BAW and SAW devices including SC, AT and ST cuts in overtone and high frequency fundamental modes at various ageing temperatures. The packages used include resistance weld, cold weld and ceramic surface mount types.

After a brief description of the measurement procedures used to gather the data, the ageing results are discussed in the context of (a) the packaging and processing methods used; and (b) commonly held assumptions regarding ageing mechanisms in crystal resonators.

The conclusion is that precise predictions of the behaviour of individual devices over long periods of time cannot be made with a high degree of confidence owing to (a) the multiplicity of mechanisms that may be present; and (b) the difficulty of determining the relative importance of those mechanisms in individual cases. Nevertheless, the accumulated data shows that with appropriate designs and process controls, devices with excellent long term ageing can be manufactured in a production environment.

1. INTRODUCTION

Quartz based frequency control products are noted for their excellent long and short term stability, which are typically an order of magnitude better than competing products, and second only to atomic clocks. Nevertheless, current applications [1,2] in long haul telecommunications, satellite navigation, and cellular communications systems require stabilities at the limit of what can be achieved in a production environment, as opposed to under laboratory conditions. In particular, in those applications such as undersea or satellite systems where periodic frequency correction is not possible, device stability over a system lifetime of many years becomes a critical parameter.

The device manufacturer is thus faced with the problem of predicting the future in order to satisfy the natural concerns of the system provider regarding the system lifetime. This paper reviews the feasibility of such predictions in the light of our current knowledge of the ageing mechanisms which may be present in typical resonators, and the ageing data for a range of frequency control devices accumulated over a long period of time.

2. DEVICES

Results are presented for the following device types:

1.000 MHz AT-cut fundamentals

5.000 MHz SC-cut third overtones

8.960 MHz AT-cut fundamentals

51.84 MHz AT-cut fundamentals

2.592 GHz SAW transversal filters

2.1 1.000 MHz: these are conventional, biconvex, fundamental mode AT-cut units packaged in a standard HC-51/U resistance weld holder. The blanks are 14.0 mm diameter, lapped, contoured and etched, with gold on nichrome electrodes. The blank is mounted in slotted ribbon mounts, bonded with polyimide paste, and sealed in dry nitrogen by capacitor discharge resistance welding. The specific application for these devices is in submerged repeater systems. The target ageing is 10 ppm over 25 years at 20 °C.

2.2 5.000 MHz: plano-convex, third overtone SC-cuts in a cold weld HC-36/U holder, designed for use in an oven controlled oscillator with turnover temperature in the range 80 to 93 °C. The design utilizes swept quartz, a 13.7 mm diameter blank, lapped, contoured and etched, but not polished, and gold on nichrome electrodes. The blank is mounted at 2 points with a low outgassing polyimide adhesive, and sealed under vacuum. All processing from blank washing on is carried out in clean room conditions.

2.3 8.960 MHz: plano-convex, fundamental mode AT-cut units, mounted at four points in a leadless ceramic surface mount package measuring approximately 8 mm square by 2.5 mm high (Fig.1). The blank diameter is 5 mm, lapped, contoured and etched, with silver on nichrome electrodes. The package is sealed in dry nitrogen by capacitor discharge resistance welding. It is anticipated that this package will increasingly replace conventional HC-49/U and HC-45/U metal holders.

2.4 51.84 MHz: fundamental mode, swept quartz, AT-cuts with an 'inverted mesa' structure produced by plasma etching. The electrode material is silver on nichrome, with the blank mounted at 4 points in a surface mount ceramic package (Fig.2) using a low outgassing polyimide adhesive. The units are sealed in dry nitrogen by seam welding a nickel lid onto a kovar brazing ring. The crystals are for application in STC Submarine Systems CANTAT transatlantic cable, where they are used as filter elements in the supervisory channels. The target ageing over the 25 year life of the system is < 15 ppm.

2.5 2.592 GHz SAWs: the results presented are from SAW filters also designed for use in the CANTAT transatlantic cable. The primary specification parameters are:

Frequency	: 2592 ± 0.25 MHz
Insertion loss	: 20.5 ± 1.5 dB
3 dB bandwidth	: 4.7 ± 0.4 MHz
Phase ripple	: <= 2° deviation from linear phase
Device ageing	: <= 22 ppm over 25 yrs @ 25 °C

Because of the very high operating frequency, the filter design uses a surface transverse wave mode rather than the normal Rayleigh mode, and the substrate orientation is chosen to optimize the turnover temperature. The SAW chip is assembled into a ceramic, surface mount package, using cantilever mounting to minimize induced strains due to thermal expansion coefficient mismatches. As with the 51.84 MHz crystal, the package was sealed in dry nitrogen by seam welding.

3. METHODS OF MEASUREMENT

3.1 Alternative techniques: the ageing characteristics of BAW and SAW quartz devices can be measured in a variety of ways. The simplest involves the storage of devices at elevated temperature, with periodic measurements performed at room temperature. The advantage of this 'storage' method is its simplicity - no custom equipment is required, other than a temperature chamber and whatever equipment is usually used for room temperature measurements. Despite this, the method is seldom used, because of the measurement uncertainties involved. These include:

(a) test fixture contact problems;
(b) effects of temperature cycling;
(c) hysteresis effects, due to devices being inactive for long periods between measurements.

An alternative approach, commonly employed for BAW devices, involves each device under test (DUT) being wired to an associated oscillator. Both DUT and oscillator are stored at elevated temperature. In an automated system, the output stage of each oscillator is enabled in turn, allowing frequency readings to be taken and the measurement data stored, all under central computer control.

To enhance the speed and resolution of frequency measurements on higher frequency devices, the oscillator output can be mixed-down using a synthesizer, mixer and low pass filter, as shown in Fig.3. This technique has been successfully employed to perform precision frequency ageing measurements of BAW quartz devices, and overcomes the primary drawbacks of the 'storage' method. However, there are still some difficulties:

(a) the only measurable parameter is frequency;
(b) custom oscillator designs are required for

different crystal types, frequencies and cuts;
(c) where the chosen ageing temperature is very high (125°C and above), the design and reliability of the oscillator can be problematic.

A third approach is to put each DUT into an ovened oscillator. This dispenses with the need for a temperature chamber, and allows devices to be aged at unique temperatures - for example, each device can be aged at its individual turnover temperature. These advantages have to be set against the increased cost of such a system.

The main drawback of the 'oscillator' ageing technique is that it is only capable of frequency measurements. To obtain ageing characteristics of other parameters, such as loss or equivalent series resistance (ESR) of a BAW device, it is necessary to perform a full one- or two-port measurement on the DUT, using a vector network analyzer or bridge method. To enable connection of multiple devices to such equipment, an RF switching matrix is required. This allows full, two-port measurements to be performed on both BAW and SAW devices, from a few kHz to 2.6GHz and beyond. However, the cost of the RF matrix makes this method extremely expensive in terms of cost per ageing position, and there are other disadvantages:

(a) limited frequency resolution using a network analyzer;
(b) long-term variations in switch characteristics can affect measurements;
(c) as with the 'storage' method, devices are inactive for the majority of the ageing period.

Any method involving the measurement of devices at elevated temperature will include uncertainties due to temperature fluctuations in the chamber. For this reason, the type of chamber usually employed will be a bath containing silicone fluid, which will have a temperature stability typically an order of magnitude better than an air oven of similar cost and capacity. Temperature fluctuations may still be a significant source of measurement uncertainty, and so it is usual to monitor the chamber temperature in some way, and use these readings to correct for such variations.

A fifth approach to ageing measurements is the 'bandrun' technique. In this approach, the units to be measured are loaded onto a carousel as used for standard overband measurements, and stored at elevated temperature. The key difference between this and the normal 'storage' method is that the frequency measurements are performed in the overband test chamber over a small range of temperatures near the devices' nominal turnover temperature. The frequency of each device at its turnover temperature can then be calculated, making temperature corrections unnecessary.

3.2 Selected techniques: the results presented here for the BAW resonators were obtained using the second approach outlined above, ie. with each crystal built in to its own test oscillator, and with oscillator and crystal immersed in a temperature bath. The measurement system is fully automated, and incorporates corrections for fluctuations in the bath temperature (Fig.3).

For the SAW filters, the devices were assembled onto ceramic thick film substrates simulating the target application, and immersed in the temperature controlled bath. The measuring system was an HP8753B network analyzer, connected to the individual devices by semi-rigid coaxial cabling and a 'tree' of RF switches. Rather than attempt to determine the centre frequency of the filter directly, the recorded frequency was the weighted mean of several hundred frequency measurements taken at intervals over the 6 dB bandwidth of the filter. The temperature of each device was monitored by a platinum resistance thermometer, and frequency measurements corrected to the nominal bath temperature by reference to the pre-determined frequency-temperature characteristics of the device.

4. AGEING RESULTS

4.1 1 MHz crystals

As part of the production process, these units were baked for 4 days at 125 °C before being placed in the ageing bath at 105 °C for 30 days during which the frequency was measured periodically. After removal from the bath for parameter measurements, a number of units were subsequently returned to the bath for a further 223 days at 105 °C. Generally speaking, the frequency shifts in the second ageing period were of the order of 1 ppm or less, but an interesting anomaly was shown by unit #0144. Fig. 4 shows the ageing data for the first 30 days, with an ageing shift at 105 °C of just over +0.5 ppm. Fig.5 shows the same crystal measured over the subsequent 223 days, and indicates quite clearly a reversal in both the sign and magnitude of the ageing.

4.2 5 MHz crystals

These units were pre-aged for 7 days at 125 °C prior to being placed in an ageing bath at 85 °C. Figs 8 and 9 show typical results over a period of 93 days, with ageing rates better than 1 $x10^{-10}$ per day.

4.3 8.960 MHz crystals

These units were baked for 7 days at 105 °C prior to be being placed in ageing at 85 °C for 62 days. Figs 6 and 7 show the ageing data for the typical case and worst case respectively, the shift in the worst case being +0.6 ppm.

4.4 51.84 MHz crystals

These units were pre-aged for 4 days at 125 °C prior to being placed in an ageing bath at 125 °C. Results over a period of 215 days are shown in Figs 10 and 11 for two of the units . While some devices show the relatively uniform behaviour typified by Fig.11, the majority are typified by Fig.10, and show a definite change in characteristics after approximately 100 days of ageing.

4.5 2.6 GHz SAWs

Fig.12 shows a typical frequency ageing characteristic over a period of 125 days at 125 °C. Fig.13 shows a histogram of the predicted 25 year ageing of a batch of 230 units based upon an initial 30 days ageing at 125 °C, while Figs 14 and 15 show similar histograms (for batches of sizes 114 and 36 respectively) based on 125 days of testing at 125 and 105 °C. These histograms show quite clearly that the predictions based on the longer time period differ significantly from those based on the shorter period, and in fact are significantly improved.

5. DISCUSSION

5.1 Ageing mechanisms

As might be expected from the central importance of long term frequency stability in applications, there is an extensive literature regarding the causes of ageing in quartz devices. Comprehensive reviews can be found in Gerber [3], and Vig and Meeker [4], while Vig and Walls review the fundamental limits on oscillator instabilities in [5]. Naturally enough, the bulk of the literature is concerned with 'state of the art' high precision devices, fabricated under very closely controlled conditions in relatively small volumes, rather than with devices made for wider use in professional and commercial applications. In these latter cases, time and cost considerations usually make it impossible to justify the use of the fabrication methods viewed as necessary for the very highest stabilities, and a key concern of this paper is to consider whether and how the work done on very high precision devices can be applied in a wider sphere.

It is generally agreed in the literature that the causes of ageing can be grouped under four headings:

(1) stress effects, for example stress relaxation in the mounting structure and the electrode films;

(2) mass loading effects, for example the adsorption of contaminants from the atmosphere;

(3) material effects, such as chemical reactions or diffusion at the electrode-quartz interface, or structural relaxations in the quartz itself;

(4) other effects, such as changes in hydrostatic pressure due to leaks in the enclosure, or permeation of gases through the holder walls.

There is considerably less agreement in the literature on the precise form of the frequency-time relationship resulting from these various mechanisms, and even less agreement on how the various parameters involved depend upon the temperature. For both mass transfer and stress relaxation, it has been argued [4,6] that the appropriate frequency-time relationship has the form

$$f = A \ln(1 + Bt) \qquad (1)$$

where f is the fractional frequency shift from time t = 0 and A and B are temperature dependent parameters. Eqn (1) leads to an ageing rate diminishing with time according to

$$df/dt = AB/(1 + Bt) \qquad (2)$$

The time dependence encapsulated in eqns (1) and (2) has been incorporated in Mil Std MIL-O-55310.

5.2 Physical basis of the 'log time' form

The form of eqn (1) is derived [4,6] from the assumption of a rate equation

$$dq/dt = K \exp(-E/kT) \qquad (3)$$

where q is a physical parameter that can either be identified as mass loading on the resonator surface, or as stress in the mounting system or electrode film which is balanced by an opposite stress in the resonator, and where E is an activation energy depending linearly on q:

$$E = E_0 + qE_1 \qquad (4)$$

In the stress case, the frequency shift is supposed to result from the force-frequency effect which is in turn due to the non-linear elastic properties of the resonator material.

With (4), eqn (3) integrates to

$$q = q_0 + (kT/E_1)\ln\{1 + [K\exp(-E_0/kT)/C]t\} \qquad (5)$$

where C is a constant of integration determined by the initial value $q_0 = (kT/E_1)\ln(E_1C/kT)$. Thus C is given by

$$C = (kT/E_1)\exp(E_1q_0/kT) \qquad (6)$$

and (5) becomes

$$\Delta q = A \ln(1 + Bt)$$

with

$$A = kT/E_1$$
$$B = (KE_1/kT)\exp[-(E_0 + E_1q_0)/kT] \qquad (7)$$

Thus in (1) A should be proportional to the absolute temperature T, while the product AB should depend exponentially on T. However, it has been argued elsewhere (in relation to SAW devices [7]) that it is parameter A that is thermally activated and that B is essentially independent of temperature.

5.3 Difficulties with the 'log time' model

(a) Mass transfer

Whether considering stress relaxation or mass transfer, there are difficulties with eqn (1) on intuitive physical grounds. Consider first mass transfer. Given a hermetically sealed enclosure containing a finite amount of contaminant, then at any given temperature it would be expected that the system consisting of the resonator, its enclosure, and the contaminant would tend towards a state of dynamic equilibrium. In this state the rate of adsorption of the contaminant on to the resonator surfaces would equal its rate of desorption, and the amount of contaminant on the resonator surface would have some equilibrium value. It would thus be expected that the resonator frequency would also tend towards an equilibrium value. According to (1) however, the frequency change is unbounded, even though the rate of change tends to zero.

In the mass transfer case, the difficulty is the consequence of the absence of any dependence on concentration in the rate equation (3). Suppose for the sake of example that q represents the number of adsorbed atoms per unit area on the resonator surface, and that E is the binding energy. Then the rate of desorption is proportional to q and inversely proportional to the average stay time τ of the adsorbed atoms on the surface, given by

$$\tau = \tau_0 \exp(E/kT) \qquad (8)$$

where τ_0 is the period of vibration of the adsorbed atoms. Thus (3) is replaced by

$$dq/dt = -q/\tau = -(q/\tau_0)\exp(-E/kT) \qquad (9)$$

which if any dependence of E on q is neglected integrates to

$$q = q_0\exp(-t/\tau)) \qquad (10)$$

In terms of frequency, this leads to

$$f = f_u[1 - \exp(-t/\tau)] \qquad (11)$$

where f_u is the asymptotic frequency shift.

This simple analysis assumes that all desorbed atoms are somehow removed from the system and are not available to be readsorbed. This might happen if for example the atoms desorbed from the resonator surface were immediately adsorbed on to the enclosure walls, with a binding energy much greater than E, but in any event it is straightforward to extend the analysis to allow for readsorption.

Consider an enclosure of volume V containing a total number N of adsorbate molecules, and with a surface area A available for adsorption. Suppose the number of adsorption sites per unit area is N_m, and that $N <= AN_m$ so that at most one molecular layer of adsorbed molecules can be formed. Additionally, suppose that the probability of an adsorbate molecule that strikes an available site actually being adsorbed rather than being reflected is α_s (the condensation coefficient [8]). At any particular time t, let q be the number of adsorbed molecules per unit area, so that the remaining number of adsorption sites is (N_m-q), and the number of adsorbate molecules remaining in the vapour phase is (N-qA). Then the desorption rate is

$$(dq/dt)_{des} = -q/\tau$$

as before, while the adsorption rate is

$$(dq/dt)_{ads} = (c_{av}\alpha_s/4V)(N-qA)(N_m-q)$$

where c_{av} is the mean speed of the adsorbate molecules in the vapour phase. The total rate of change of q is thus the sum of these two expressions, and depends quadratically on q. It is easy to see that if q = 0, dq/dt is > 0, while if $q = N_s$, dq/dt is < 0. Thus in equilibrium when dq/dt = 0, q must have some intermediate value.

The general form of the rate equation is

$$dq/dt = aq^2 + bq + c$$
$$= a\{[q + (b/2a)]^2 - [(b/2a)^2 - c/a]\}$$
$$= a[(q + q_1)^2 - q_2^2]$$

This can be integrated to give

$$q = -q_1 - q_2\coth[aq_2(t + t_0)]$$

Since coth(x) tends to 1 as x tend to infinity, the asymptotic value of q is $q_{eq} = -(q_1+q_2)$, when dq/dt = 0. More precisely, since coth(x) -> 1 + 2exp(-2x) for large x, then for large t, q behaves as

$$q \to q_{eq} - 2q_2\exp[-2aq_2(t + t_0)]$$

so that the final approach to equilibrium is exponential as in the simple case represented by (10).

(b) Stress relief

Similarly, in the case of stress relief, it would be intuitively supposed that the frequency would tend asymptotically to the value corresponding to the zero stress state, rather than increase or decrease indefinitely. It would appear that the logarithmic relationship has been adopted from the study of creep in metals. In this type of work it is typical to study creep under constant load (cf. the review paper by McLean [9] and the paper by Michel et al [10]). In such conditions it is reasonable to expect a continuing extension of the sample, but these conditions do not apply in a resonator. Here the initial stresses in the mounting structure or electrode films are balanced by opposite stresses in the resonator. Thus the load on the film or the mounting is provided by the quartz, and as film or mounting adjusts by some sort of plastic deformation, the stress and strain in the quartz will reduce in proportion.

5.4 Curvefitting and prediction

Given a set of reasonably well behaved data, it will always be possible to fit the data as closely as desired, provided that the approximating function has enough adjustable parameters. Curvefitting in this sense (interpolation) is not a problem, but using the approximating function to extrapolate outside the range of the experimental data can only be justified on physical grounds. Given the hypothesis that a particular type of mechanism is responsible for the observed data, the functional form of the dependence must be derived by physical arguments, usually

involving a number of unknown physical quantities. Curvefitting can be used to derive values for these unknown physical parameters. Future behaviour can then be predicted on the basis of the original hypothesis. However, the fact that the particular functional form fitted the original data is not a guarantee that the hypothesis was correct, and certainly not a guarantee of the predictions made. On the other hand, if the predictions made turn out to be incorrect, then it will be proven that the original hypothesis was in itself not correct.

For prediction on the basis of a given hypothesis to have any validity, there are two requirements:

(1) gather enough data to perform a curve fit on the basis of the selected hypothesis, and determine the appropriate parameters;

(2) test the hypothesis by collecting further data and comparing that data with the predictions from the model determined in (1).

If the predictions are confirmed, then at least some support will be given to the original hypothesis and further predictions from the model. If not the hypothesis is disproved and an alternative must be sought. In the case of the data presented here, it is more often than not the case that although a reasonable fit can be obtained to the model expressed by eqn (1) on the basis of early ageing data, this fit is not confirmed by subsequent data. This is particularly true in the case of the 51.84 MHz fundamental units which almost all show a change in their ageing characteristics after about 100 days, not predicted by the model fitted to the early data. This general observation is consistent with the similar observation made by Filler and Vig [11] in their study of the long term ageing of various crystal oscillators over period of several years.

6. CONCLUSIONS

On both physical grounds and on the basis of the ageing data presented here, the ageing model expressed by eqn (1) must be regarded as unreliable. It must be concluded that precise predictions of the behaviour of individual devices over long periods of time cannot be made with a high degree of confidence due to (a) the multiplicity of mechanisms that may be present; and (b) the difficulty of determining the relative importance of those mechanisms in individual cases.

The likelihood of multiple mechanisms also indicates the danger of placing too much reliance on the 'acceleration factors' sometimes used to deduce ageing at one temperature from data taken at another temperature. Also, if in the case where mass transfer is the dominant mechanism, it is supposed that the observed ageing represents an approach to a dynamic equilibrium, the characteristics of the ageing, and in particular its sign, will depend just as much on the previous thermal history of the device as on the actual ageing temperature.

Nevertheless, despite these cautions, the accumulated data shows that with appropriate designs and process controls, devices with excellent long term ageing can be manufactured in a production environment.

7. REFERENCES

1. H.Hellwig, 'Time and Frequency Applications', IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol 40, No 5, pp 538-543, 1993

2. J.R.Vig, 'Military Applications of High Accuracy Frequency Standards and Clocks', IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol 40, No 5, pp 522-527, 1993.

3. E.A.Gerber, 'Long Term Stability and Aging of Resonators', in Precision Frequency Control, E.A.Gerber and A.Ballato (eds), Vol 1, Chapt 6, p. 208 et seq, 1985.

4. J.R.Vig and T.R.Meeker, 'The Aging of Bulk Acoustic Wave Resonators, Filters and Oscillators', Proceedings of the 45th Annual Symposium on Frequency Control, pp 77-101, 1991.

5. J.R.Vig and F.L.Walls, 'Fundamental Limits on the Frequency Instabilities of Quartz Crystal Oscillators', Proceedings of the 1994 IEEE International Frequency Control Symposium, pp 506-523, 1994.

6. R.L.Filler, 'Aging Specification, Measurement, and Analysis', Proceedings of the 7th Quartz Devices Conference, pp 93-104, 1985.

7. P.L.Scarff, M.B.Carey, J.E.Huff, 'An Aging Model for Surface Acoustic Wave Devices', IEEE Transactions on Ultrasonics, Ferroelectricity, and Frequency Control, Vol 40, No 6, pp 630-641, 1993.

8. M.Prutton, Surface Physics, Chapt 6, p.114, 2nd Ed, Oxford University Press, 1983.

9. D.McLean, 'The physics of high temperature creep in metals', in Reports on Progress in Physics, Vol XXIX, Part 1, pp 1-34, 1966.

10. JP.Michel, M.Zafrany, V.Candelier, J.Lamboley, L.Xavier, F.A.Kuhnast, M.Nassik, J.M.Fiorani, 'Influence of Part Stability on Aging of Quartz Crystals', Proceedings of the 8th European Frequency and Time Forum, pp 832-842, 1994.

11. R.L.Filler and J.R.Vig, 'Long Term Aging of Oscillators', IEEE Transactions on Ultrasonics, Ferroelectricity, and Frequency Control, Vol 40, No 4, pp 387-394, 1993.

FIG. 1

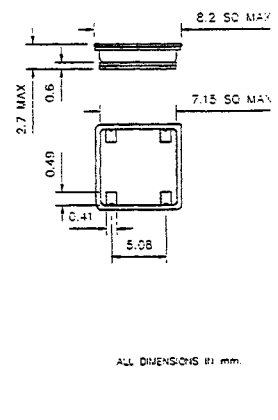
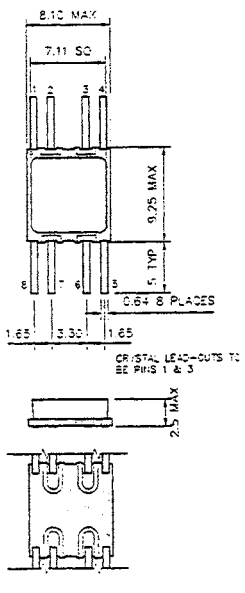


FIG. 2



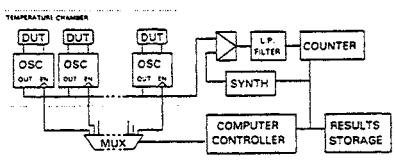


FIGURE 3 - OSCILLATOR AGEING SYSTEM CONFIGURATION

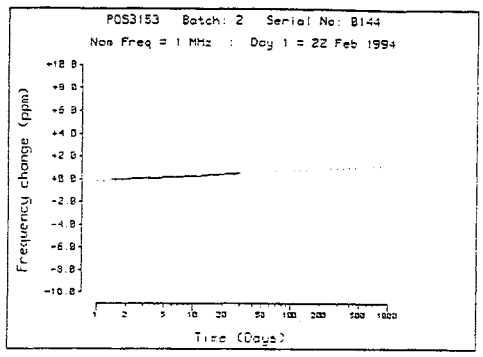


FIGURE 4

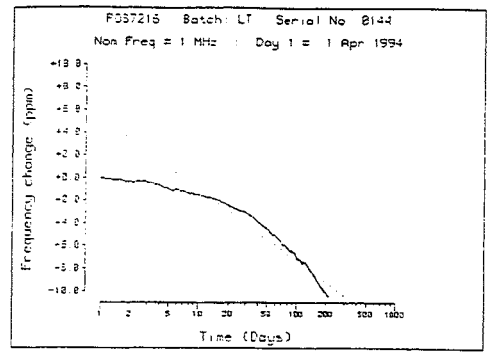


FIGURE 5

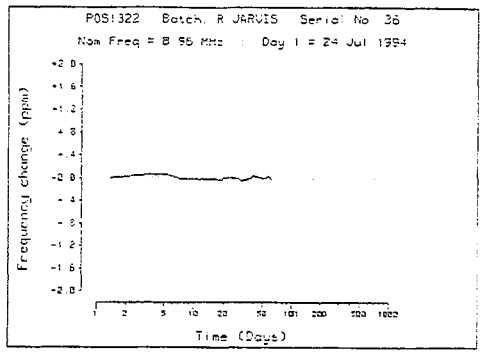


FIGURE 6

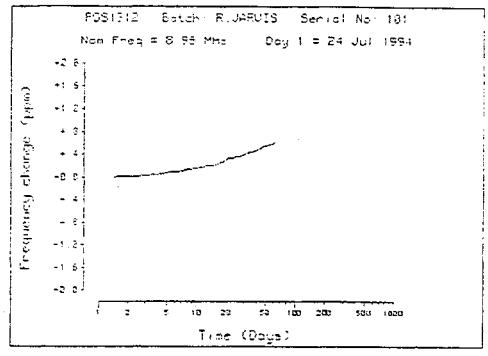


FIGURE 7

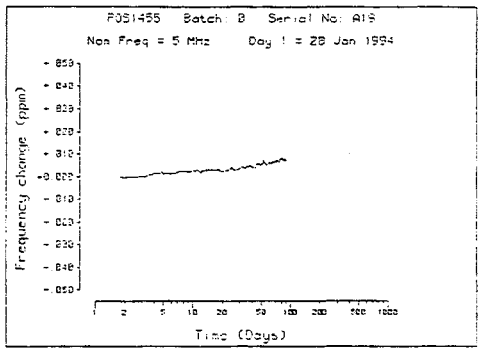


FIGURE 8

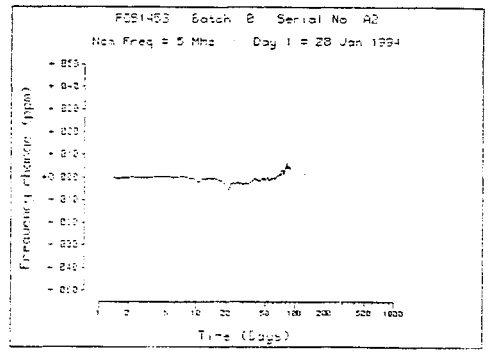


FIGURE 9

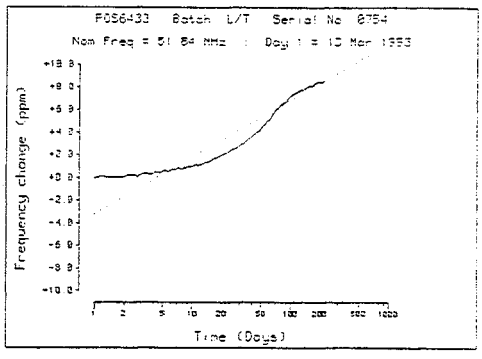


FIGURE 10

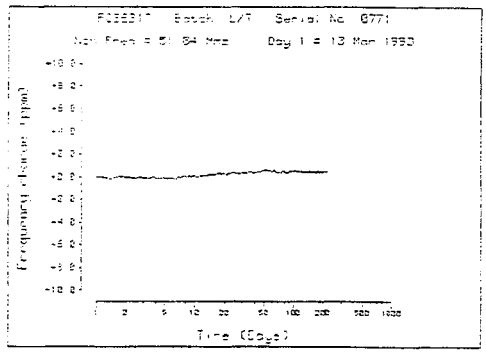


FIGURE 11

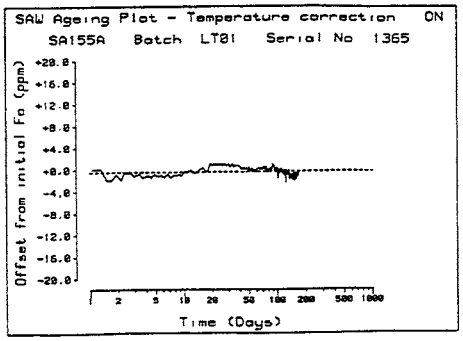
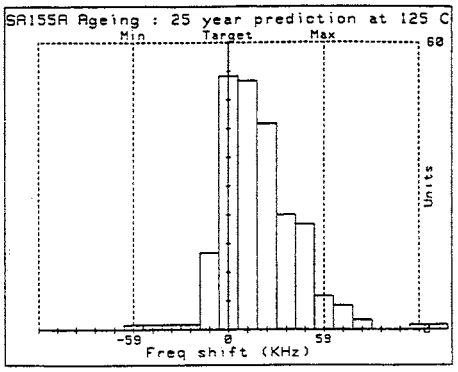


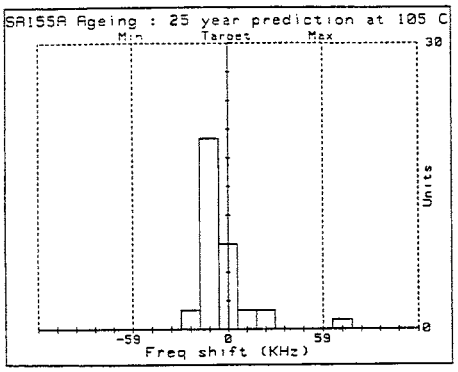
FIGURE 12



Production Unit Subset (720 hours at 125 deg C)

N_units = 230 Mean = 19.0 Sigma = 23.7

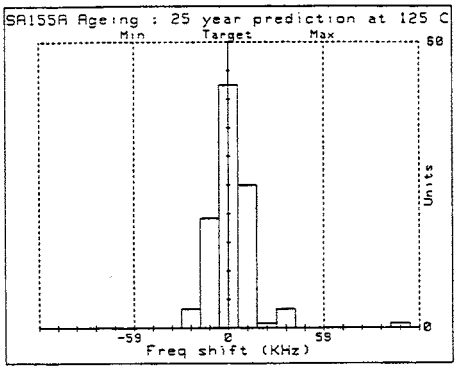
FIGURE 13



Life-Test Units (3000 hours at 105 degrees C)

N_units = 36 Mean = -3.3 Sigma = 15.2

FIGURE 14



Life-Test Units (3000 hours at 125 degrees C)

N_units = 114 Mean = 2.8 Sigma = 14.5

FIGURE 15

INSITU X-RAY DIFFRACTION STUDY OF SWEEPING OF IMPURITY IONS IN QUARTZ CRYSTALS.

M.T. SEBASTIAN[a,b], R.A. BECKER[a] and H. KLAPPER[a]

(a) Institut fur Kristallographie, RWTH, 5100 Aachen FRG
(b) Regional Research Laboratory, Trivandrum 695 019, India

Abstract

A high temperature and high voltage vacuum chamber has been fabricated for insitu X-ray diffraction study of sweeping of impurity ions in quartz crystals. X-ray diffraction studies are made on synthetic quartz crystals both in α and β phases using graphite and silver electrodes. X-ray topographs showed the contrast due to space charge collection near the anode side of the crystal which disappeared during irradiation. Sweeping experiments performed using silver electrodes showed contrast due to the migration of silver ions near the cathode side. It is found that the contrast due to space charge collection appears only in the α phase and disappears as it transforms to the β phase. The mechanism for the appearance of the contrast and its disappearance during X-ray irradiation is discussed.

Introduction

Among the oxides, quartz is a unique material. It is extremely stable and found in nature as large crystals or grown in the laboratory with even higher quality. Its most prominent property is piezo-electricity in combination with excellent mechanical behaviour. This property allows quartz to be used in a variety of electronic devices like crystal oscillators and filters for precision frequency control, sensors of temperature, pressure and mass. In most of these applications imperfections play an important role and quartz crystals of very high quality are often required. Aluminium is invariably present in both synthetic and natural quartz which exists as a substitute for silicon. From an ionic point of view, Aluminium is a +3 entity in a +4 site. The trivalent aluminium can make only three of the bonds normally made by a silicon and hence a negative charge on the site due to the non-bonding oxygen. This charge is compensated by an alkali ion (Li^+, Na^+) or a proton (H^+) electrostatically bound in an adjacent c-axis channel. Temperature variations, electric field or irradiation causes the weakly bound alkali ions to move through the large structural channels. This leads to frequency changes in resonators made from such quartz crystals. The easy migration of impurity ions through the large c-axis tunnels has been exploited by several authors [1-3] to purify quartz crystals. Commercially the process is called sweeping. In the present paper we report the fabrication of a high temperature and high voltage vacuum chamber for the insitu x-ray diffraction study of sweeping in quartz crystals.

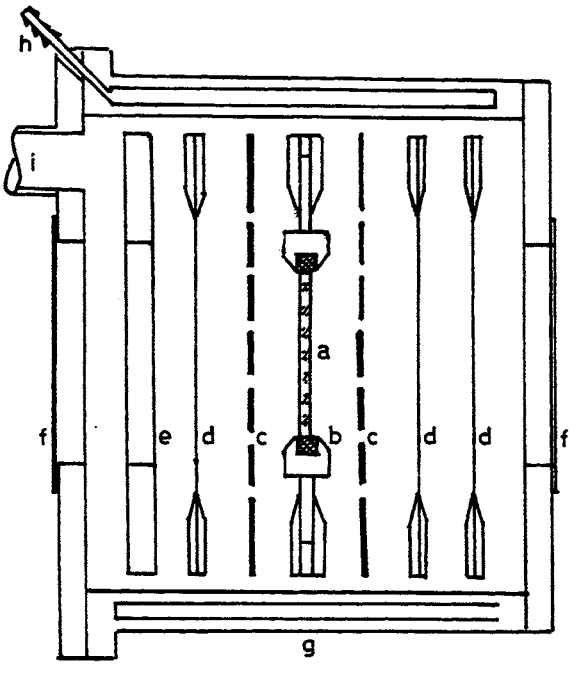


Fig.1. Schematic sketch of the chamber a) crystal b) soft graphite felt electrodes c) graphite heater plates d) Al-heat shield foils (e) ceramic carrier (f) beryllium windows (g) body of the chamber with water cooling (h) water-in and outlet (i) outlet to the vacuum pump.

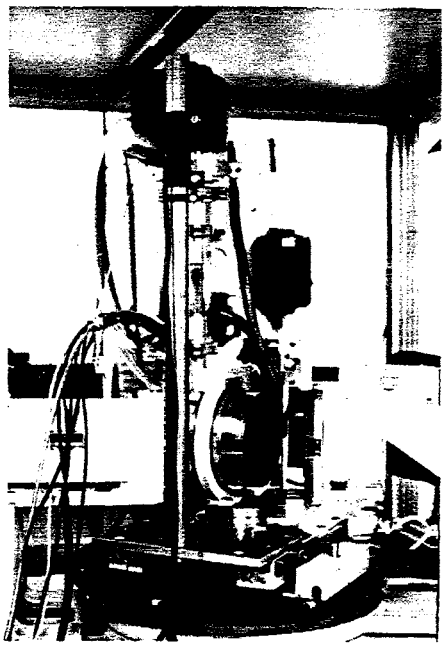


Fig.2. Photograph of the high temperature and high voltage chamber mounted on a commercial Lang camera.

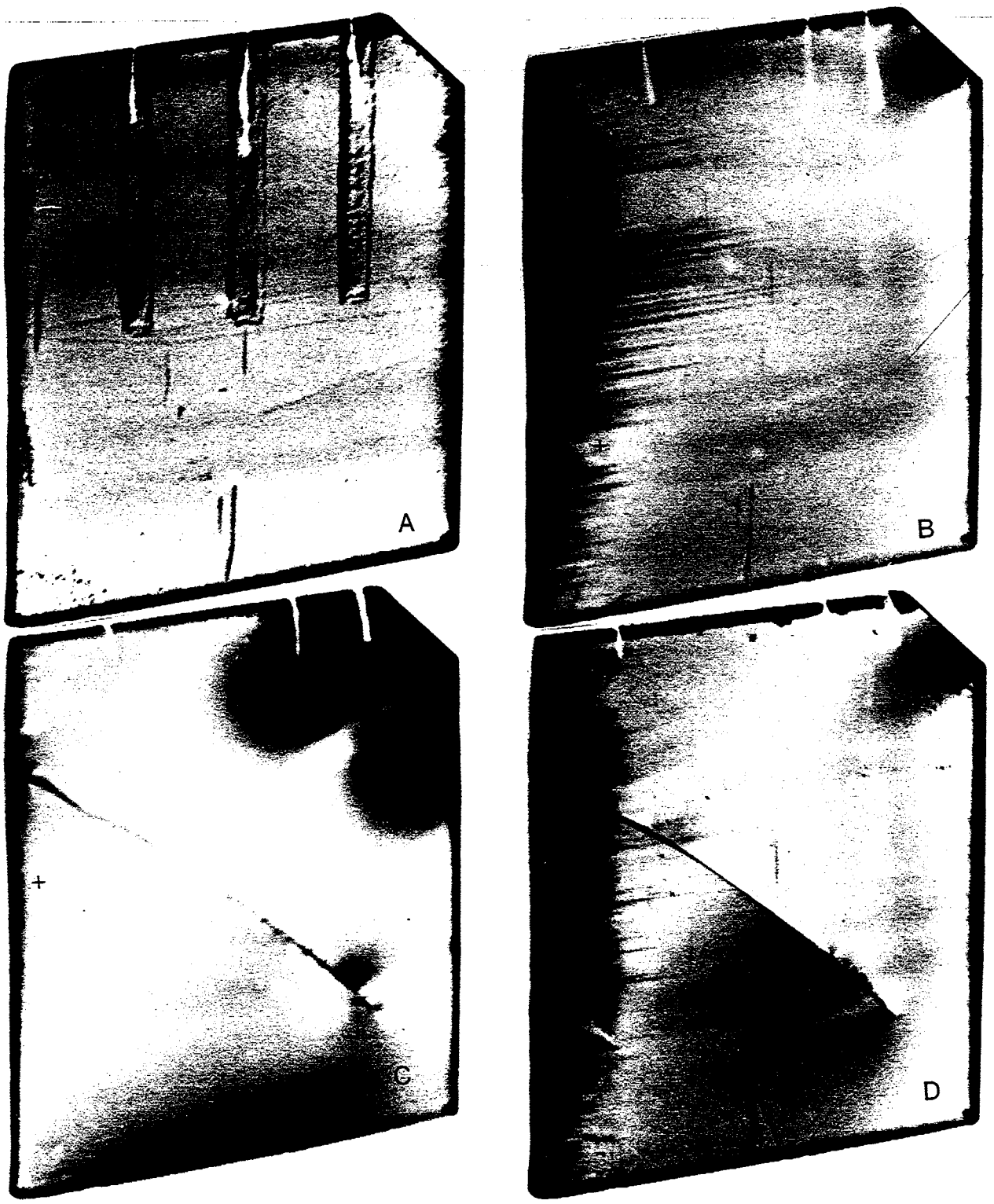


Fig.3. 10.1 X-ray topographs recorded from a Y-cut synthetic quartz crystal using MoKα radiation. (a) RT without electric field. The three vertical contrasts are due to electrodes for measuring the negative space charge developed on the crystal. (b) 545°C with 2.7 kV/cm after 4 days (c) 573°C with E = 2.7 kV/cm after 6 days (d) 580 °C with E = 2.7 kV/cm after 7 days (e) 573°C with E = 2.7 kV/cm after 8 days.

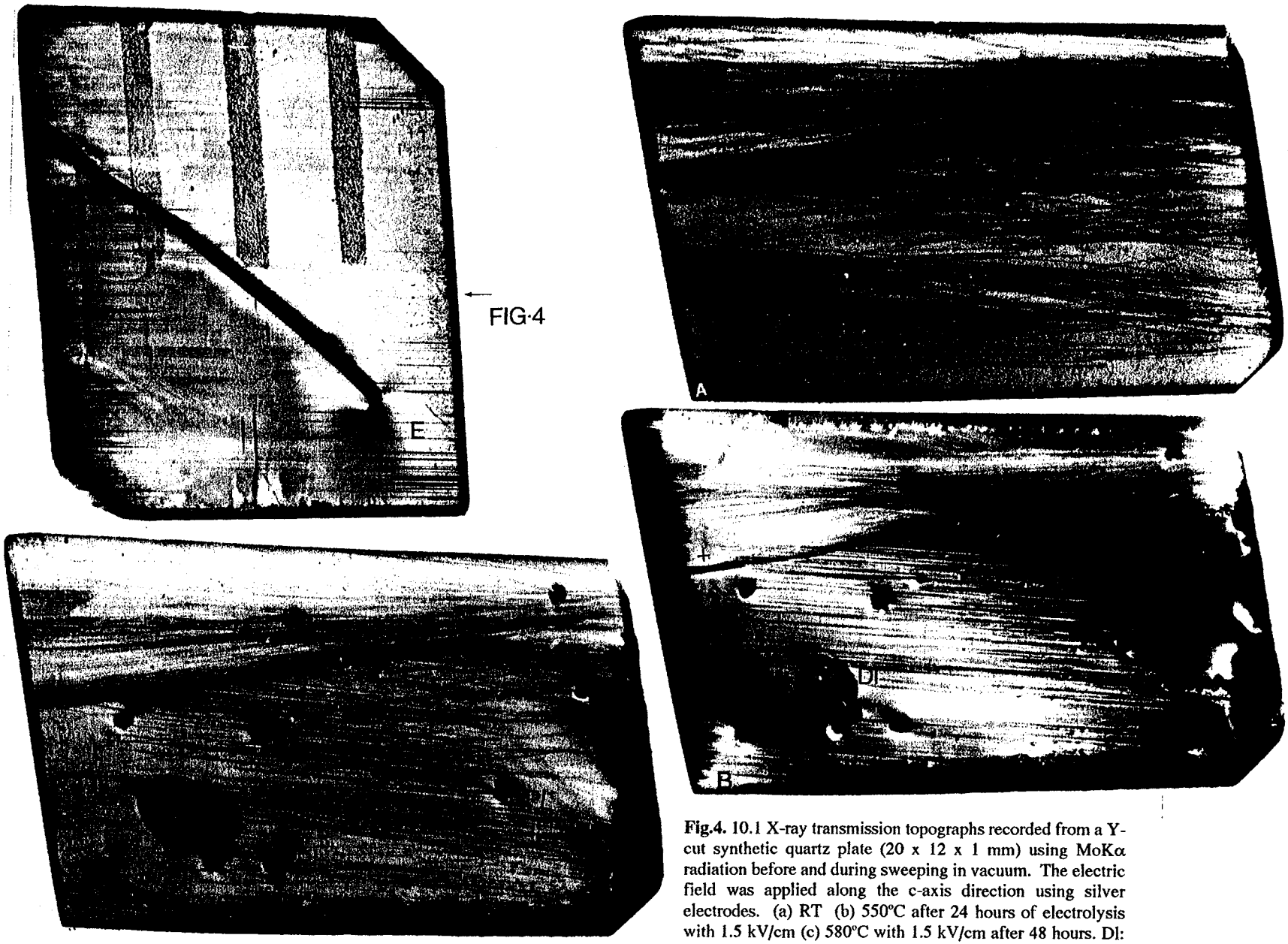


Fig.4. 10.1 X-ray transmission topographs recorded from a Y-cut synthetic quartz plate (20 x 12 x 1 mm) using MoKα radiation before and during sweeping in vacuum. The electric field was applied along the c-axis direction using silver electrodes. (a) RT (b) 550°C after 24 hours of electrolysis with 1.5 kV/cm (c) 580°C with 1.5 kV/cm after 48 hours. Dl: dislocation loops.

Experimental

Figure 1 gives a schematic sketch of the chamber. The specimen crystal is pressed between graphite felt electrodes. The use of soft graphite felt electrodes gives very good electrical contacts even for curved crystals and also prevents the problems arising from thermal expansion of the crystals at high temperatures. Figure 2 shows the photograph of the chamber. The chamber can be rotated on a commercial Lang camera for minor adjustments to bring the crystal into the diffracting conditions. The heater plates are made of graphite and the furnace is available up to 700°C. The outer framework of the chamber is made of stainless steel and is cooled by running cold water. The temperature is detected by two chrome-alumel thermocouples. The beryllium windows are wide open with sufficient size for recording traverse topographs of crystals up to the size 20 x 20 mm. The chamber can be rotated to make it possible to record topographs with the reflections of a series of lattice planes belonging to a zone axis. When the specimen is heated the crystal orientation changes slightly and a re-adjustment is needed. All adjustments required for recording the section and traverse topographs can be performed at high temperatures as in an ordinary Lang camera. A high voltage of up to 3 kV/cm of the crystal can be applied to the crystal in vacuum (10^{-4} torr) obtained using a turbomolecular pump. In the present investigation Y-cut synthetic quartz crystal plates of about 1 mm thickness were used. The crystals are polished, etched and then mounted between the graphite electrodes in the fabricated chamber. First an X-ray topograph was recorded at room temperature without electric field. This is followed by recording several topographs of the same or other reflections with insitu electric field at different temperatures in vacuum. The exposure time for recording topographs with insitu electric field was about half of that without field. This is because of the enhancement in the diffracted intensity under the action of the static electric field [4-6].

Results and discussion

Figure 3 shows the 10.1 x-ray topographs recorded from a quartz plate at different stages of sweeping. Fig. 3(a) was recorded with E=2.7 kV/cm along the c-axis at 545°C. Strong striation contrast sprouting from the anode side (left side) of the crystal appears. The contrast slowly disappears as it transforms to the β phase (see Fig c and d) and reappears as the crystal cools down to the α phase (see fig e). Fig. 4 shows the 10.1 transmission topographs recorded from a Y-cut synthetic quartz plate recorded before and during sweeping using silver electrodes. Silver paste was applied on both sides of the crystal before mounting between the soft graphite felt. The crystal contains a large number of dislocations. The topographs recorded with insitu electric field 1.5 kV/cm along the c-axis shows a showing dark contrast near the cathode side (right side of the Fig.3) This contrast is due to the lattice deformation caused by the accumulation of silver ions. As the electrolysis progressed dark striation contrast also appeared at the left side (anode side). The contrast disappeared as the crystal transforms to the β phase (Fig.4d) and re-appeared on cooling below the transition temperature.

Several dark spotty contrast appeared on the topographs (Fig 4) when some of the crystals were heated above 500°C. Most of the spotty contrast are along dislocation lines. In some crystals dislocation loops (marked Dl in the figure) also appear which are found to be associated to precipitates. The spotty contrast is due to precipitation of inclusions and are up to 1 mm in size. The contrast around fluid inclusions and silver is related to elastic deformation of the quartz lattice due to an increase of pressure in the fluid inclusion. TEM studies [7-9] have shown that annealing of wet quartz at atmospheric pressure induces the precipitation of a large density of water bubbles and the nucleation of a large density of dislocations or dislocation loops.

The strong contrast with striations parallel to the c-axis and sprouting from the anode (left) side of the crystal is present on all hkil topographs. The contrast appeared only on prolonged heating at temperatures between 400 and 572°C with insitu electric field along the c-axis direction. The contrast completely disappeared when the topographs were recorded above 573°C (in the β phase) and reappeared on cooling below the transition temperature. The contrast is present only when the electric field is along the c-axis. The striation contrast on the topographs sprouting from the anode end of the crystal is attributed to lattice deformation caused by space charge collection [4-6]. During solid state electrolysis the relatively smaller and weakly bound alkali ions move towards the cathode through the large structural channels. If additional mobile positive ions are not available at the anode side, electrolysis will result in the collection of negative space charge at the 4th non-bonding oxygen near the aluminium. This space charge collection deforms the atomic planes inhomogeneously giving rise to striation contrast on the topographs. X-ray irradiation destroys the striation contrast by the production of Al-hole (Al-e^+, centres. The striation contrast disappears as the crystal transforms to the β phase. It is not known how the lattice deformation due to negative space charge collection is relieved in the β phase. Further work is needed in this direction.

References

1. A. Kats, Philips Res. Repts. **17** (1962) 133.
2. R.N. Brown, J.J. O'Connor and A.F. Armington, Mater. Res. Bull. **15** (1980) 1067.
3. D.R. Koehler and J.J. Martin, J. Appl. Phys. **57** (1985) 5209.
4. M.T. Sebastian and H. Klapper, 4th Eur. Time Freq. Forum (1990) 659.
5. M.T. Sebastian, A. Zarka and B. Capelle, J. Appl. Cryst. **21** (1988) 326.
6. M.T. Sebastian and H. Klapper, Phys. Stat. Sol. **b186** (1995).
7. R.D. Baeta and K.H.G. Ashby, Phys. Stat. Sol. **18** (1973) 155.
8. A.C. McLaren, R.F. Cook, S.T. Hyde and R.C. Tobin, J. Phys. Chem. Miner. **9** (1983) 79.
9. S.H. Kirby and J.W. McCormick, Bull. Minerol. **102** (1979) 124.

DEFECTS CHARACTERIZATION IN SYNTHETIC AND NATURAL QUARTZ : THE ROLE OF IONIZING RADIATIONS

Part II of "Non destructive study of Al-hole concentration and frequency change of quartz resonators".

J.J. BOY*, M. A. PASQUALI*, B. VIARD**, K. YAMNI**
*ENSMM/LCEP - Route de Gray - 25030 BESANÇON Cedex - France
**UFR Sciences/LCCM - Route de Gray - 25030 BESANÇON Cedex - France

ABSTRACT

Treatments as sweeping or irradiation have been employed in order to change the point defects configuration in high quality quartz crystal. The paramagnetic defects Al-hole (whose the most prominent is $[AlO_4]^\circ$), created during these processes, can be qualified by Electron Paramagnetic Resonance spectrometry (EPR, so called ESR). The production of these centers is also monitored by 632.8 nm laser absorption technique as function of progressive irradiation doses. Finally, the F.T. Infra-Red spectrometry is used to evaluate the competition between the formation of Al-hole and Al-OH defects. The Inductive Coupled Plasma spectrometry gives us, as for its, the concentration of the usual impurities (Al, Fe or Alkali ions).

As results, we have exhibited strong differences in the characteristics presented by natural and synthetic quartz material, in one hand, and by +X and Z-zones of the same synthetic crystal, in other hand. Furthermore, we have tested resonators manufactured in different blanks qualified by the techniques presented above. Theirs electrical characteristics have been correlated with the kind of defects.

Keywords : Infra-Red, ICP and ESR spectrometries, Laser absorption measurements, quartz crystal.

1. INTRODUCTION

Irradiation and sweeping have employed as treatments on high quality quartz crystal used to manufacture ultra-stable resonators. The produced defects as Al-OH and Al-hole can be monitored by IR-FT, 632.8 nm laser absorption or ESR techniques, in comparing results before and after each treatment. In the present paper, this characterization procedure has been applied on Y and AT-cuts taken from pure Z and +X zones of synthetic quartz crystals. In order to evaluate the influence of the treatments in the performances of resonators, specific measurements of Q-factor, motional resistance or frequency shift versus temperature have been performed. The tested resonators are AT-cut BVA_4 (i.e. with deposited electrodes) working at 10 MHz 3rd overtone.

Point defects in *as-received* (or as-grown) synthetic quartz are usually present as Al—M^+ (M = Li, Na, K,...) and as-grown OH^- centers. The stability of these centers have been investigated by several authors in using irradiation and sweeping process. Halliburton et al. [4] have investigated the absolute concentration of Al—defects in synthetic quartz as function of the irradiation temperature. They concluded that the ratio of Al—OH and Al—hole formation depends on the types of electron traps present in the sample and on the convertion of the original Al—M^+ centers occurred by irradiation at temperatures above 200 K. Sibley et al. [5] also have reported the effects of radiation on the Al-OH and OH^- bands production and their temperature dependences. Some papers of Lipson et al. [1-3] concern the distribution of Al—OH defects in synthetic quartz produced by additional radiation doses and sweeping process. These authors have employed low temperature IR-FT spectrometry and also He-Ne laser absorption for mapping the Al-OH and Al-hole distribution in high-quality synthetic quartz. Investigations of the dose dependence irradiation and the rates of as-grown OH^- and Al-OH defects production have conduced these authors to suggest that the Al—OH formation is provided by an additional hydrogen-ion source different of the original OH^- centers.

After Kats [6], many authors have established a parallel study between synthetic and natural quartz. Brown and Kahan [7] have applied sweeping and high radiation doses to monitor the changes of IR spectra and to distinguish swept synthetic, unswept and natural quartz. Bahadur [8] has verified the modifications of some specific IR bands of natural and synthetic quartz exposed to different radiation doses at 77 and 300 K. In addition to the radiation exposure, he applied very low field sweeping (of about 30 V/cm) for alkali removal and also annealing treatments.

2. SAMPLE PREPARATION

Synthetic quartz crystals were obtained from SAWYER Research Products. The samples were taken from pure-Z and +X zones (Fig. 1). These blanks were prepared, resulting in specific AT and Y-cuts of rectangular shapes. Their dimensions depend on the kind of measurement, the smallest ones being used for ESR spectrometry. But, for the AT-cut samples and as far as possible, we have analyzed the zone used to realize the resonators, whichfor the diameter is smaller than about 15 mm.

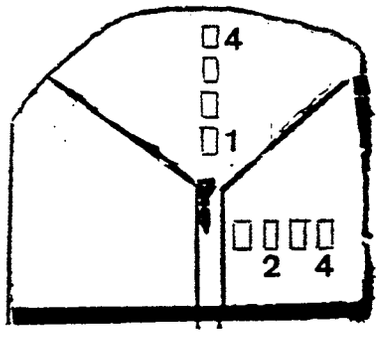


Fig. 1 : Scheme of Y-cut from synthetic quartz

The results issued from infrared and 632.8 nm laser absorption measurements were carried out from 5 mm thick AT and Y-cut samples before and after each treatment.

Y-cut plates were exposed at 100 Krad radiations (from Cesium source). After analysis, a half part of each plate has been used to prepare a set of small samples to ESR investigations. Their rectangular shapes (4x3x2 mm) were cut in order to keep the main dimension parallel to the c axis of the quartz crystal. At least, an additional dose of ionizing radiations (from X-rays source) was applied to these samples.

A Z-cut natural quartz plates (with a thickness of 5 mm) were prepared from each original block. Each sample was exposed to a 1 Mrad γ-irradiation from a Co^{60} source and, as previously described, infrared and laser absorption measurements were employed to monitor the point defects changes.

3. EXPERIMENTAL PROCEDURE

a. Infrared measurements :

The data are obtained between 3100 and 3700 cm^{-1} with a NICOLET IR-FT spectrometer, model 750. The samples are put inside a cryogenic cell cooled to the nitrogen liquid temperature. The window used for IR analyses is 0.5 cm diameter.

b. The laser absorption technique :

As described in previous works [9, 10], this technique is performed with an optical set-up composed by a 2mW He-Ne laser emitting a monocromatic 632.8 nm radiation, two micrometric motion units controlled by an HP computer and a silicon photodiode detector. The data acquisition is obtained as shown in Figure 2. The sample is analysed with the beam perpendicular to the Z-axis of the AT and Y samples. The absorption values for one plane are calculated from the following relationship :

$$\sigma_L = \frac{1}{d} \cdot \frac{1}{N} \cdot \sum_{n=0}^{N} \log_{10} \frac{I_0}{I_n}$$

where :

d is the thikness of the sample (generally 0.5 cm),

I_0 and I_n are the measured intensities outside and through the sample at the point number n,

and $\overline{\sigma_L}$ is the mean value of the all σ_L.

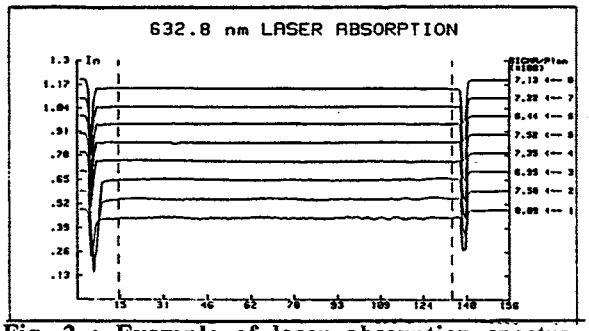


Fig. 2 : Example of laser absorption spectra (I_n versus Point number)

c. ESR experiments :

The spectra are realized by a spectrometer with a VARIAN rectangular cavity operating at 9.25 GHz. The magnetic field range was 3200 ± 100 G and the modulation frequency was 200 KHz to almost of our experiments. Except for one spectrum, all ESR data were taken at 20 K.

d. The impurity concentrations :

The Al and alkali metals contents (Na, K and Li) were determined by ICP techniques and presented in the previous paper [9].

At least, we indicate that the electrodiffusion (so called sweeping) process has been performed at nitrogen atmosphere on 5 mm thick AT-cut samples inserted between platinum electrodes. A 1500 V/cm has been applied and the temperature kept at 505 °C during 4 or 7 days.

4. RESULTS AND DISCUSSION

4.1. IR spectra of synthetic quartz samples :

Figures 3 and 4 show comparative infrared spectra of samples extracted from +X and Z zones. They correspond to different treatments :

AS-RCVD means "as-received",

100 Krad for 100 Krad γ-irradiation,

100 Krad+X-rays for an additional dose of X-rays.

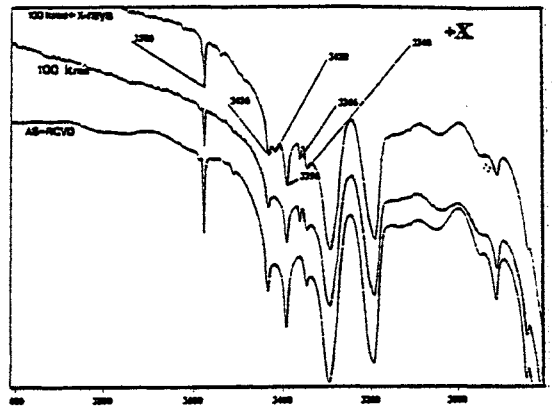


Fig. 3 : IR spectra for +X zone

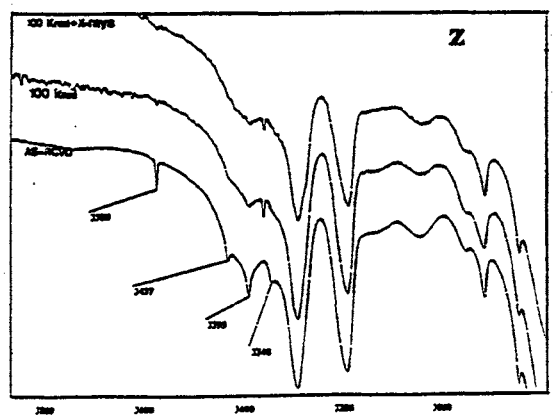


Fig. 4 : IR spectra for Z zone

The table 1 presented below indicates the relative absorption coefficient of each band, relatively to the absorption at 3200 cm^{-1} (which is linked to Si—O vibration, independant of any treatement and not so close to 3306 cm^{-1} as the other Si—O vibration).

If the three s bands (3580, 3436, 3348 cm^{-1}) tend to vanish rapidely in the Z zone, s_2 remains constant after 100 Krad dose as in the +X zone. For this zone, the s bands are not really affected by the additional radiation dose. Of course, e_2 band (at 3366 cm^{-1}) appears while the

radiation dose increases, their respective heights being almost identical.

So, in the Z zone, we observe a quasi-disappearance of the as-grown OH^- defects from the first irradiation, All the bands show a saturation effect at this dose. For the +X zone, we observe first that the intensity of each s band is stronger than this of Z zone ; so, the as-grown OH^- defects are more numerous in this zone. Furthermore, we note, here too, that s and e bands are not really affected by the additional radiation dose and the 4 s bands remain present.

We are in agreement with other authors who have verified that high doses can induce a saturation effect, whereas the s bands disappeared practically under high irradiation conditions [3], [7].

Table 1 : Relative absorption coefficients in synthetic quartz

Bands (cm^{-1})	Relative absorption coefficient (* 10^{-1})									
	As-recieved		γ-irradiation (100 Krad)		γ-irradiation and X-rays		AT-cut unswept		AT-cut swept	
	+X	Z	+X	Z	+X	Z	+X	Z	+X	Z
3580 (s_4)	5.00	1.59	2.79	# 0	2.78	# 0	7.12	1.59	4.12	1.72
3436 (s_3)	2.58	0.62	1.48	# 0	1.18	# 0	3.33	0.58	3.14	0.63
3396 (s_2)	4.39	2.50	3.44	1.33	3.50	1.40	4.55	2.32	1.96	1.88
3348 (s_1)	1.06	0.49	1.15	0	1.18	# 0	0.91	# 0	# 0	# 0
3366 (e_2)	0	0	1.15	1.17	1.20	1.09	0	0	26.67	0.94
3306 (e_1)	0	0	0	0	0	0	0	0	13.92	0
σ (cm^{-1})	0.076	0.098	0.087	0.10	-	-	0.076	0.086	0.086	0.089

- : not measured

Figures 5 and 6 show the IR spectra of small samples taken from different positions in relation to the seed of +X and Z zones (as indicated in figure 1). They have been also irradiated as described above. The number of Al—OH defects created by irradiation increases with the distance to the seed of the crystal. Far from the seed and particularly in +X zone, the s bands disappear, almost of the hydroxyl radicals participating to the formation of Al—OH defects (revealed by creation of a very deep e_2 peak and a little e_1 (at 3306 cm^{-1}) absorption peak). For e_2, the ratio between the relative absorption coefficients for both positions are 4.46 for +X zone and 1.52 for the other one.

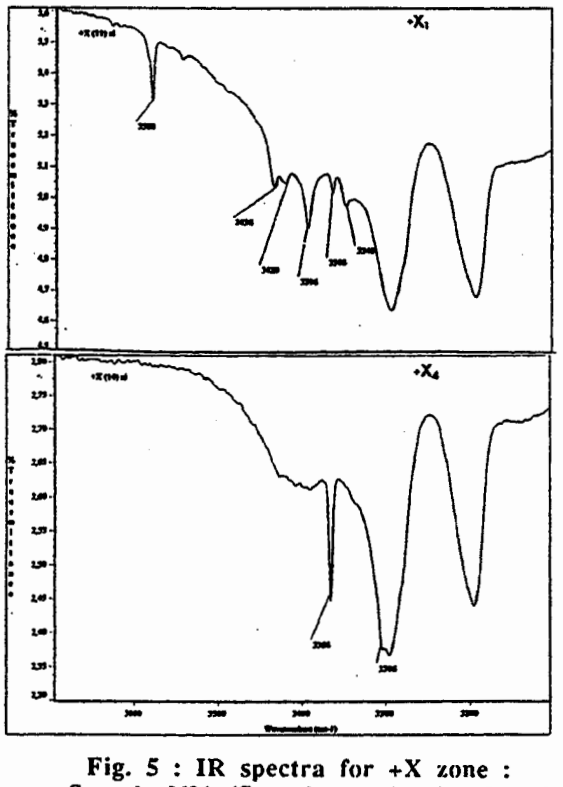


Fig. 5 : IR spectra for +X zone : a. Sample N°1 (See also table 1) ; b. N°4

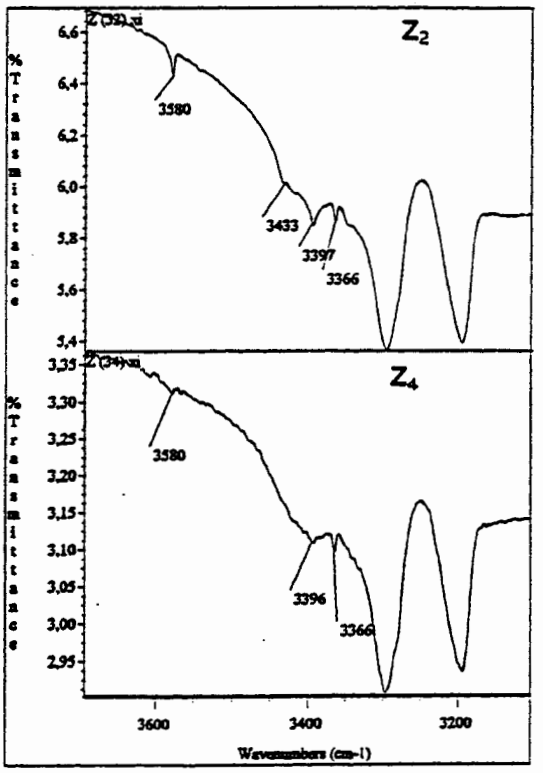


Fig. 6 : IR spectra for Z zone : a. Sample N°2 ; b. N°4 (See also table 1)

The IR spectra shown in the next figures (7 and 8) concern AT-cut samples extracted also from +X and Z zones. Each sample, swept as indicated above, exhibits some absorption peaks whose the relative absorption coefficients are presented in the right part of table 1. After sweeping, we note stronger absorptions due to Al—OH and a formation of e_1 peak at 3306 cm^{-1} just in +X zone. In order to monitor the effect of sweeping, ratios between coefficients at 3366 and 3580 cm^{-1} were calculated : they are 6.47 and 0.55 in +X and Z zones respectively.

Changes in the original conditions suggest that the Al—OH concentration for swept samples can saturate bfore irradiation.

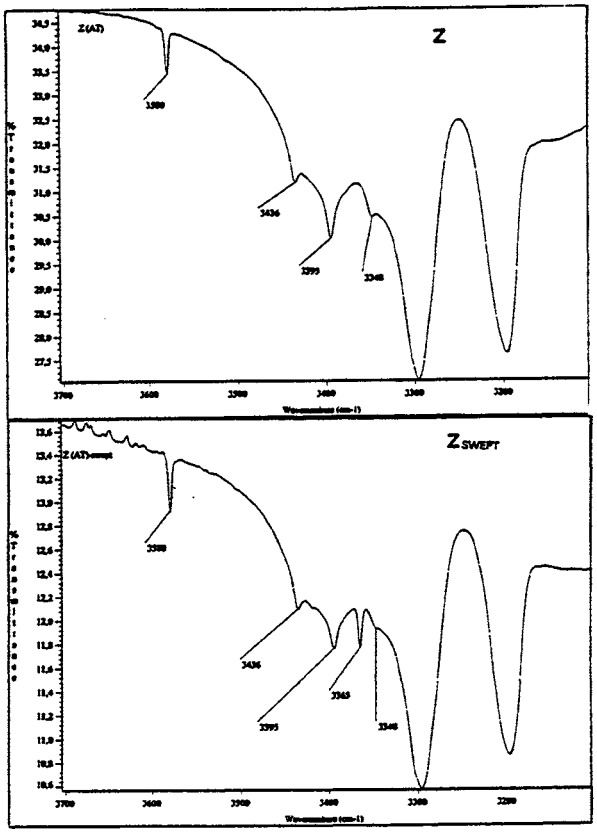


Fig. 7 : IR spectra for AT-cut from Z zone : a. "as-received" ; b. After sweeping

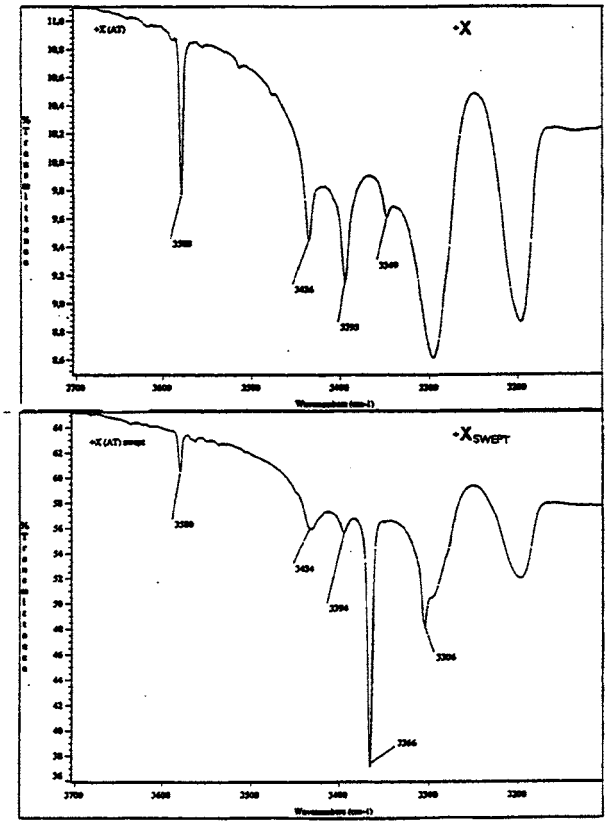


Fig. 8 : IR spectra for AT-cut from +X zone : a. "as-received" ; b. After sweeping

4.2. ESR spectra of synthetic quartz samples :

As the next figures 9 and 10, they show the +X and Z zones ESR spectra whichfor the magnetic field is parallel to Z-axis. The Hyperfine Structures (HFS) exhibited, at room temperature, by +X sample (Fig. 9a) indicate an intense paramagnetic center. In Z zone, these HFS lines are about 10 times more intense close to the seed than far from this. Furthermore, the amount of defects created just by X-rays exposure is similar in +X and in the nearest part to the seed of Z zone.

Indeed, though the results of laser absorption at 632.8 nm indicate a probably, but small formation of paramagnetic centers with a 100 Krad radiation doses, we have not seen HFS lines in any sample just after the first stage of irradiation.

At least, we can see that a charateristic HFS of Al-hole defect is clearly recognized at the right of the left part of each low temperature spectrum.

Other ESR spectra, whichfor their studies are in progress, have been performed in rotating Z-axis toward magnetic field direction and could be presented in other paper.

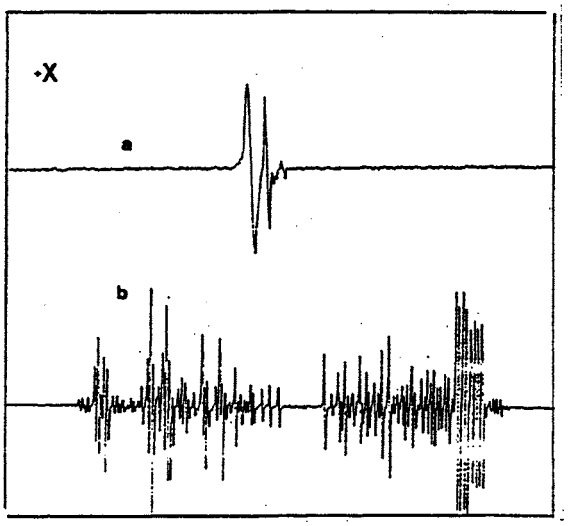


Fig. 9 : ESR spectra of +X zone (Gain : 4000) a. At room temperature ; b. With T = 20 K

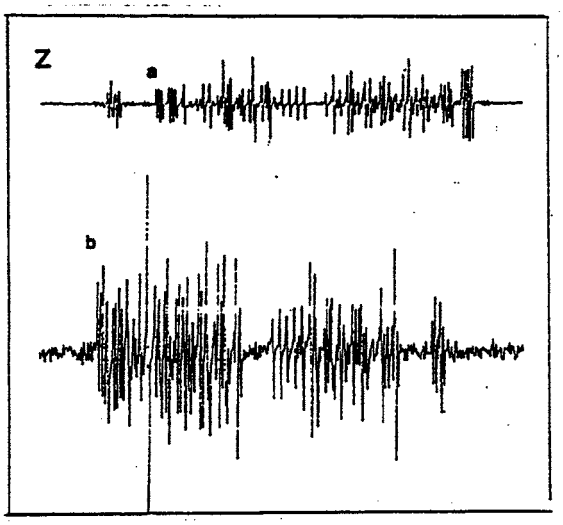


Fig. 10 : ESR spectra of Z zone
a. Seed side (Gain : 4,000)
b. Far from the seed (Gain : 40,000)

4.3. ICP analyses of synthetic quartz samples :

The Al and alkalis concentrations shown in the next table confirm that the impurities concentration is higher in +X zone than in Z zone (for more details, see also [9]).

Table 2 : Impurities level (in ppm weight)

Element	+X zone	Z zone
Al	2.92	1.57
Li	0.61	0.08
Na	3.36	0.001

4.4. IR spectra of natural quartz samples :

In order to compare the radiation response of synthetic and natural quartz, IR spectra and 632.8 nm absorption measurements have been carried out also on 4 natural quartz samples from Brazil. The main results are presented in the table 3.

Also, the figure 11 shows the IR spectra for a high-quality natural quartz.

Table 3 : Relative absorption coefficients in natural quartz

	Relative absorption coefficient (* 10^{-1})							
Bands (cm^{-1})	NQ7		NQ12		NQ16		NQ21	
	B.I.	A.I.	B.I.	A.I.	B.I.	A.I.	B.I.	A.I.
3595	2.17	2.00	8.28	5.17	2.39	2.16	2.37	2.09
3474 (Li)	8.17	2.83	44.8	15.4	12.2	2.94	2.54	0.75
3430	7.17	7.00	9.83	12.6	5.44	6.08	2.71	2.84
3366 (e_2)	92.3	95.7	162	140	79.6	79.8	43.2	43.9
3306 (e_1)	24.5	25.2	40.9	36.6	20.9	21.4	15.3	14.9
σ (cm^{-1})	0.081	0.24	0.11	0.25	0.051	0.26	0.081	0.34

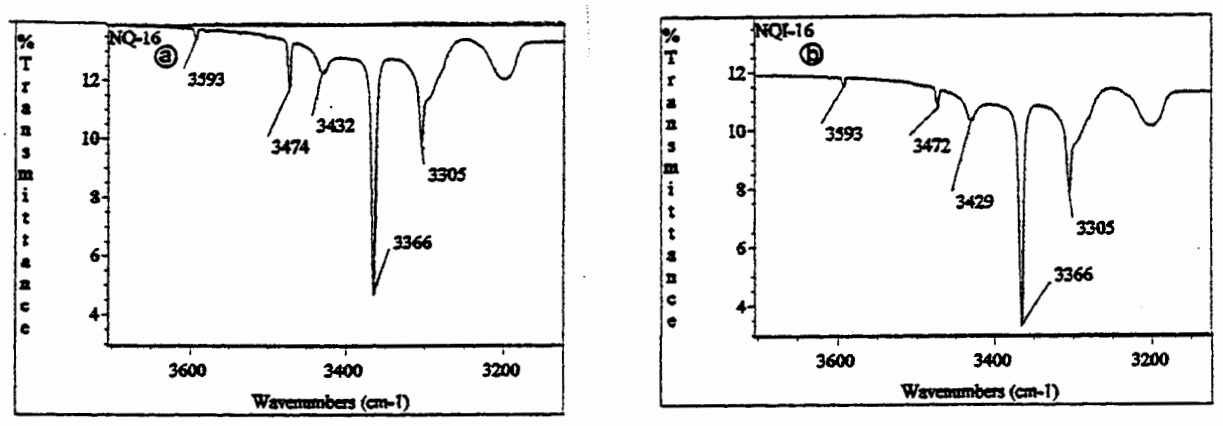


Fig. 11 : IR spectra for Z-cut from natural quartz : a. "as-received" ; b. After γ-irradiation (1 Mrad)

On these spectra, we observe that the Al–OH bands increase slightly and the other absorption peaks remain constant or exhibit a small reduction. While these bands are produced by irradiation or electrodiffusion process in synthetic quartz, they are present in as-received natural quartz (which has been, perhaps, naturally irradiated). The most remarkable effect of irradiation of natural quartz is the strong reduction of 3474 cm^{-1} band which is associated to Li-OH center.

At least, the results shown in table 3 suggest that the effect of ionizing radiation on natural quartz creates an important amount of Al-hole defects, as indicated by the intense absorption observed at 632.8 nm. We have to note that all samples presented a high degree of transparency before irradiation and just only one (NQ21) showed a pronounced smoky coloration after γ-irradiation.

5. AT-CUT RESONATORS

Here, we present just the first measurements about the electrical characteristics of AT-cut 10 MHz resonators with deposited electrodes and sealed in a "classical" mount.

To characterize the effect of treatments as irradiation and sweeping on the resonance frequencies, we have realized 8 resonators issued from both +X and Z zones of synthetic quartz. In each zone, 2 of them are manufactured in as-grown crystal and 2 are issued from swept blank. At least, we have irradiated (100 Krad, Ce source) one resonator from each different kind of quartz blank.

The next table (Table 4) presents the motional resistance of each kind of resonator measured in a π-bridge at atmospheric pressure and at about 40 °C.

Table 4 : Motional resistances (Rm in Ohms) of AT-cut resonators

	As-recieved	As-RCVD + Irrad.	Swept	Swept + Irrad.
Z zone	115	130	135	155
+X	122	126	122	145

We can remark that the motional resistances of unswept resonators are systematically better than these of swept resonators. So, if certain authors have noted that sweeping can improve the long term drift behaviour of high quality resonators, it is important to said that the conditions of sweeping keep an essential part on the performances. Furthermore, irradiation damage slightly this electrical parameter.

At least, the next figures (12.a and 12.b) illustrate the temperature behaviour of the resonance frequencies of certain resonators.

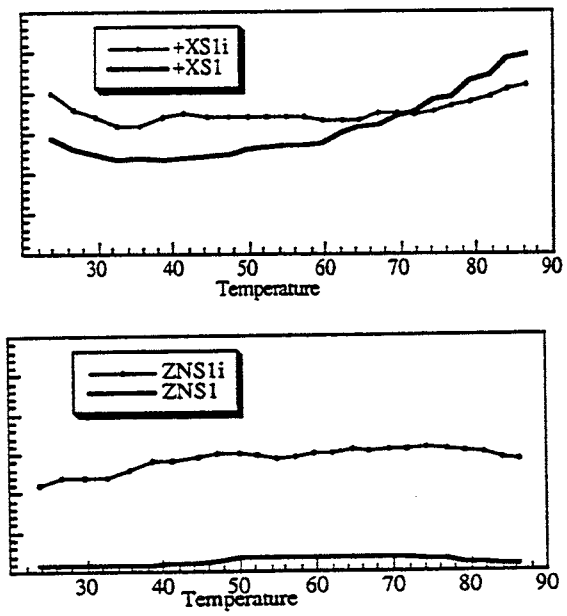


Fig. 12 : Rm vs temperature before and after irradiation a. Swept in +X zone ; b. As-RCVD in Z zone

6. CONCLUSION

This work is a part of previous publication for which the main purpose is first the characterization of the defects produced by treatments as irradiation and sweeping of high quality synthetic quartz and second, the study of their influences on performances of resonators manufactured in the qualified material.

We can summarize the principal observations as follow :

- on the as-received state, both zones (+X and Z) of synthetic quartz present the as-grown OH^- absorption peaks,
- after 100 Krad γ-irradiation, these s bands dissapear in Z zone and show a strong reduction in +X zone. Furthermore, an additional irradiation dose in +X zone does not affect the intensity of s bands,
- in spite of the disapearence or the reduction of s lines, the e_2 band is produced in the 2 zones and also, they show a stabilization after 100 Krad γ-irradiation,
- there is an inhomogeneous distribution of impurities in each zone as observed by ICP spectrometry. Indeed, the produced defects by irradiation and sweeping increase toward the seed. This is confirmed by IR and ESR analyses,
- sweeping of as-received samples produced similar effects than irradiation : reduction of the s bands and formation of e bands, but the absorption intensity for both lines is more pronounced after sweeping. Comparatively, the 3366 cm^{-1} band of +X swept sample shows a signal 22 times greater than the same line of the irradiated sample,
- natural quartz samples present a strong absorption at 632.8 nm indicating Al-hole formation from the original Al—centers,
- we can be supposed that the Li-OH center associated to the 3474 cm^{-1} absorption band is not as stable as the other ones under irradiation. In order to speculate the compensation of OH and holes centers formation, it would be interesting to correlate the concentration of each original and produced defects,
- althougt the impurities concentrations of natural quartz have not been determined, the IR and 632.8 nm absorption results indicate a considerable concentration of Al in comparison with synthetic crystal, sufficiently to create Al-hole defects.

ACKNOWLEDGEMENTS

The authors thank to Prof. F. Jandron and Prof. M. Clergeaux of "Acoustique et Optique de la Matière Condensée" laboratory (P. et M. Curie University) for performing the ESR analyses. Also, Prof P. Zecchini for helpful suggestions about the IR data presentation (Franche-Comté University). They express thanks to C.R.T.S. of Besançon and EMBRARAD (Brazil) for irradiation. M.A.Pasquali would express her gratitude to CAPES (Brazil) for supporting by means a scholarship.

REFERENCES

1. A. KAHAN, H.G. LIPSON : "Defect Centers in irradiated and swept quartz" - Proc. 39th Ann. Freq. Control Symp. (1985), pp. 255, 258.
2. H.G. LIPSON, A. KAHAN : "Infrared characterization of Aluminum and Hydrogen defect centers in irradiated quartz" - J. Appl. Phys., 58 (2), (1985).
3. H.G. LIPSON, F. EULER : "Dose dependence of radiation - Induced defect changes in quartz" - Proc. 42d Ann. Freq. Control Symp. (1988), pp. 169, 175.
4. L.E. HALLIBURTON, N. KOUMVAKOLIS, M.E. MARKES, J.J. MARTIN : "Radiation effects in crystalline SiO_2 : the role of Aluminum" - J. Appl. Phys., 52 (5), (1981).
5. W.A. SIBLEY, J.J. MARTIN, M.C. WINTERSGILL, J.D. BROWN : "The effect of radiation on the OH^- infrared absorption of quartz crystals". J. Appl. Phys., vol. 50 (8) (1979) , pp 5449, 5452.
6. A. KATS : "Hydrogen in alpha-quartz" - Philips Res. Rep. 17, 133 (1962).
7. R.N. BROWN, A. KAHAN : "Optical Absorption of Irradiated Quartz in the Near I.R." - J. Phys. Chem. Solids , vol. 36 (1975) , pp 467-476.
8. H. BAHADUR : "Infrared characterisation of natural and cultured quartz : the effect of electrodiffusion and irradiation" - J. App. Phys. 66 (10) (1989), pp. 4973, 4982.
9. K. YAMNI, J.J. BOY, M.A. PASQUALI, P. ZECCHINI : "Non-Destructive Study of Al-hole Concentration and Frequency change of Quartz Resonators (Part I)". Proc. 8th Europ. Freq. and Time Forum (1994) , pp 203, 214.
10. P. ZECCHINI , H.G.LIPSON : "Defect Distribution Mapping in Quartz". 43rd Ann. Symp. Freq. Control (1989) , pp 490-496.
11. Y. SHIMODA, T. UNO : "Impurity and defect center concentration in high-purity quartz : their change processes by sweeping and irradiation" - J. Appl. Phys., vol. 29 (6) (1990), pp. 1159, 1164.

STUDY OF IMPURITIES IN QUARTZ BY ELECTRIC MEASUREMENTS.

C. POIGNON, G. JEANDEL and G. MORLOT.

L.E.M.T.A. (U.R.A. 875)
Faculté des Sciences B.P. 239.
54506 VANDOEUVRE-LES-NANCY cedex. FRANCE.

ABSTRACT.

We have measured thermionic currents by the T.S.D. (Thermally Stimulated Depolarization) method in quartz samples in order to study the mobility of ionic impurities in this material. High quality and standard quality samples have been studied. In order to purify the crystals, the 'sweeping' technique is used. We compare current density of a sample swept in air with that of a sample swept under vacuum; the two crystals are cut perpendicular to Z (c) axis. The influence of polarization time and of electric tension on the current density has been studied for a high quality untreated quartz sample.

Keywords: quartz , impurities , sweeping , Ionic Thermally Current.

1. INTRODUCTION.

Quartz which is a piezoelectric material is widely used in much technological equipment ranging from clocks to telecommunications or location satellites. Those sophisticated systems require a very high temporal stability of frequency, whereas degradation is often observed; this reduction of quartz resonator performances is due to the defects present in the crystal. Some impurities can move under the effect of an ionized irradiation or under the effect of an electric field. This mobility brings about variations in frequency of the resonator.

These impurities are essentially alkali ions M^+ (Li^+, Na^+, K^+) and protons H^+ which occur as compensating charges in an interstitial position. A positive ion offsets the defect of the charge which is created by the substitution of a silicium ion Si^{4+} for an aluminum ion Al^{3+} during crystal growth.

The ability which some interstitial impurities have to move along channels parallel to **c** has been studied in the first by King [1] and is used in industry to purify the crystals. The application of an electric field to the terminals of a heated sample brings about the moving of ions that are less bonded to the crystalline structure towards the surface. The extremities of the crystal are subsequently cut.

In order to try to better characterize the sweeping, the T.S.D. (Thermally Stimulated Depolarization) method [2,3] has been employed. We used it for both high quality and standard quality samples of quartz furnished by the S.I.C.N. (Société Industrielle de Combustible Nucléaire à Annecy).

2. THE T.S.D. METHOD.

2.1. Sweeping.

Sweeping is the industrial process which can purify the quartz crystals by diminishing the number of ionic impurities. The crystal structure (class 32) presents large channels parallel to **c**. Along these channels, small interstitial impurities can move. An electric field is applied (parallel to **c**) to the terminals of a sample which is heated to about 500°C. Thermal agitation allows the interstitial ion to jump over the potential barrier created by the presence of the associated substitutional impurity and by the crystal [4]. The released particle can move along the channel towards the surface following the direction of the electric field. After several days, the sample is cooled to ambient temperature; then the electric tension is removed. The extremities of the crystal, where the interstitial impurities are accumulated, are cut. Electrodes in platinum are generally applied to the two faces perpendicular to Z (**c**) of the sample. Sweeping can be made under vacuum or in air.

2.2. Some defects detection techniques.

Defects which are present in quartz are detectable with different techniques. Dislocations are observable by X ray topography [5]. The centres created by the association of a substitutional ion with an interstitial ion have been studied by acoustic loss measurements [6,7] or by dielectric relaxation measurements [8].

Infrared spectroscopy gives information on the concentration of [OH] in the crystal [5,9]. Chemical analysis quantifies the elements which are present in a sample reduced to the plasma state [5]. When a quartz crystal is irradiated, a hole can be created in the p-orbital of an oxygen ion adjacent to an Al^{3+} ion. The Al-hole center is detected, for example, by E.S.R. techniques. An electron can be trapped in an oxygen vacancy too. This defect called E'_1 is also detected by these techniques [10].

The T.S.D. method allows studying the mobility of ionic impurities which are moved under the effect of an electric field at a given temperature. Quartz is an insulating material, the conductivity is ionic. So, thermionic currents are measured with this T.S.D. technique [11,12,13].

2.3. Principle.

When an electric field is applied to a sample, either a displacement of charge inside the material from one equilibrium position to another or a displacement of charge around the equilibrium position (rotation of dipoles) is created.

The T.S.D. method consists in applying an electric voltage to the terminals of a sample at a fixed temperature [2,3]. The dipoles are not oriented initially; they tend to align parallel to the electric field. Under the effect of both thermal agitation and field, the interstitial charges which are the least bonded to the crystal structure migrate by successive jumps to the electrode (cathode if the charges are positive). A depolarizing field is created by these charges; it reduces the number of oriented dipoles. The sample is then cooled, the tension is removed and the electrodes are short-circuited through an electrometer. The temperature is raised linearly. The released particles migrate inside the material. The measured depolarizing current is due to the relaxation by the rotation of dipoles and the movement of space charges.

3. RESULTS AND DISCUSSION.

3.1. Characterization of samples and experimental conditions.

Standard quality (n°511) and high quality (n°272) quartz samples have been studied. Some crystals were purified under vacuum, others in air. Plates are metallized by evaporation in a vacuum: chromium electrodes are realized like this. Plate thickness is 0,75mm. They are cut perpendicular to the Z (c) axis.

An electric tension of 1600 V is applied to the terminals of a sample for 14 hours at 350 K. Then, the crystal is cooled to the temperature of liquid nitrogen, the electric tension is removed and the electrodes are short-circuited. Current measurement begins a few minutes after removal of the tension at this temperature (Figure 1).

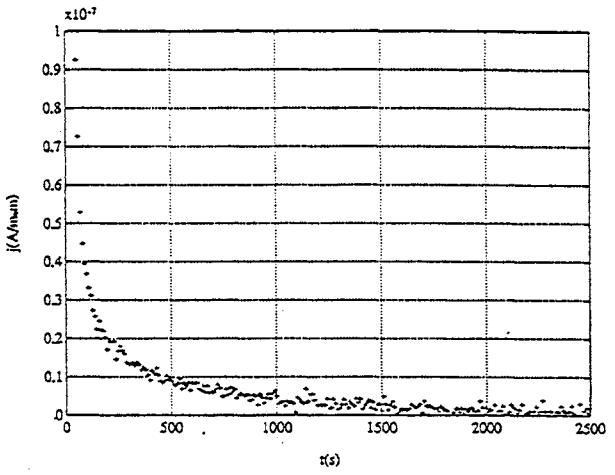


Figure 1: measurement of current a few minutes after removal of electric tension at 78 K.

The current stabilizes; the crystal structure of quartz can be considered as frozen. The temperature is then raised at a rate of 3 K/min and a measurement of thermionic current is made [12,13,14].

3.2. Standard quality samples of quartz (n°511), one swept in air and the other swept under vacuum.

A measurement of thermionic current has been carried out on a quartz sample swept under vacuum and another on a sample swept in air. Figure 2 shows two curves of current densities; they are functions of the temperature. The two crystals were cut near to the cathode.

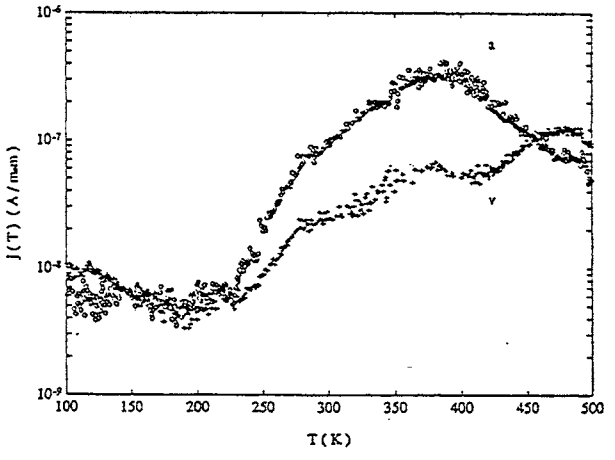


Figure 2: current density T.S.D. for a plate of 511 sample swept under vacuum (v) and for a plate of 511 sample swept in air (a).

The current density of the sample swept in air is higher than that of the sample swept under vacuum.

Let us consider that all the Al^{3+} substitutional ions are initially compensated by alkali ions M^+ [15]. When the sweeping was carried out under vacuum, an alkali ion left the Al^{3+} trap and was replaced by a proton H^+ previously present inside the crystal. For the sweeping in air, another source of protons exists: the water vapour present in the atmosphere. The protons can penetrate inside the crystal at the anode junction [15,16]. Finally at the end of a sweeping under vacuum, the concentration of alkali ions decreased, whereas the one of the hydrogen remained constant. On the other hand, at the end of a sweeping in air, the concentration of alkali ions decreased but that of the hydrogen increased. This can explain the fact that the current density of the sample swept in air is greater than that of the sample swept under vacuum.

The current density curve of the plate swept under vacuum ((v) on Figure 2) possesses three maxima at: 278K, 375K and 478K. The other curve ((a) on Figure 2) has a large peak of which the maximum is at 375K. This peak seems to include the other peaks.

3.3. Influence of some parameters on the measurement of thermionic current.

3.3.1. Evolution of current density as a function of electric tension.

Figure 3 represents curves of current density which are realized with different levels of polarization. An electric tension of 100V, 200V, 1200V or 1600V was applied to a high quality untreated sample for 14 hours at 350K.

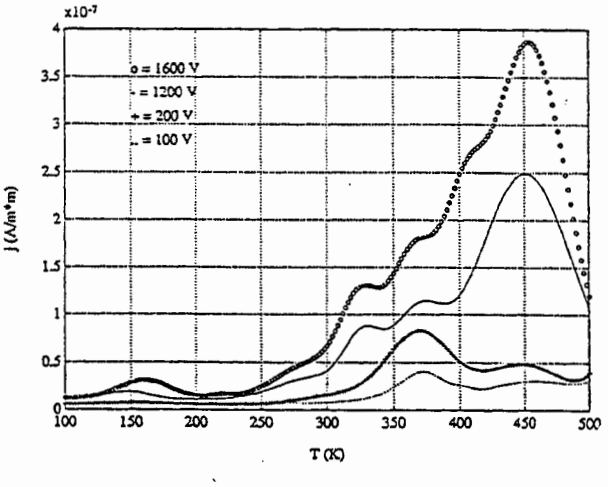


Figure 3: T.S.D. current densities for a sample n°272 untreated. The different applied tensions are 100V, 200V, 1200V and 1600V.

An evolution of current density is obvious: the higher the voltage, the greater the number of moved impurities. Current density measured after a polarization of 100V is small and is not sensitive to the increase in temperature up to about 300K.

3.3.2. Influence of polarization time on the measurement of current density.

Figure 4 shows the evolution of current density as a function of the polarization time at 350K using an electric tension of 1600V.

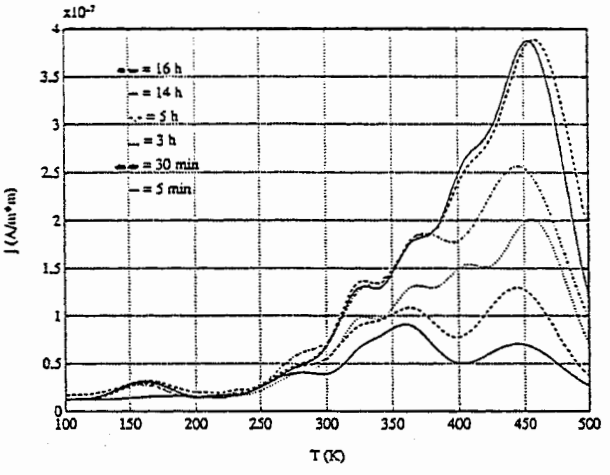


Figure 4: T.S.D. current densities for an untreated sample n°272. The different polarization times at 350K are 5min, 30min, 3h, 5h, 14h and 16h.

The current density increases with the polarization time. However there seems to exist a time beyond which current density increases practically no more. In fact, we establish a few differences between the curves corresponding to the times of 14 hours and 16 hours. Current density measured after a polarization of 1600V for 5min at 350K (figure 4) is clearly greater than that measured after a polarization of 100V for 14h at 350K (figure 3).

4. CONCLUSION.

The obtained results complement the others which have been previously published [12,13,14]. The sweeping in air leaves more ionic impurities which can be later moved in the Z-direction than that under vacuum. In fact, current density measured on the sample swept in air is higher than that measured on the sample swept under vacuum. On the other hand, the sweeping in air and under vacuum leads to the same current density in the Y-direction [13].

The effect of polarization (time and intensity) has been studied: current density increases with the

polarization time and tension. There seems to exist a time beyond which the quantity of moved impurities no longer evolves.

More ionic impurities are moved after a short (5min) polarization of 1600V than after a long (14h) polarization of 100V.

BIBLIOGRAPHY.

[1] KING, BELL SYSTEM TECH. J. 38,573 (1959).

[2] MULLER P., "Relationship between thermally stimulated depolarization and conductivity", PHYS. STAT. SOL. (a) 23,165 (1974).

[3] MULLER P., "Electric relaxation currents in high-resistivity solids", PHYS. STAT. SOL. (a) 67, 1, 11-65 (1981).

[4] PLATA J., BRETON J. and GIRARDET C., "Theoretical model for the electrodiffusion of M^+ (M = Li, Na, K) ions in a quartz crystal.", PHYSICAL REVIEW B, vol. 38, n°5, 3482-3493 (1988).

[5] ZECCHINI P., YAMNI K. and VIARD B., "Modification induced by sweeping and observed in quartz crystals according to their α_{3500}-value.", J. de PHYSIQUE, vol. 4, Suppl. JP III, n°2, 189-194, (1994).

[6] FRASER D., PHYSICAL ACOUSTICS V (ed. MASSON 1968), 59-108.

[7] MARTIN J.J., "Aluminum-related acoustic loss in AT-cut quartz crystals.", J. APPL. PHYS., vol. 56, n°9 (1984).

[8] PARK AND NOWICK, "Dielectric relaxation of point defects in α-quartz" , PHYS. STAT. SOL. (a) 26, 617 (1974).

[9] BACHHEIMER J.P., "Thermal (500-580°C) stability of hydrogen in synthetic quartz", J. de PHYSIQUE, vol. 4, Suppl. JP III, n°2, 173-176, (1994).

[10] CHATAGNON B.,"La résonance paramagnétique électronique du centre E'_1 dans le quartz.", Thèse I.N.P.L. Nancy (1986).

[11] BUCCI C., "Ionic thermocurrents in dielectrics", PHYSICAL REVIEW, vol.148, n°2, 816 (1966).

[12] JEANDEL G., WARIN A., POIGNON C., AUBRY J.P., MORLOT G., "Mesure de courants thermoioniques dans le quartz", PROC. OF 6th EUROPEAN FREQUENCY AND TIME FORUM, Noordwijk, NL, 389-392,.(1992).

[13] POIGNON C., JEANDEL G. & MORLOT G., "Electric relaxation by movement of impurities in quartz" , PROC. OF 7th EUROPEAN FREQUENCY AND TIME FORUM, Neuchâtel, 667-670, (1993).

[14] POIGNON C., JEANDEL G. and MORLOT G., "Mobilité des impuretés ioniques dans le quartz.", J. de PHYSIQUE, vol. 4, Suppl. JP III, n°2, 159-167, (1994).

[15] HALLIBURTON L.E., "Defect models and radiation damage mechanisms in alpha-quartz.", Cryst. Latt. Def. and Amorph. Mat., vol.12, 163-190, (1985).

[16] PLATA J., BRETON J., ALVIRA E., DELGADO V. and GIRARDET C., "Surface effects on the electrodiffusion of alkali-metal ions and protons in quartz.", PHYSICAL REVIEW B, vol. 43, n°6, 5053-5062, (1991).

SESSION 12

ATOMIC STANDARDS II

Chairman : A. DORENWENDT

FIRST RESULTS ON THE HIGH C-FIELD CESIUM RESONATOR AT THE POLITECNICO DI TORINO

G. Costanzo, E. Rubiola, A. De Marchi
Politecnico di Torino, Torino, Italy.

Abstract

The first results of the high C-field Cesium beam experiment under development at the Politecnico di Torino are here reported, together with a brief review on the crucial problems of this scheme and on the approach taken for their solution. The first Cesium resonance curve is shown, and improvements that are likely to be achieved in the near future are indicated. No departure from the traditional description of the hyperfine Zeeman effect, based on the Breit Rabi law, is unambiguously confirmed by measurements at the present level of uncertainty.

1 Introduction

The philosophy and the main features of the high C-field concept for an accurate Cesium beam frequency standard have been discussed in several papers [1, 2, 3, 7] and will be recalled here only by very general lines. An account of experimental work is given in [4].

The standard works on the $m_F = -1; \Delta m_F = 0$ transition at the magnetic field for which its frequency is minimized. From the Breit Rabi law this field value is expected to be 81.987 mT, and the transition frequency is expected to be $\nu_{-1min} = \nu_{00}\sqrt{15/16} \approx 8.900727438257$ GHz. The quadratic coefficient at the minimum is likewise expected to be $4.4147 \cdot 10^{10}$ Hz/T^2, or 441.47 Hz/gauss^2.

The microwave cavity employed in the prototype under construction is a single interaction cylindrical multilambda Rabi cavity, which was analyzed theoretically and shown to be capable of guaranteeing accuracies of the order of 10^{-14} [2, 3]. It resonates on the TE_{017} mode and is expected to have a very flat phase distribution across the central area around the axis, as shown in [5], except for the section of cavity around the feeding port, where it should behave more like is shown in [6]. The use of a very wide Cesium beam is encouraged by this configuration. The beam apertures in the realized cavity are cylindrical, 8mm in diameter and 13mm long. Longer cutoff sections will be employed in the near future.

The magnet producing the C-field is described in detail in [7], and was measured with an proton NMR probe to be uniform to better than 10^{-3} for the extension of the cavity (0.4 m). The magnet used in this experiment is the second prototype that was realized, as the first turned out to be wound on a leaking vacuum containment Copper tube and, since it could not be recovered, had to be discarded. In the new magnet a continuous correction winding was tightened outside the cooling jacket, with taps at each turn for shim current injection. A few more details on the magnet and its uniformity measurement and optimization are given in the following section 2.

[0]This work was partially founded by:

- Consiglio nazionale delle Ricerche,
- Agenzia Spaziale Italiana,
- Istituto Nazionale di Fisica della Materia.

State selection and analysis are made with gradient magnets with a wide V shaped gap, which produce a gradient of up to 50 T/m along the symmetry axis of the V. This is sufficient to adequately separate at the detector even an 8mm wide beam when the magnet is used as the A (state preparation) magnet. Unfortunately, the same cannot be stated for the use as a B (analyzing) magnet, because the obtainable angular separation between the trajectories of the two hyperfine levels is not sufficient to clearly resolve the two beams at the closeby detector. For this reason the signals observed in the first experiments were very small, resting on a big background. Substitution of the analyzing magnet with optical detection is expected to yield in the near future much better fringe contrast and signal to noise ratio.

The first Cesium resonance signals recently obtained from this apparatus are shown and discussed in section 3.

2 C-field uniformity

The main source of inaccuracy in the high field configuration is expected to be the uncertainty in the assessment of the frequency bias introduced through the quadratic Zeeman effect by the magnetic C-field inhomogeneities along the beam path within the cavity. An rms uniformity of $3{\cdot}10^{-6}$ is needed for a bias of $3{\cdot}10^{-14}$, and an rms uncertainty in its measurement of $3{\cdot}10^{-7}$ results in an uncertainty in the estimate of the bias of $3{\cdot}10^{-15}$.

In order to minimize this bias, accurate measurements of the magnetic field spatial variations are needed, and shim coils must be optimally placed and excited according to some optimization procedure. The first task was achieved by steper motor driven longitudinal scanning of a proton NMR probe along the axis of the magnet. The second was approached both in a heuristic way and through a rigorous mathematical optimization procedure [8, 9].

In order to minimize the uncertainty in the field uniformity that might be introduced (after measurements are taken) by mechanical instability with time and varying temperature gradients, the Copper winding was epoxied down during assembly, and temperature stabilization of the whole magnet is being planned.

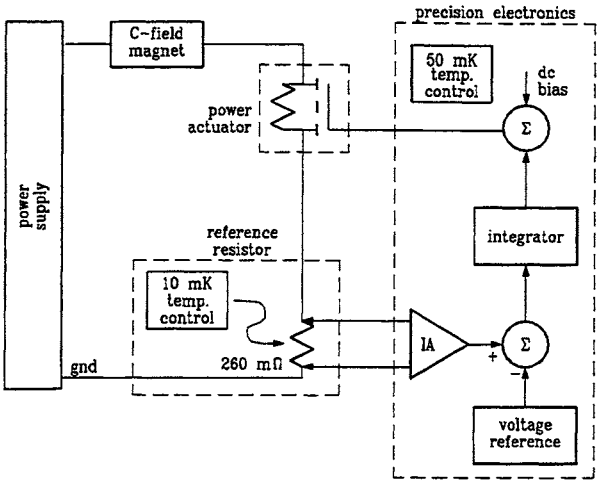


Figure 1: Block diagram of the current control.

2.1 Current stabilization

In order to measure C-field uniformity with resolution of the order of 10^{-7}, the field — and consequently the exciting current — must be stabilized at least at the same level for the total measurement time. Therefore, a current stability of about 10^{-7} must be guaranteed for averaging times ranging from a few seconds to about two hours.

Current stabilization is achieved by load regulation according to the scheme of Fig. 1. A signal proportional to current is sensed at the voltage terminals of a reference resistor, amplified and compared with a voltage reference in the summing node of another amplifier. The output of this is the error signal, which is integrated in order to drive the power actuator.

The overall design guidelines can be drawn from a preliminary analysis, as follows. The reference resistor must have a stability of 10^{-7} for a contribution of 10^{-7} to current stability. In non-climatized laboratory conditions, voltage measurements seem to be limited by undesired thermocouples that can originate emf variations up to about 100 nV; therefore, a specification of 100 nV stability for the electronic circuitry seems to make sense. Consequently, a voltage of 2 V must be available across the reference resistor for a current stability of 10^{-7}. Under these

hypotheses, the minimum resistance value turns out to be 200 mΩ.

A 0.26 Ω reference resistor was built for this purpose, keeping in mind that its power dissipation will be about 20 W in operating conditions. Two manganin wires are wound around an alluminium cylinder and thermally coupled to it. Two Peltier devices stabilize this cylinder at the optimum temperature of 31.2 °C within 10 mK. Around this optimum, the R vs. T characteristics of the resistor is quadratic, with a second order coefficient of $-4.5{\cdot}10^{-7}$ K^{-2}. Our 0.26 Ω resistor was compared to a 0.1 Ω standard in a classical Kelvin bridge configuration. The temperature of the standard was monitored in order to remove its uncertainty contribution. The resistance stability turned out to be $1.2{\cdot}10^{-7}$ and is satisfactory. In fact, this value is conservative since it includes also at least some uncompensated thermal effects, zero detector noise and power supply residual instability.

As for the precision electronic circuitry, the instrumentation amplifier and the summing amplifier turn out to be the most critical stages, while the voltage reference LTZ1000 can be considered ideally stable (long term stability $3{\cdot}10^{-7}$ and thermal drift $2{\cdot}10^{-8}$/°C).

The I.A. and the summing amplifiers are based on a selected version of OP177, which shows a thermal drift smaller than 100 nV/°C, open loop gain of 130 dB and CMRR of 120 dB. By careful resistance matching, a common mode rejection ratio of 110 dB was obtained in the final circuit. The I.A. gain and the summing amplifier gain ratio are directly responsible for current stability; since they rely upon resistance ratios, these ratios must be kept constant within 10^{-7} (in the proposed implementation, useful resistance values should be in the 1–10 kΩ range for best operational amplifiers stability). After a few attempts made with thermally coupled Vishay resistors (1 ppm/°C) and house-built manganin resistors arrays, it became clear that in our case a ratio stability of 10^{-7} needs temperature stabilization or compensation. It was then decided to temperature stabilize all the precision circuitry.

The power actuator, made with a 0.56 Ω resistor in parallel to an IRFP254 power MOS seemed to be a good choice because it has zero driving current and small dynamics. This last feature is useful in case of control failure and it makes the circuit safe for human intervention without switching the power off, despite of the high energy stored in the magnet (10 J).

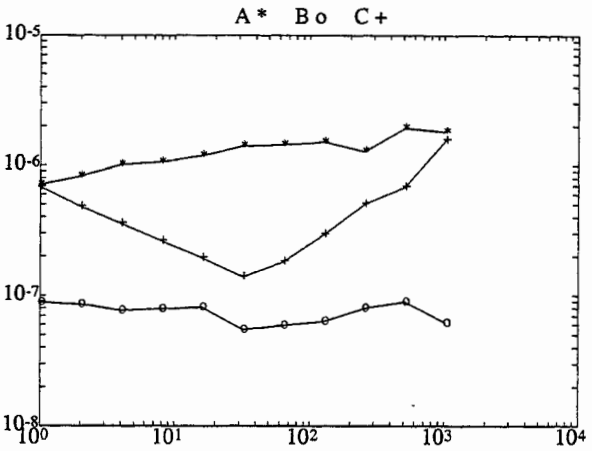


Figure 2: Measurement results. A: open loop voltage. B: closed loop voltage. C: closed loop magnetic field.

Experimental results, described in terms of Allan deviation $\sigma(\tau)$, are reported in Fig. 2 and commented below.

Curve A is the stability of the voltage across the reference resistor in open loop conditions, which is limited by the power supply.

Curve B is the stability of the voltage across the reference resistor in closed loop conditions. The current stability may be slightly worse, as it may be affected by resistor instabilities.

Curve C is the stability of the magnetic field, measured with the NMR probe resting in the center of the magnet. The leftmost part, with slope -1, shows the resolution of the NMR frequency measurement. For averaging times longer than about 10^3 s, the stability of the field seems to be limited by the stability of the magnet itself, probably because of temperature fluctuations.

2.2 Uniformity optimization

The field uniformity measured as described, under closed loop current control, on the axis of the magnet as it was built (that is without any correction shim coil) turned out to be about $3{\cdot}10^{-4}$ rms. In order to show a typical result as can be obtained with shims, a uniformity comparison

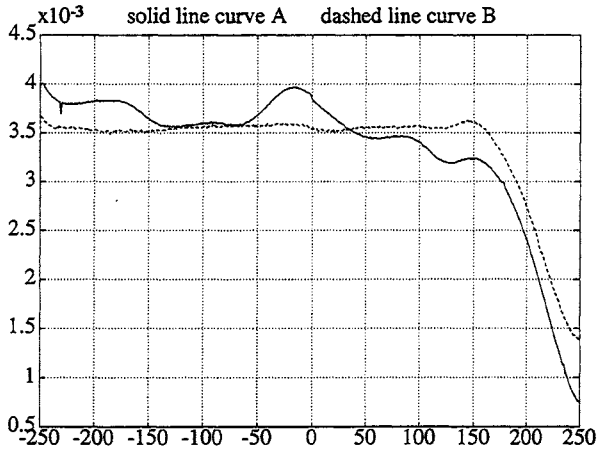


Figure 3: Uniformity measurement results. A: no correction coils. B: 18 correction coils excited by 7 different current values combined in different ways. The horizontal axis gives the distance in mm from the centre of the magnet. The vertical axis is $\delta B/B$.

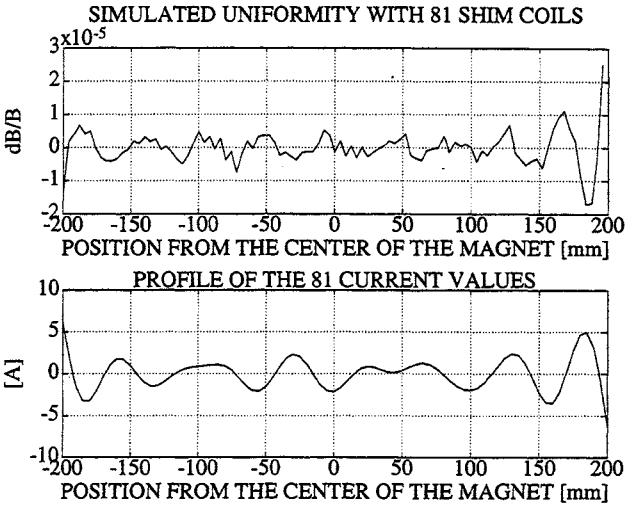


Figure 4: Uniformity simulation. Top: simulation results. Bottom: longitudinal profile of the 81 current values needed to obtain the uniformity shown.

with and without correctin is reported in Fig. 3 for the first magnet that was realized. No such result has yet been achieved for lack of time with the new magnet.

The optimized uniformity shown in Fig. 3 was obtained with a trial and error procedure, assisted by calculations based on measured field profiles for the single coils which were obtained by difference between two scans of the magnet. The obtained rms uniformity was evaluated to be about $5{\cdot}10^{-4}$ for the uncorrected magnet, and $1.5{\cdot}10^{-5}$ after correction.

The more rigorous optimization procedure described in [8, 9] has led for the moment being only to simulations, which appear nevertheless quite encouraging. A simulated result is shown in Fig. 4 together with the longitudinal current profile that is needed to obtain it. The best rms uniformity projected at this point is about $5{\cdot}10^{-6}$, which would be right on target.

Evaluation of the frequency bias introduced in operation by field disuniformity should be feasible with an uncertainty of $3{\cdot}10^{-15}$, since field measurements are made with a resolution of $3{\cdot}10^{-7}$. However, perplexities still remain about the possibility of guaranteeing mechanical stability of the magnet to the level necessary to build confidence in the fact that the field uniformity during the Cesium beam standard operation be the same as what was measured with the NMR probe before assembly.

3 First results

The first resonance signals obtained from the apparatus consist of a very small flop-out feature (less than 0.4 mV_{pp}) on top of a background of a couple of hundred mV. These voltage values are developped by the detector current on 1 GΩ. A typical trace is shown in Fig. 5.

In these first experiments no shim coils were excited to improve field uniformity, and the tube was positioned vertically, with the Cesium source down bottom and the hot wire detector up top. This was done mainly for laboratory size reasons, despite of the fact that no theory has yet been developed to describe the first order Doppler effect in the multilambda Rabi cavity with varying atomic velocity. In the present case, however, with a design target beam velocity of 100 m/s (not measured), the relative velocity variation from the first to the second half of the cavity is a few parts in 10^4, and should therefore be of no concern, since it was shown in [2, 3] that a 10^{-2} difference in the conductivity of the two ends is acceptable for Doppler shift cancellation at the level of 10^{-14}.

In order to study the variations of resonance

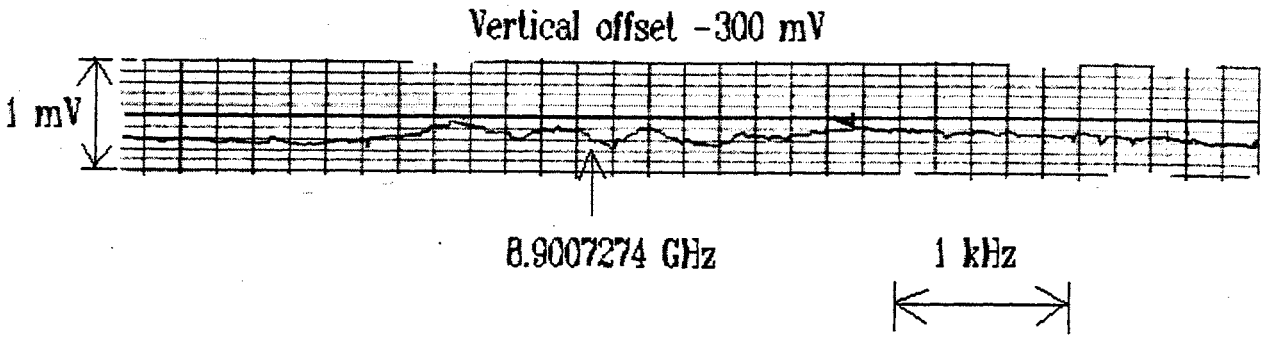


Figure 5: Cesium resonance signal. The sweeping time was 10 s/kHz and the time constant was about 0.3 s.

frequency with magnetic field, the latter was trimmed by passing a current from a separate power supply in the whole correction winding. Measured frequency shifts are plotted in Fig. 6 versus magnetic field. The uncertainty of the latter is limited by the 1% accuracy with which the trimming current was measured. The vertical scale gives frequency differences in Hz from the frequency of the minimum. The frequency of the minimum itself is uncertain by a few hundred Hz due to the fact that the central interference fringe could not be unambiguously identified because the signal was too poor. At the present uncertainty level no departure from the Breit Rabi law can be inferred with any level of confidence.

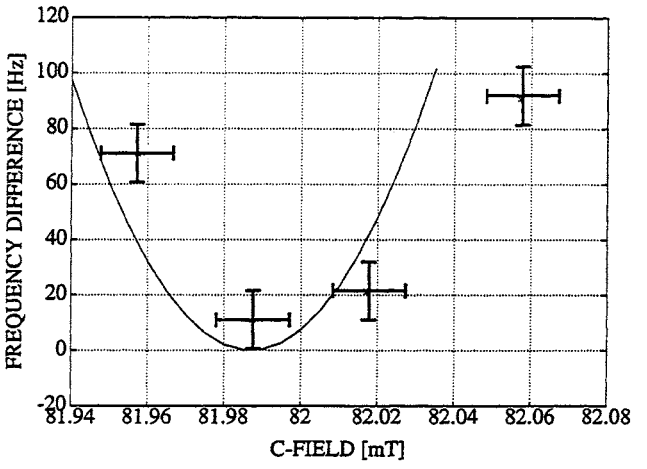


Figure 6: Measured resonance frequency variations with field around the minimum. The parabola represents the theoretical behavior according to the Breit Rabi law. The vertical positioning of experimental points, relative to the calculated curve, is uncertain by a few hundred Hz.

4 Conclusions

This paper reports improvements introduced in the high field Cesium beam experiment after the 7^{th} EFTF meeting in Neuchâtel and the first results obtained for the microwave resonance signal. The disappointing contrast of the observed resonance fringes appears limited by the poor performance of the adopted gradient magnets design and/or realization. Better signals are expected as soon as this problem is overcome. The uncertainty left by available observations in the description of the Zeeman effect around the minimum does not allow to identify unambiguously any departure from the Breit Rabi law. Better signals are needed for that purpose.

Acknowledgments

The authors would like to acknowledge discussions and help from many colleagues, in particular D. Rovera of LPTF, R. Drullinger of NIST and L. Cutler of HP.

References

[1] A. De Marchi, "The high C-field concept for an accurate cesium beam resonator," *Proc. 7^{th} EFTF* p. 541–548, Neuchâtel, Swiss, 1993.

[2] A. De Marchi, E. Bava, P. Tavella, "Accuracy analysis of the atom-microwave interaction in a TE_{017} cylindrical Rabi cavity for a high C-field cesium beam resonator," *Proc. 7th EFTF* p. 489–493, Neuchâtel, Swiss, march 16–18 1993.

[3] E. Bava, A. De Marchi, P. Tavella, "Prospects for the use of multilambda Rabi cavities in atomic beam frequency standards," *Proc. 47th FCS* p. 109–113, Salt Lake City, USA, jun 2–4 1993.

[4] G. Costanzo, *Ottimizzazione di un campione atomico di frequenza a fascio di cesio*, PhD Thesis, Politecnico di Torino, 1995.

[5] G. Vecchi, A. De Marchi, "Spatial phase variations in a TE_{011} microwave cavity for use in a cesium fountain primary frequency standard," *IEEE Trans. on Instrum. & Meas.*, vol. 42 no. 2, p. 434–438.

[6] A. Khursheed, G. Vecchi, A. De Marchi, "On the dependence of spatial phase variations with the RF source in cavities for cesium fountain standards," *Proc. 9th EFTF*, Besançon, France, march 8–10 1995.

[7] A. De Marchi, G. Costanzo, C. Lombardi, A. Del Casale, F. Maddaleno, E. Rubiola, M. Repetto, M. Chiampi, "High uniformity magnet for a high C-field cesium beam resonator," *Proc. 7th EFTF* p. 367–371, Neuchâtel, Swiss, march 16–18 1993.

[8] G. Costanzo, M. Nervi, M. Repetto, F. Levati, "Use of automated optimisation techniques in the design of the magnetic system for a caesium frequency standard," *3rd Workshop on optimization & inverse problems in electromagnetism*, CERN, Geneva, Swiss, september 1994. To be published in *International Journal of Applied Electromagnetics in materials.*

[9] G. Costanzo, A. De Marchi, M. Nervi, M. Repetto, "High homogeneity solenoidal magnet for Cs frequency standard," *IEEE Trans. on Mag.*, vol. 30 no. 4, july 1994, p. 2628–2631.

STUDY OF THE S/N PERFORMANCE OF PASSIVE ATOMIC CLOCKS USING A LASER PUMPED VAPOUR

Gaetano Mileti and Pierre Thomann

Observatoire Cantonal Neuchâtel - 58, rue de l'Observatoire, CH-2000 Neuchâtel

Abstract

We have studied theoretically and experimentally the potential of replacing the discharge lamp by a laser diode in a miniaturised rubidium clock. Our results constitute a significant improvement of the resonance signal amplitude reached in such devices (of the order of 20% of the DC photo current) in comparison with signals produced with a discharge lamps (0,5% typically). Our measurements also show that the shot-noise limit can be reached even with commercial laser diodes, thus yielding $\sigma_y = 10^{-13}\ \tau^{-1/2}$ *as a fundamental limit for this kind of miniature standards. However, the frequency noise of the laser can degrade this performance through the light-shift. We will report on some critical aspects of this problem, such as the operating frequency and the conditions of stabilisation of the laser.*

I. Introduction

In parallel with the development of conventional rubidium frequency standards for industrial and space applications /1/, ON is investigating on the potential of using a laser instead of the traditional discharge lamp /2/. We report in this paper the results of our study concerning the achievable frequency stability with available commercial laser diodes. Even if several authors already explored this field /3/, some obscure domains still exist, for example the discrepancy between expected and actually measured stability /4/. Furthermore, a quantitative analysis of the improvements that this approach could bring, and an identification of the critical parameters, have not been presented yet.

First, it appeared necessary to understand how the laser properties - such as the spectral distribution of light, the intensity noise or the frequency noise - influence the characteristics of the double resonance signal. This comprehension is possible only on the basis of a theoretical model that can predict quantitatively the signal as a function of the different experimental conditions. As the laser diode field is still developing, the availability of a simple but complete model will simplify the task of estimating the impact of a type of laser present in the market.

The paper is organised in the following way. Section 2 reminds the theoretical and historical background concerning miniature optically pumped gas-cell atomic clocks. Section 3 presents the experimental set-up and some typical results. These results are then interpreted and discussed in Section 4. Section 5 collects our conclusions.

II. Theoretical and historical background

Predicting the stability of an atomic clock can be divided in two steps. The first step consists in modelling the response of the atomic medium in order to describe quantitatively the resonance signal. The second step consists in deducing from this signal the achievable performances of the clock. This second aspect is now pretty well-known, especially because both the resonance signal and the stability can be measured experimentally, so that theory and experiments can be compared. As a result, predicted and observed values of the Allan variance, for a given detected signal, are approximately the same. It is still not the case for the first step. Especially with optically pumped rubidium frequency standards, the number of different and simultaneous atomic processes together with the strong inhomogeneous aspect of these interactions, obstructed quantitative comparisons between theoretical and experimental signals.

One of the main complications in the theoretical description of the atomic system in a passive rubidium clock came from the spectrum of the irradiating beam. In a miniaturised device, the integrated filter approach imposes optical pumping conditions that are nearly impossible to model in any realistic way.

The use of laser sources instead of the traditional discharge lamps did not only bring a potential improvement of the performances of passive vapour clocks, it also - in principle at least - simplifies the theoretical description of the full process of optical-microwave atomic double resonance. Furthermore, since most of the effort has already been made in the previous studies of conventional clocks and is now available /5/, it was just a matter of reducing the formulas at an appropriate degree of complexity.

We found several publications on this subject, but not a clear separate treatment of the two quoted aspects : understanding the measured resonance with an atomic model first and secondarily predicting the achievable stability . These were the objectives of the research presented here. The next step, depending on the technological developments in the laser field, would be to build a real full clock optimising the desired performances, on the basis of this understanding.

Atomic response

The atoms in the vapour are subjected to different interactions apart from the two resonant electromagnetic fields. Primarily there is the so-called C-field, which is used to separate the field-independent transition (clock transition) from the other field dependent transitions. Secondly there are collisions between the interrogated atoms (spin exchange), or with the walls of the cell, or with a foreign buffer gas. All these collisions create relaxations of the atoms and define the physical parameters called longitudinal and transverse relaxations rate γ_1 and γ_2, which determine the ultimate line Q and strength of the double resonance signal. These important parameters can be calculated (/5/ and /6/) and one generally optimises the buffer gas pressure and nature, in order to reduce the rate of wall collisions and minimise the temperature coefficient. The smaller the cell, the more the buffer gas pressure has to be increased, and so, the more the line Q will be reduced. Note that C-field inhomogeneities

also increase the relaxation rates and affect the signal. In conventional rubidium standards, the foreign gas in the lamp, in the filter and in the cell are also adjusted in order to minimise light-shift effects. With the laser, the buffer gas in the absorption cell can also influence the light-shift, but in a different and more controlled way (see remarks on long term stability in Section IV).

Typical values of the relaxation rates encountered with 1 cm^3 rubidium cells filled with nitrogen and/or argon are comprised between 700 and 1000 s^{-1}. Using these parameters, the resonance signal produced by a small volume of vapour can be calculated easily, using formulas available in the literature.

In the case of laser pumping, we can simplify these formulas by the fact that the light is resonant only with one of the D1 or D2 lines, and suppose that when the laser is tuned to one of the hyperfine level (for example F=2 if Rb87 is used), the absorption rate corresponding to the other transition is negligible. For a three level atom, the variation of light intensity in an optically thin slice of vapour of length Δl and a density n of rubidium atoms is given by equation (1).

$$\Delta I = -n\,\Delta l\,h\,\nu_{Rb}\left[\frac{\Gamma}{2} - \frac{\Gamma^2}{4\gamma_{//}} + \frac{\Gamma^2}{4\gamma_{//}}\frac{\Omega^2}{(\omega-\omega_{Rb})^2+\gamma_\perp^2+\Omega^2}\right] \quad (1)$$

The first term is the absorption in absence of optical pumping, the second term is a reduction of this absorption due to optical pumping, and the third term is an additional absorption dip (double resonance signal) created by the microwave repopulation of the absorbing hyperfine sublevel. The meaning of the parameters is indicated below.

$\Gamma=(h\nu_L)^{-1}\int I_L(\nu)\sigma_{Rb}(\nu)\,d\nu$ *pumping rate*	$\Omega^2=4\omega_1^2\gamma_\perp/\gamma_{//}$ ω_1 : *microwave Rabi frequency*
$\gamma_{//}=\gamma_1+\Gamma/2$ *total longitudinal relaxation rate*	ω_{Rb} *Rubidium clock frequency*
$\gamma_\perp=\gamma_2+\Gamma/2$ *total transverse relaxation rate*	ω *Interrogating frequency*

To describe the real rubidium 87 atoms, we have to take into account the fact that there are eight ground-state sublevels instead of two, and much more than one excited sublevel, depending on which optical line is used. For example, if we consider the D2 transition of Rb87, we should consider 56 absorption rates (or atomic cross sections $\sigma(\nu)$). With a few approximations however, this number can be reduced to two absorption rates. First one has to use the symmetry properties of the atomic transfer coefficient regarding the light polarisation and the Zeeman sublevels of each hyperfine levels. These properties allow to deal with only six cross sections corresponding to the three allowed transitions for each hyperfine level in the ground state (F=1→F'=0, F=1→F'=1, F=1→F'=2, F=2→F'=1, F=2→F'=2, and F=2→F'=3). Each one of these frequency dependent cross sections have different amplitudes and centre frequencies. Because of Doppler and collisional broadening, the three absorption cross sections for each hyperfine ground state sublevel can be summed as if they were all centred on an average frequency. Note that this approximation corresponds to the measured absorption spectra of the rubidium vapour.

Thus, the local variation of light intensity with real Rb87 atoms can be expressed in a form similar to (1) but depends on which hyperfine sublevel is depopulated by the laser. In the case we depopulate F=2 we obtain (2), with a different meaning of the total longitudinal relaxation rate, and a factor q, given below.

$$\Delta I = -n\,\Delta l\,h\,\nu_{Rb}\left[\frac{5\Gamma}{8} - \frac{15\Gamma^2}{8(3\Gamma+8\gamma_1)} + \frac{q\,\Omega^2}{(\omega-\omega_{Rb})^2+\gamma_\perp^2+\Omega^2}\right] \quad (2)$$

$$q=\frac{4\Gamma^2\gamma_1(\Gamma+\gamma_1)}{(\Gamma^2+7\Gamma\gamma_1+8\gamma_1^2)(3\Gamma+8\gamma_1)} \qquad \gamma_{//}=\frac{\gamma_1(\gamma_1+\Gamma)(3\Gamma+8\gamma_1)}{(\Gamma^2+7\Gamma\gamma_1+8\gamma_1^2)}$$

The light intensity transmitted by the cell can be computed as a function of the interrogating frequency. The theoretical resonance signal is obtained by solving the differential equation (3).

$$\frac{\Delta I}{\Delta l} \approx \frac{dI}{dl} = F(I,n,\omega_1,\omega,l) \quad (3)$$

The solution of (3), evaluated at *l=length of the cell* is the signal used to lock the quartz-crystal oscillator in a rubidium frequency standard. From its characteristics, from the type of perturbations and noises and from the type of interrogation-detection scheme, it is possible to deduce the achievable frequency stability /5/.

Short term stability

In a passive lamp pumped rubidium frequency standard, the ultimate short term stability is usually limited by the shot noise level at the detector /5/. As the shot noise depends on the square root of the incident light intensity, the Allan deviation can be expressed as a function of the spectral density of noise of the relative frequency deviation as :

$$\sigma_y(\tau)_{shot\text{-}noise} = \sqrt{\frac{S_{yy}^{shot\ noise}}{2}}\cdot\tau^{-1/2} = \frac{\sqrt{2\cdot e\cdot I\cdot\eta}}{\sqrt{2}\cdot D\cdot\nu_{Rb}}\cdot\tau^{-1/2} \quad (4)$$

$$\approx 1\times10^{-10}\cdot\frac{Shot\ Noise}{Discriminator\ slope}\cdot\tau^{-1/2}$$

If we assume a 1 cm^2 detector with a quantum efficiency η of 0.5 A/W, the typical measured photocurrent is 150 µA, giving a shot noise of 7 pA/√Hz. At this light intensity, the discriminator slope varies between 0.4 and 0.8 nA/Hz, and the limit short term stability is then :

$$\sigma_y(\tau)_{shot\text{-}noise} \approx 1-2\times10^{-12}\cdot\tau^{-1/2} \quad (5)$$

This stability has actually been achieved and represents the state of the art in compact and miniature Rb standards /7/. It is important to mention the fact that the discriminator slope D used in the above formula depends on the shape of the signal and on the exact modulation/demodulation parameters of the interrogation scheme /5/.

Replacing the discharge lamp with a laser diode implies a reduction of the spectral width of the pumping light. It is so expected that much less light would produce a similar resonance signal. As a result, the shot noise level would fall back to the pA/√Hz level. This more efficient optical pumping implies a potential increase of the discriminator slope, which could put the potential Allan Variance smaller than $10^{-13}\ \tau^{-1/2}$. The present study investigates on the possibility and the limits of such improvements.

III. Experimental results

Our experimental work represents an investigation of the parameters involved in the short term stability of a future laser pumped rubidium frequency standard. We were trying to give an answer to the following questions :

1. Is it possible to reach the shot noise level with commercial laser diodes ? What is the effect of laser intensity noise ?
2. What effects are to be expected from the frequency noise of the laser and from the light-shift ?
3. What double resonance signal level is achievable with a laser instead of a lamp, and how does it compare to the theoretical predictions ?
4. What are the most important criteria to consider for the stabilisation of the laser ?

It was obvious from the beginning that no exhaustive answers could be given. However, the availability of a finalised conventional rubidium based on the same experimental set-up could offer us useful comparisons between laser pumped and lamp pumped devices.

Practical implementation

The experimental set-up is essentially the same described in our previous study on light-shift /2/ (1993). We used commercial compact-disk type laser diodes tuneable to the D2 rubidium lines at 780 nm. For the purpose of this study, cylindrical cells (length and diameter : 14 mm) containing Rb^{87} and 30 mbar N_2 were used. The cell temperature was varied between 55 and 75 °C. We employed a magnetron type resonant cavity /1/, and a silicon photodetector. The full atomic resonator had a volume of less than 20 cm^3.

Observation of a residual spectrum

Measuring absorption spectra, we observed that the Rb87 vapour could not absorb totally the light emitted by the laser diode (Fig. 1). There is a residual light whose relative intensity lies between 10 and 20%. This percentage is of course bigger at the output of the absorption cell. We verified accurately that it was not stray light, and observed this effect for various diodes from different producers. As the residual light could be filtered out with a narrow interference filter (0,2 nm), we concluded that it came from the spontaneous emission and/or from the longitudinal modes of the laser cavity. As expected, the relative intensity of this residual light increases when the laser current is closer to the threshold.

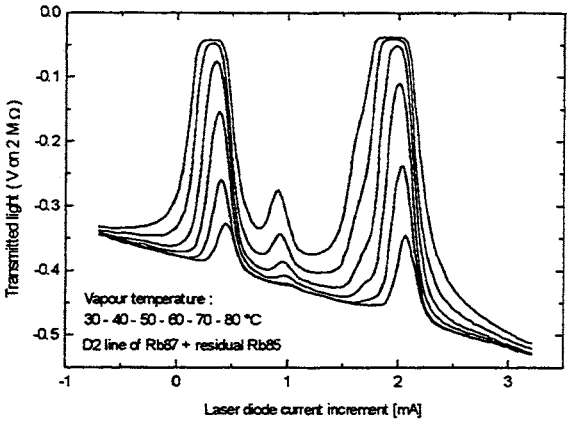


Fig. 1 *Residual spectrum of commercial laser diodes : the emitted light can not be totally absorbed by the vapour. The curves correspond to different temperatures.*

Measurements of intensity noise

The intensity noise has been measured as a function of the photocurrent. Figure 2 displays the PSD for different values of the DC photocurrent, and Figure 3 illustrates the same results for a fixed Fourier frequency (typical quartz modulation frequency). Note that to obtain these results particular attention has been devoted to reduce the feed-back effects. We observe that the intensity noise of the laser (at fixed current and temperature) is larger than the shot noise for a photocurrent bigger than 1 μA (Fig. 3).

The square of the relative noise spectral density at higher photocurrents is the Relative Intensity Noise (RIN), and has been measured for different values of the photocurrent. As expected, it decreases if the current is increased (for a current equal to the double of the threshold the RIN is 1.5 10^{-13} Hz^{-1}).

As the light level which allows operation at the shot noise level depends on the RIN, the optimal laser intensity used in a rubidium clock will depend on this particular property of the diode.

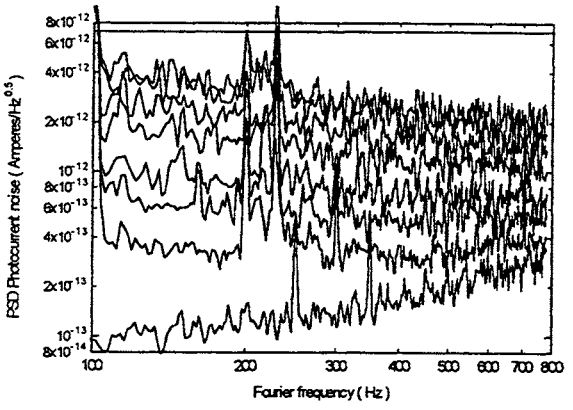


Fig. 2 *PSD for different DC photocurrent : 0 - 0,2 - 0,5 - 1 - 2 - 3 - 4 and 5 μA. The upper line is the lamp shot noise level in a miniature rubidium.*

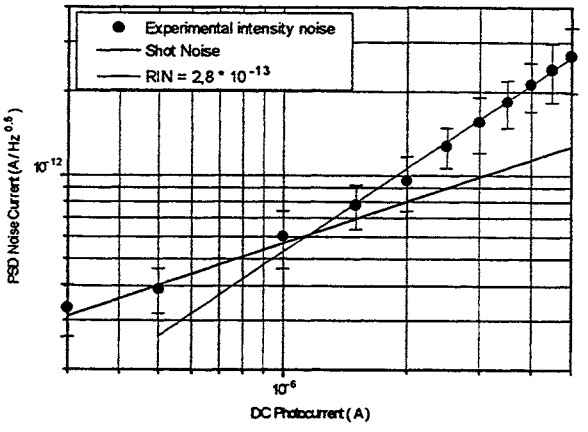


Fig. 3 *Noise @ 300 Hz versus DC photocurrent.*

Measurements of frequency noise

Following a standard method using the rubidium atomic lines as a frequency discriminator, we also measured the frequency noise of the laser. The result is displayed in Figure 4.

The typical value @ 300 - 600 Hz lies between 20 and 30 kHz/√Hz.

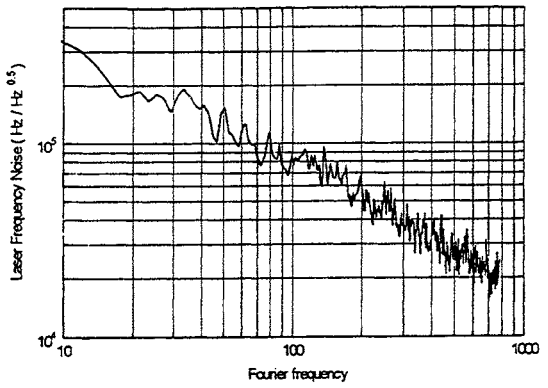


Fig. 4 *Laser diode frequency noise.*

Methods to reduce intensity and frequency noise

The intensity noise of the laser can be - at least in principle - eliminated in an electronic manner. The method consists in using a second beam coming from the same laser but which does not propagate through the atomic vapour. This beam is detected with a separate photocell, and the produced current is either used in an electronic substractor and/or a divider which eliminates all the correlated noises present in the principal photodetector. We did only preliminary measurements with this technique, and encountered difficulties to obtain convincing results when only one of the beam propagates through the rubidium vapour. In fact, the relative importance of the "residual spectrum" in this case is not the same for the two beams, which reduces the cancellation of the noise related to its correlation. If needed, further investigations have to be made in this direction.

The frequency noise measured above corresponds to a free running laser. In any clock, however, the laser has to be stabilised, and we can expect that increasing the bandwidth of the servo-loop for the laser would reduce the frequency noise in the critical spectral region for this application. Figure 5 illustrates the result obtained modulating the laser at 50 kHz. The laser is locked to a Rb 87 line and the frequency noise is measured using a cell containing Rb85 as frequency discriminator.

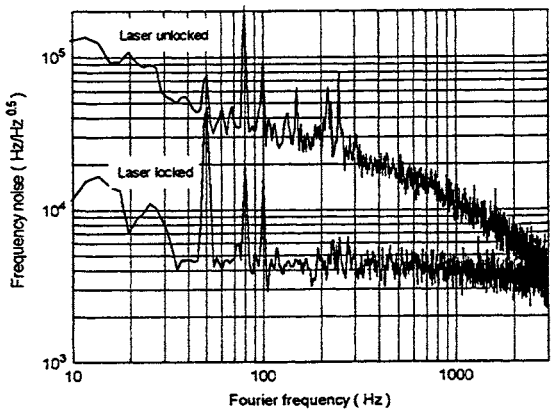


Fig. 5 *Laser frequency noise reduction.*

In a real rubidium clock, frequency noise and intensity noise have to be minimised simultaneously and it is possible that compromises will have to be found. Furthermore, it is also to be expected that - as the laser will be precisely in an atomic absorption line - FM to AM will occur and deteriorate the short term stability of the clock.

Measurement of the discriminator slope

We measured the resonance signal in the following manner. The laser is tuned to the transition 5S-F=2→5P3/2, and the microwave is swept across the resonance. Its amplitude and width are detected as a function of the intensity of the "homogeneous" laser beam. In the range of the considered intensities, the signal amplitude varies from 0 to 2 μA (18 to 27 % of the DC photocurrent), and its FWHM from 700 to 900 Hz. As a comparison, the same device with a discharge lamp produced hardly 1% of signal, and a FWHM ranging from 1.2 and 1.5 kHz. The discriminator slope in miniature conventional rubidiums is typically 0,5 nA/Hz (Amplitude / FWHM = 0,4 - 0,8 nA/Hz). What is important to note is that there is only a slight increase of signal when passing from the discharge lamp to the laser, the improvement lies in the fact that much less pumping light is needed to achieve the same atomic inversion.

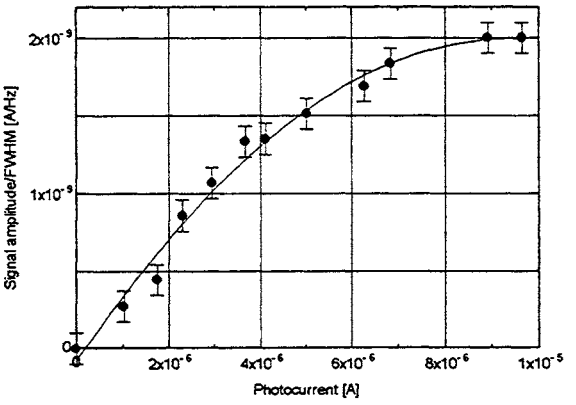


Fig. 6 *Amplitude/width of the resonance signal.*

Measurement of light-shift

We already reported an accurate physical study of the light-shift in a laser pumped rubidium vapour /2/ (1993). In order to concentrate on the effects in a real clock, we measured the light-shift in the corresponding experimental conditions. The frequency of the locked interrogating quartz is compared to a reference clock while the laser frequency is swept across the two main resonances of the D2 line of Rb87. The result, displayed in Fig. 7 yields an experimental value of the light-shift per Hertz of laser detuning and per μA of DC photocurrent. We will express in the following section the achievable stability on the basis of this data and of the frequency noise of the laser. This light-shift measurement will also induce some remarks on the long term frequency drift.

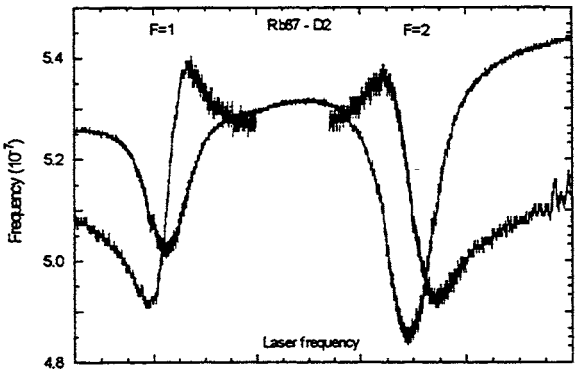


Fig. 7 *Light-shift and transmitted light.*

IV. Theoretical study and discussion

In the first part of this section we present the results of the numerical calculations based on (2). These results are then compared to the experimental results. The second part of the section is devoted to the calculation of the limit short-term stability that can be expected as a function of the RIN and frequency noise of the laser. A third paragraph contains remarks concerning the other performances of a future laser pumped rubidium frequency standard.

Calculated resonance signal

From the width of the measured resonance signal at low microwave and light intensities, it is possible to deduce the transverse relaxation rate γ_2. The longitudinal relaxation rate γ_1 can also be deduced experimentally /8/.

Calculating the resonance signal produced by an optically thin slice of rubidium vapour, one observes that it is theoretically impossible to produce a relative resonance signal larger than some percents of the transmitted light. A relative signal of the order of 20% is the result of a thick vapour, that we can consider as a succession of slices. Each slice produces a signal of some percents of the incoming light. The detector receives the sum of all the signals, but only the residual light transmitted by the last slice. There is an optimal optical thickness, where the cumulated signal becomes of the same order of magnitude as the total transmitted light, and an apparent broadening of the line occurs. This optimum is found experimentally by adjusting the vapour temperature or the cell length. It depends on the buffer gas pressure and composition, which finally determines the dependence of Γ on the number of incident photons.

Figure 8 displays typical results obtained numerically for 1 - 2 cm^3 of vapour. These results correspond to the experimental measurements. In fact, we deduce typically a signal amplitude of 1 μA, corresponding to 20% of the transmitted light, and a discriminator slope of 1 nA/Hz. From these calculations, it is also possible to estimate how many atoms in the vapour are optically pumped in optimal conditions, and how many are subjected to a microwave transition at resonance. These numbers give us an understanding on the atomic scale of the different interactions, and are useful in order to compare crude estimations of the achievable clock stability to more accurate calculations.

Short term stability predictions

As numerical and experimental results for the double resonance signal correspond qualitatively and quantitatively, we can extrapolate the stability achievable with a laser-pumped vapour of Rb87 having a volume of 1-2 cm^3 using the results of Section 3. We used the expression (4) replacing the shot noise by the intensity noise for different values of laser RIN, and the measured discriminator slope. We also derived from the light-shift measurements the effect of the laser frequency noise. Figure 9 summarises the results. The indicated parameters are the RIN of the laser diode and the frequency noise of the laser diode. We also indicated the shot noise limit. For these calculations, we supposed that the laser was tuned to the zero light-shift frequency. We will justify this approximation in the following section.

The disk in Fig. 9 indicates the stability achievable with commercial laser diodes : $2.10^{-13}\ \tau^{-1/2}$. Such a result will be obtained only if the low FM and AM can be realised simultaneously, and no FM to AM conversion occurs. Our preliminary measurements seem to indicate that it is possible.

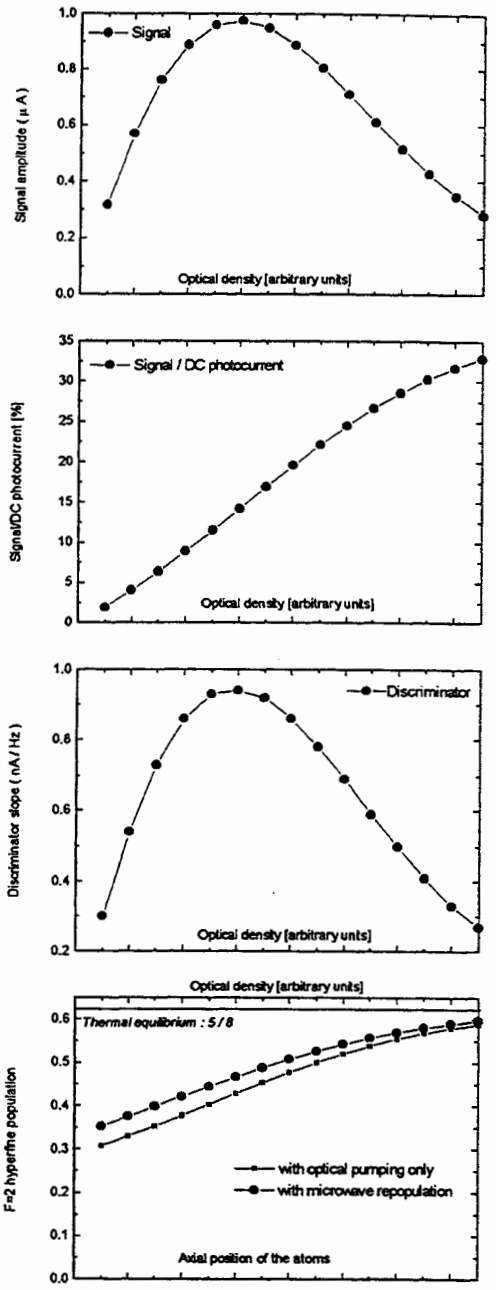


Fig. 8 *Example of numerical results for 1-2 cm³ vapour of Rb87 with 30 mB nitrogen, and a laser diode tuned to the D2 transition, depopulating F=2.*

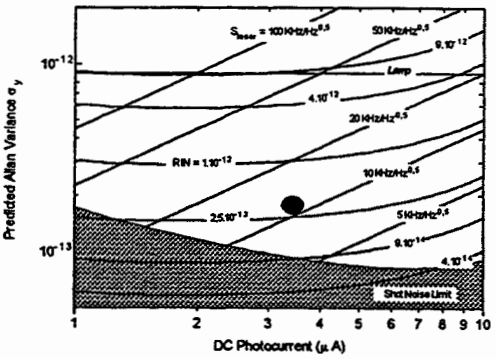


Fig. 9 *Short-term stability predictions. The meaning of parameters is indicated in the text.*

This shot noise limit can then be calculated for higher volumes of Rb87, (Fig. 10) and for different values of the discriminator slope (change of the buffer gas broadening or use of wall coating).

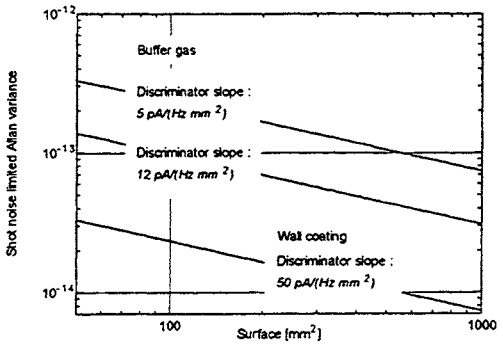


Fig. 10 *Shot noise limit : effect of the volume and of the discriminator slope (L = 10 mm). The measured discriminator slope correspond to the middle curve.*

Remarks on the long term stability

If the results of figure 7 are analysed theoretically, we can see that when the laser is tuned to one of the two resonances, the light-shift produced by the other hyperfine line is not negligible (a dispersive curve attenuates slower than an absorptive curve). We expect from these curves, for example if the laser is tuned to the F=2 line, a detuning of 55.4 MHz to the zero light-shift, and a residual light-shift of 2.41 Hz / μA. This corresponds to the result obtained in a real light-shift measurement, when the laser is effectively tuned to the absorption line centre and the clock transition is measured as a function of the laser intensity. We could also predict and measure a similar consequence for the power-shift (-10-11/dB). However, the required stability of light and microwave power for a good long term performance does not seem to be too restrictive.

These observations are important because they indicates that it could be necessary to operate the laser in a real clock at a detuning of some tenth of MHz with regard to the zero light-shift frequency. We could verify however that, for such laser detunings, the effect on the short term stability of the intensity noise through the light-shift (neglected in the previous section) was minor as compared to the other processes.

Remarks on the other performances

If we look through the standard specifications of miniature rubidium clocks, we can predict the following benefits that the use of laser diodes could bring.

The fact that the absorption cell can be separated from the light-source, allows to consider the diode as a part of the electronic package. Thus, the physics package becomes very small and contains a reduced number of elements (no filtering region). The device used for the present study - for example - had a volume of less than 20 cm³. Apart from the fact that the understanding of the measured effects is easier, the thermal sensitivity should be reduced and have less coupling effects. Since the volume that has to be contained in the magnetic shields is smaller, the magnetic sensitivity should be decreased. This has also a consequence on the weight of the whole unit. The power consumption for heating the cell is reduced, and the warm up should be shortened.

The negative aspects are related to the unknown characteristics of the laser diodes : availability, price, ageing, lifetime and reliability.

V. Conclusions

Our study showed a good agreement between theoretical and experimental double resonance signals obtained with laser pumping of a rubidium vapour. We showed that for a same configuration a slight improvement of the discriminator slope is obtained as compared to lamp pumping. The main difference is an important reduction of the DC level of light, yielding an expected short term stability at the $10^{-13}\tau^{-1/2}$ level (two orders of magnitude better than conventional rubidiums). Our intensity noise and frequency noise measurements suggest that this improvement is possible even with commercial laser diodes. Furthermore, our theoretical analysis illustrates the limits of this technique and gives explicitly the progress that have to be made in order to exploit the potential of using monochromatic light sources. The problem of frequency noise and light-shift has also been studied and related to the short-term and long-term performances of a future standard.

VI. Acknowledgement

We thank Dr Giovanni Busca and Dr Robert Drullinger for stimulating discussions and the collaborators in ON who offered their scientific and technical assistance.

VII. References

/1/ G. Mileti, I. Rüedi, and H. Schweda, (1992), Proc 6th EFTF, Noordwijk , pp 515/19.
P. Rochat, H. Schweda, G. Mileti, G. Busca, (1994), Proc. 8th EFTF **1**, Münich, 385/405.
P. Rochat, H. Schweda, G. Mileti, G. Busca, (1994), Proc. IEEE International FCS, pp /.

/2/ J. T. Liu, P. Thomann, L. Zhang, G. Busca, (1990), Proc. 4th EFTF, Neuchâtel, pp 157/59.
P. Willemin, G. Busca, J. T. Liu, P. Thomann, (1991), Proc. 5th EFTF, Besançon, pp 181/86.
G. Mileti, P. Thomann, (1993), Proc. 8th EFTF, Münich, pp 377/84.

/3/ J. C. Camparo, R. P. Frueholz, (1986), J. Appl. Phys., **59** (10), pp 3313/17.

/4/ L. L. Lewis and M. Feldman, (1981), Proc. 35th AFCS, pp. 612/24.
M. Hashimoto, M. Ohtsu, (1987), IEEE J. Quantum Electronics, **23**, pp446/51.
P. J. Chantry, B. R. McAvoy, J. M. Zomp, I. Liberman, (1992), Proc. IEEE International FCS, 114/22.
C. Szekely, R. E. Drullinger, F. L. Walls, J. P. Lowe, A. Novick, (1993), Proc. 7th EFTF, Neuchâtel, pp 593/97.
S. Kinugawa, T. Imamura, H. Suga, Y. Ohuchi, Y. Koga, Y. Saburi, (1993), Proc. International Symposium on Atomic Frequency Standards and Coherent Quantum Electronics, Nara, pp 81/82.

/5/ J. Vanier, C. Audoin, (1989), The Quantum Physics of Frequency Standards, Hilger.

/6/ W. Franzen, (1959), Phys. Rev., **115**, pp 850/56.
J. Vanier, J. F. Simard, J. S. Boulanger, (1974), Phys. Rev. A, **9**, pp 1031/40.
M. B. Gornyy, D. L. Markman, B. G. Matisov, (1985), Radiotecnica i elektronika, **9**, pp 1780/88.

/7/ W. J. Riley, (1981), Proc. 13th PTTI, pp 609/55.
See also /1/.

/8/ P. Thomann, G. Busca, (1981), J. de Physique, C8, Tome 42 , suppl. No 12, pp. C8-189/197

New Russian National Time and Frequency Standard and Time and Frequency Service in Russia

B.Al'shin, V.Borisochkin, B.Gaigerov, G.Yolkin, N.Koshelyaevsky, S.Pushkin, V.Tatarenkov, G.Aprishkin[†], V.Koutcherov[‡], Yu.Medvedev[*], N.Tel'pukhovski[**], V.Stepanov[***]

GP "VNIIFTRI", Mendeleevo, Moscow region, [†]NPO "DALSTANDARD", Khabarovsk, [‡]GOSSTANDARD, Moscow, [*]ATZUS, Golitzino, Moscow region, [**]VS NIIFTRI, Irkutsk, [***]NPO"SNIIM", Novosibirsk, Russia

SUMMARY

The certification tests of the whole National Time and Frequency Service (NTFS) were completed in 1994 in Russia. This report presents the results of it, some details of composition, performances and features of NTFS.

The national time and frequency standard laboratory plays a key role in NTFS. As a matter of fact, it establishes the reference frame for an atomic timing through out the whole territory of Russia. It independently reproduces the national unit of time, the second; generates and keeps the national time scale UTC(SU); and via the set of secondary state laboratories and technical means (on land and satellite based) ensures unification of time and frequency measurements. In addition the laboratory acts as certification center for the most accurate time and frequency standards, comparison instruments, time scale algorithms, etc.

The accuracy of the new national time and frequency standard, estimated as the ability to independently reproduce a unit of time, the second, basing on laboratory primary cesium frequency standards only in accordance with the SI definition was found to be $\leq 5\times10^{-14}$ [1].

The other important characteristic, the national time scale UTC(SU) stability, was also significantly improved and was certified as $\sigma_y(2, \tau) \leq 1\times10^{-14}$ for 1 day $\leq \tau \leq$ 100 days [2]. As a matter of fact, this characteristic is considerably better then stated. The most striking feature of the national time scale is that it based on a free running ensemble of super stable H-masers only.

Time links primarily based on "common view" GLONASS and GPS sessions, with leading laboratories inside Russia and abroad, provides frequency resolution at a level of about 1×10^{-15} and time uncertainty better than 10 ns [3].

The secondary state time laboratories were established to provide reliable and accurate time and frequency dissemination. Their main task is to generate UTC(k) which steers UTC(SU) and to distribute it via special radio and TV stations. The certification tests showed the ability of these laboratories to steer UTC(SU) at an uncertainty level better than a few tens ns and to keep an autonomous atomic time scale with a stability $\leq 10^{-14}$.

KEY WORDS: Time and Frequency Standard, Time Scale, Time Transfer, Time Dissemination.

1. NATIONAL TIME AND FREQUENCY SERVICE OF RUSSIA

The National Time and Frequency Service of Russia was founded about fifty years ago and nowadays operates according to Federal Law of Insurance of Measurement Unification under the supervision of the Committee of the Russian Federation on Standardization, Metrology and Certification (GOSSTANDARD). It consists of two main components.

The set of time laboratories, including the national time and frequency standard laboratory and secondary state laboratories widespread from the European part of the country to the Far Eastern region, Fig.1.

The set of radio and TV stations which are located in the vicinity of the above

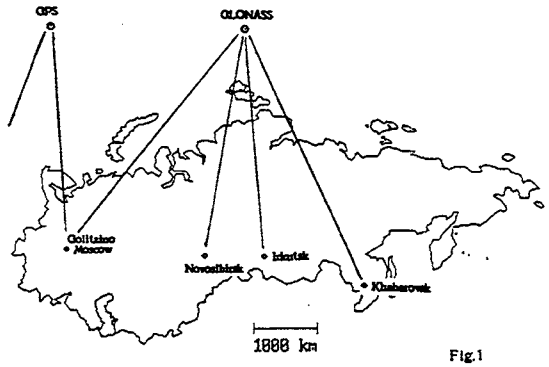


Fig.1

mentioned laboratories emit standard time and frequency signals under the control of these laboratories. The metrological characteristics of these stations and their schedules meet requirements of Circular B [4]. Apart from these stations all on land and space (GLONASS) navigation systems, which are controlled by the Department of Defense (DOD), emit signals which steer the time unit and time scale generated by the National Time and Frequency Standard.

The other significant component of NTFS activity according to Federal Law of Insurance of Measurement Unification is the determination of the Earth's rotation parameters - Universal Time and the Pole Coordinates. This aspect will not be discussed here.

2. THE NATIONAL TIME AND FREQUENCY STANDARD

2.1 Primary Cs Standards

The independent reproducing of the unit of time, the second, which is defined in the SI system using transition (F=3, m_f=0 ⇔ F=4, m_f=0) of rest and unperturbed Cs_{133} atom [5] is realized in laboratory by two standards; MCsR 102 and MCsR 103. Both instruments use a classical magnetic separation system, but are designed quite differently. Such design differences provide an opportunity to clarify and exclude various source of systematic uncertainties in the instruments.

During a period of about one year preceding final certification tests, both instruments were carefully studied. The result of this study as a budget of systematic errors is presented in Table 1.

The detailed analysis of the above mentioned effects one may found in [1].

The results of the comparison of the time unit reproduced by both laboratory standards, MCsR 102 and MCsR 103, relative to the time unit kept by the ensemble of H-masers, with stability $\leq 1\times10^{-14}$ for averaging time from 1 month to 1 year, are presented at Fig.2.

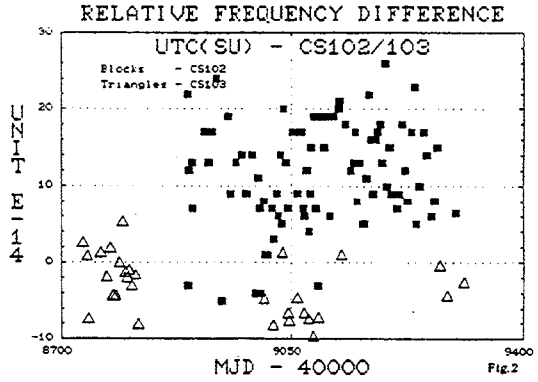


Fig.2

These results confirm in an indirect way the uncertainty estimation mentioned at Table 1. They show that the time unit duration reproduced by MCsR 102 standard coincidences with that of the TAI system (one must keep in mind that the time unit of the UTC(SU) system is less than that of TAI by about 1×10^{-13}), and show the existence of a systematic difference between the two Cs standards to be about 1.5×10^{-13}.

2.2 The ensemble of H-masers and national time scale stability

Perhaps the most striking feature of the National Time and Frequency Standard is utilization for more than twenty years of the ensemble of continuously operating H-masers for time scale generation.

Nowadays, this ensemble consists of 10 commercial H-masers, 6 instruments of CH1-80 type and 4 other instruments (new generation) from the complex CH0-110. All instruments operate under automatic cavity frequency control based on beam intensity modulation.

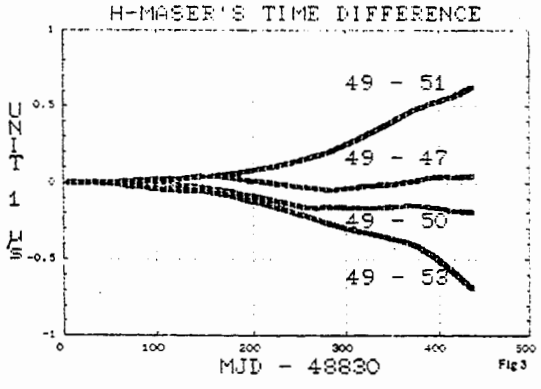
The typical results of H-masers intercomparison (1 ppS signals) is presented in Fig.3. The constant frequency and time differences between H-masers are removed. Moreover for the first two months, frequencies of H-masers are artificially

Table 1

The sources of systematic frequency errors, their values and uncertainties in MCsR102 and MCsR103

Unit x 10^{-14}

	Source	Value		Uncertainty	
		MCsR102	MCsR103	MCsR102	MCsR103
1.	C-field	~ 40 000			
1.1	C-field uniformity			0.3	2.0
1.2	$\overline{H}$ measurement uncertainty			1.3	2.0
1.3	H^2 replacement by $\overline{H}^2$			0.5	0.7
2.	Second order Doppler	27	23	2.0	2.0
3.	Phase shift	18	25	3.0	4.0
4.	Cavity pulling			1.0	0.5
5.	Rabi pulling			2.0	3.0
6.	Gravitational red shift	2.5	2.5	0.0	0.0
7.	RF spectrum			1.0	1.0
8.	Electronics			2.0	2.0
9.	"RF power" shift			–	6.0
		Total:		**5.0**	**8.9**

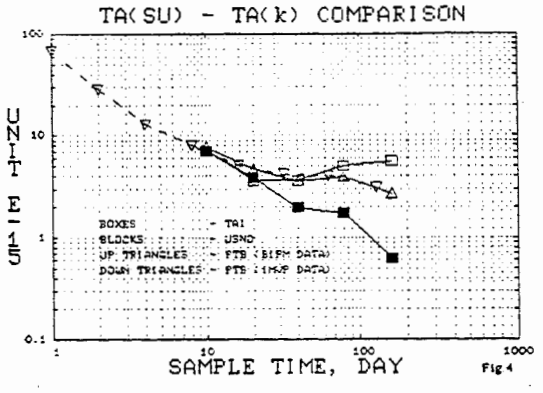


equalized (for demonstration purpose only).

The H-masers numbered 51 and fewer belong to the CH1-80 type. Those numbered 52 and more H-masers are of new generation. The presented data displays unprecedented stability of clocks; an uncertainty gained during one year doesn't exceed a few hundred nanoseconds! Stability estimation for pairs of H-masers based on modified Allan variance [6] gives values of about $3\text{-}4\times10^{-15}$ for averaging time from 10 to 100 days. This means that for TA(SU) and UTC(SU) one may expect at least two times better stability. Some H-masers have no detectable (at level about 3×10^{-17}/day) frequency drift. For the other instruments, it lies at a level $\sim 10^{-16}$/day, but is very stable, so there is no problem to take into account for time scale generation.

Since it's beginning TA(SU) and UTC(SU) have been free running time scales based on the ensemble of H-masers which have been deliberately shifted in frequency and time. The time unit is obtained by simply averaging of selected number of time units kept by these H-masers and their preliminary determined rates. The only difference between TA(SU) and UTC(SU) is a leap second which is introduced into the latter according to IERS circulars.

The next picture, Fig.4, represents TA(SU) stability plot relative to time scales of some leading laboratories in the world. If



one takes a more perfect reference he will immediately get a better result. Sometimes it even seems too difficult to correctly estimate stability level of time scale based on the above mentioned ensemble of H-masers.

Much more detailed information on the H-maser's ensemble of the national time and frequency standard of Russia may be found in [2, 7].

2.3 Time links characteristics

Since the early 1990's, satellite based links began to play the major role. This means

first of all GLONASS and secondly GPS. Nowadays, there are several GLONASS manually controlled navigation receivers of type A-724M in the national standard laboratory, which were slightly updated for time transfer needs. All other more or less significant time laboratories in Russia and the CIS are equipped with the same instruments. This provides the possibility to arrange a sort of overlapping sessions under the special GLONASS tracking schedule, which is elaborated and distributed by the DOD secondary time laboratory. Certification tests based on purely domestic data give time uncertainty 15-20 ns for averaging time one day and more and frequency resolution of at least $\leq 1\times10^{-14}$ for a one month averaging time [3]. The other results worth mentioning are based on the GLONASS/GPS time link comparison [8,9]. These give time uncertainties of about the same value as those mentioned above.

Starting in the middle 1991, a type TTR 6 GPS time transfer system was installed in our laboratory under special agreement with BIPM. Since than, the national time and frequency standard laboratory has had a very accurate time link with the international time-frequency community. Based on direct data exchange with laboratories of PTB and DLR, the uncertainty of this time link was estimated at 5 ns for a one day averaging time, and a frequency resolution better than 1×10^{-15} for averaging time of more than 1 month [3].

In addition to these time links, there are two others of much more limited validity. First is a TV time link within Moscow region, which is based on differential reception of signals transmitted by Ostankino TV station. The second is a micrometeorite (MM) two-way the TV time link is ≤ 10 ns and the frequency link at a distance up to two - two and a half thousand kilometers. The time resolution of resolution is $\leq 1\times10^{-14}$ for averaging time 10 days and more. The micrometeorite time link has a resolution of about ≤ 30 ns for one arm operation per session.

3. STATE SECONDARY TIME LABORATORIES: THEIR MAIN EQUIPMENT AND PERFORMANCES

There are at least four state secondary time laboratories in Russia (see map at Fig.4): ATZUS, Golitztino (G), Moscow region, DOD; VS NIIFTRI, Irkutsk (Im); NPO "DALSTANDARD", Khabarovsk (Kh); NPO "SNIIM", Novosibirsk (Nm).

The main tasks of these laboratories are to maintain the steering UTC(k) which follows UTC(SU) within predetermined uncertainties, and to keep as stable as possible an autonomous atomic time scale TA(k). In addition they have to disseminate time and standard frequencies via controlled special local radio and TV stations.

The main time keeping and time transfer equipment of these laboratories is depicted in Table 2.

Table 2
State secondary time laboratories. their equipment and performances

Secondary laboratory name and abbreviation	Clocks	Time transfer equipment	Certified UTC(k) tracking uncertainty	Certified TA(k) stability
ATZUS, G	12 H-masers	3 GLONASS TV	100 ns	$< 2\times10^{-14}$
VS NIIFTRI, Im	7 H-masers	2 GLONASS MM	100 ns	$< 2\times10^{-14}$
NPO "DALSTANDARD", Kh	8 H-masers	1 GLONASS MM	200 ns	$< 4\times10^{-14}$
NPO "SNIIM", Nm	5 H-masers	2 GLONASS MM	200 ns	$< 4\times10^{-14}$

Based on this equipment, laboratories succeed in achieving metrological performances reflected in the two following pictures; Fig.5 and 6. Noteworthy that the stability of an atomic time scales TA(k) and individual clocks in the majority of secondary laboratories is comparable to that of the

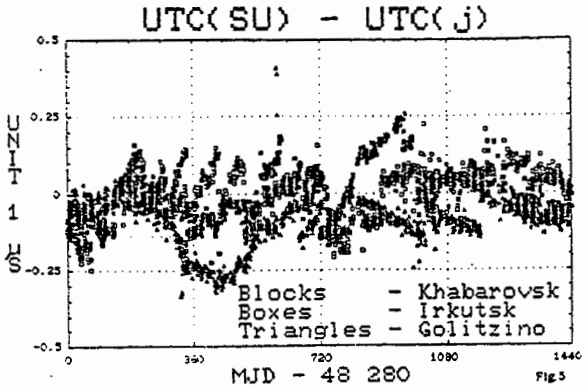


Fig.5

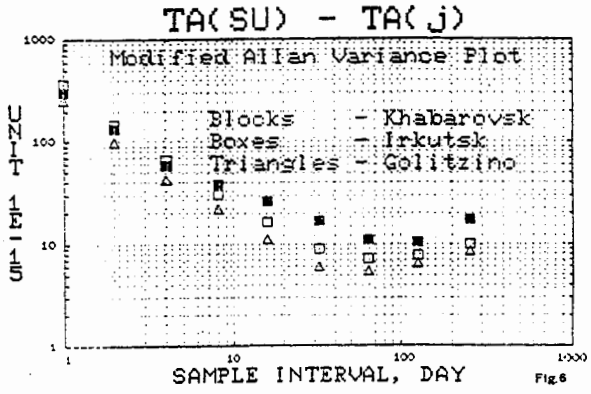


Fig.6

best laboratories contributing to TAI.

More detailed analysis of their activity may be found in our annual reports [10].

4. TRANSMISSION OF THE TIME AND STANDARD FREQUENCY SIGNALS

Transmission of time and standard frequency signals is the important part of the NTFS activity [11]. It is realized in direct way through two special LF stations: RBU and RTZ; two special HF stations: RWM and RID and through other radio and TV stations which emit such a signal as an auxiliary task. A typical frequency uncertainty of these signals are about 5×10^{-12} and time uncertainty from one to a few tens microseconds. Along with these there are many Loran-like and Omega-like Russian navigation stations which are controlled by DOD and operate in UTC(SU) time reference system and which consequently one may use for timing. The complete list of emitting stations and much more detailed information on their location, type and uncertainties of signals and schedule may be found in [4, 12].

Since 1990 satellite navigation system GLONASS began to play more and more significant role for time dissemination.

According to the GLONASS Interface Document forecast of UTC(SU) broadcasted in navigation message is accessible with an uncertainty of about 1 μs [13]. The laboratory of the national time and frequency standard in cooperation with secondary laboratories monitors GLONASS satellite's signals and distribute via special circulars the only legal information on GLONASS timing. Next picture, Fig. 7, shows the result of monitoring of GLONASS Time and UTC(SU) forecast basing on data from the laboratory of the national time and frequency standard.

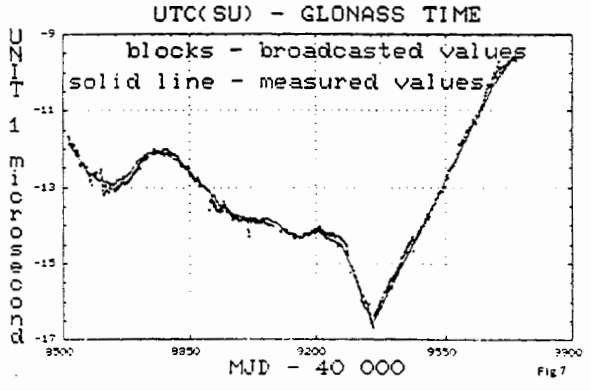


Fig.7

With removed obvious outliers this data shows stability of GLONASS Time of about 4×10^{-14} for averaging time 10 day and more and UTC(SU) forecast uncertainty of about 60 ns for one day averaging time.

Acknowledgement. The authors would like to thank all colleagues for co-operative work which leads to new accuracy and stability level of the National Time and Frequency Service of Russia. Our special indebtedness to BIPM for kind placing in our disposal TTR 6 GPS time transfer system.

References.

1. Г.Елкин, Ю.Яковлев, В.Барышев, В.Пуртов, Метрологические цезиевые реперы частоты МЦ 2 и МЦ 3: Опыт длительной эксплуатации в составе Госэталона, Труды 5-го Российского симпозиума "Метрология Времени и Пространства Э, 11 - 13 октября

1994, Менделеево, Россия.

2. А.Альшина, Е.Загирова, Н.Кошеляевский, Л.Мирошникова, С.Пушкин, A.Bauch, P.Hetzel, J.Hahn, H.Nau, Исследование долговременной стабильности частоты национальной шкалы времени России и отдельных хранителей эталона по результатам международных сличений, Труды 5-го Российского Симпозиума "Метрология Времени и Пространства", 11-13 октября 1994, Менделеево, Россия.
3. А.Альшина, Н.Зорина, Н.Кошеляевский, С.Пушкин, В.Титов, Ю.Медведев, Н.Тельпуховский, A.Bauch, P.Hetzel, P.Daly, W.Lewandowski, C.Thomas, J.Hahn, H.Nau, Оценка точностных характеристик оперативных каналов сличения Национального эталона времени и частоты России, Труды 5-го Российского Симпозиума "Метрология Времени и Пространства", 11-13 октября 1994, Менделеево, Россия.
4. БЮЛЛЕТЕНЬ В $^{09}_{1991}$ Государственная комиссия единого времени и эталонных частот СССР, Эталонные сигналы времени и частоты. Характеристики и программы передач через радиостанции, телевидение и сеть звукового вещания, Москва 1991.
5. Le Systeme International d'Unites (SI), 6-e Edition, Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92310 Sevres, France, 1991.
6. Measures for Random Instabilities in Frequency and Time (Phase), International Telecommunication Union, Radiocommunication Study Groups, Question ITU-R 104/7.
7. A.Al'shina, B.Gaigerov, N.Koshelyaevsky, S.Pushkin, Russian National Time Scale Longterm Stability, Proceedings of the Twenty Fifth Annual Precise Time and Time Interval Application and Planning Meeting, 29 November - 2 December 1993, Marina Del Ray, California, USA.
8. Отчет ОКР"Создание объектов и средств Государственной системы единого времени и эталонных частот "ЦЕЛЬ" (второй этап) в части, возложенной на Госстандарт", НПО "ВНИИФТРИ", Москва, 1994.
9. P.Daly, N.Koshelyaevsky, W.Lewandowski, G.Petit and C.Thomas, Metrologia, **30**, 89-94, 1993.
10. Отчет о постоянной метрологической работе 07.06.17.01 "Государственная служба времени и частоты СССР (России)", Раздел I, "Хранение, поддержание и совершенствование Государственног эталона времени и частоты", НПО "ВНИИФТРИ", Москва, 199x.
11. В.Борисочкин, В.Кучеров, В.Татаренков, Г.Черенков, Система передач эталонных сигналов частоты и времени России, Труды 5-го Российского Симпозиума "Метрология Времени и Пространства", 11-13 октября 1994, Менделеево, Россия.
12. Annual Report of the BIPM Time Section, Volume x, 199x, Bureau International des Poids et Mesures, Pavillon de Breteuil, F-92132 SÈVRES, France.
13. Интерфейсный Контрольный Документ "Глобальная Навигационная Спутниковая Система ГЛОНАСС"

THE USE OF HYDROGEN MASERS IN TAI COMPUTATION

J. Azoubib and C. Thomas

Bureau International des Poids et Mesures
Pavillon de Breteuil, 92312 Sèvres Cedex, France

Summary

Formerly, the generation of TAI was almost entirely based on data from commercial caesium clocks and from primary frequency standards. The hydrogen masers maintained in timing laboratories were used for local studies in which excellent short-term stability was required. Over the last few years, the development of auto-tuning modes for the cavities has led to improved middle-term and long-term stability of active hydrogen masers, through a large reduction in their intrinsic drifts.
Since July 1991, data from hydrogen masers received regularly at the BIPM have been introduced into the TAI computation, with a view of improving the stability and reliability of the resulting time scale. After some fluctuations, their global relative contribution, expressed in percent of the total weight, reached a nearly constant value of about 17% in 1993. The effect of introducing hydrogen masers has been investigated at the BIPM by comparison of the free time scale EAL, first stage of the TAI computation, with another free time scale, ESMH, obtained with the same algorithm and clock data, but excluding the data from the hydrogen masers. It appears that fluctuations in the global relative contribution of hydrogen masers in EAL led to some local instabilities of EAL. However, for the full period under study, the time scale ESMH is globally as stable as EAL and would have led, after frequency steering corrections, to a fictitious atomic time TESMH as accurate as TAI. These results can easily be explained: just one hydrogen maser experienced a large frequency drift relative to EAL, and the particular weighting procedure in use gives a low weight to any drifting clock.
If the number of hydrogen maser data reported to the BIPM happens to increase and the stability of EAL to improve, it may become necessary to refine the procedures for frequency prediction and weight determination used in the algorithm for hydrogen masers. A first test on real data, which introduces an estimated drift for each individual hydrogen maser, has already been made at the BIPM. This shows that it is a demanding task to make a good estimation of the frequency drift of an hydrogen maser, and one which may require a long period of observation, but such estimates would help to improve the reliability of the time scale and reduce its residual frequency drift.

1. Introduction

The first step in the establishment of TAI is the computation of a free atomic time scale, EAL (échelle atomique libre), obtained as a weighted average of about 200 free-running atomic clocks spread world-wide, using the algorithm ALGOS which is optimized for long-term stability [1]. To maintain its accuracy, TAI is derived from EAL by frequency steering. The steering corrections, determined by comparing the EAL frequency with primary frequency standards, are of the same order of magnitude as the EAL fluctuations. For example, occasional fractional frequency corrections of $0{,}5 \times 10^{-14}$ have been required in recent years.

An iterative, post-processing procedure allows the computation of EAL. This treats, as a whole, measurements taken over a basic sample period of $T = 60$ days, although measurements are available more frequently. A definitive computation of the deferred-time time scale EAL is thus made available six times a year. By averaging frequency values over 60 days, both the noise of synchronization links and the white frequency modulation of the clocks themselves are smoothed out. With this integration time, the scale operates in the region corresponding to the random walk frequency modulation of caesium clocks. For this reason stability is not improved by averages of longer duration.

1.1. Basic equations of EAL computation

Basic data for the calculations are the time differences between clocks, measured at ten-day intervals on the MJD ending with 9. The time scale EAL is updated for each of the seven dates t included in the 60 day period under consideration $t = t_0 + nT/6$, $n = 0, 1, \ldots 6$ where t_0 is the last date of the previous two-month interval. At date t, EAL is defined as:

$$EAL(t) = \sum_{i=1}^{N} p_i \left[h_i(t) + h'_i(t) \right] \Big/ \sum_{i=1}^{N} p_i , \qquad (1)$$

where $h_i(t)$ represents the reading of clock H_i at date t, N is the number of clocks, p_i is the weight attributed to clock H_i and $h'_i(t)$ is a time correction applied at date t to ensure time and frequency continuity of the scale when the weights of individual clocks are changed and/or when the total number of clocks is changed [2].

Suppose the time difference $x_i(t)$ between each clock H_i and EAL, at date t, is written as:

$$x_i(t) = EAL(t) - h_i(t), \qquad (2)$$

and the measured time difference at date t between clocks H_i and H_j, is written as:

$$x_{ij}(t) = h_j(t) - h_i(t). \qquad (3)$$

Equations (1), (2) and (3) lead to a deterministic system with N equations and N unknowns for each date t included in a given 60 day period. The solution is unique and the results are the N time differences $x_i(t)$ which give access to EAL from each individual clock H_i at a given date t. For the whole 60 day interval, this procedure gives rise to the quantities $x_i(t_0 + nT/6)$, $n = 0, \ldots, 6$ for each clock. The least squares slope of these quantities is referred to as the frequency $y_i(t_0 + T)$ of the clock H_i, relative to EAL, for the 60 day interval $[t_0, t_0 + T]$.

The correction term $h'_i(t)$ for clock H_i is the sum of two terms, with

$$h'_i(t) = x_i(t_0) + y_{ip}(t).(t - t_0), \qquad (4)$$

where $y_{ip}(t)$ is the frequency of clock H_i, relative to EAL, predicted for the period $[t_0, t]$.

The practical computation is possible when a weight and a predicted frequency, relative to EAL, are attributed to each clock contributing to the generation of EAL for the 60 day interval $[t_0, t_0 + T]$.

The predicted frequency $y_{ip}(t)$ for date t included in the interval $[t_0, t_0 + T]$ is kept constant for the whole 60 day interval $[t_0, t_0 + T]$. It is chosen to be the frequency $y_i(t_0)$ computed for the previous 60 day interval $[t_0 - T, t_0]$:

$$y_{ip}(t) = y_i(t_0). \tag{5}$$

The assumption that clock H_i is most likely to behave over the coming 60 day interval as it did on the previous one is linked to the choice of the sample duration T: for such an averaging time the predominant clock noise is random walk of frequency, and the optimal estimate for the future is the last value computed over a sample of duration T.

The weight p_i attributed to clock H_i is chosen to optimize the long-term stability of the resulting time scale and to minimize the annual fluctuation of its apparent frequency with respect to primary frequency standards. In addition, the weight determination takes into account the behaviour of the clock in the current 60 day period. An intentional null weight is attributed to a new clock for its two first 60 day intervals of operation, before being effectively introduced in the weighted average. Then the weight of clock H_i is computed using a temporary value p_i' given by:

$$p_i' = \frac{10000}{\sigma_i^2(6,T)}, \tag{6}$$

where individual variances $\sigma_i^2(6,T)$, conventionally expressed in $(\text{ns/d})^2$, are computed for the frequency $y_i(t_0 + T)$ and the five previous 60 day frequencies, the factor 10000 being arbitrarily chosen. The weight p_i attributed to clock H_i is equal to p_i' except in two cases:

* If $\sigma_i(6,T) \leq 3{,}16$ ns/day, then p_i is set to the value 1000 which is the upper limit of weights (the clock weights published in the successive volumes of the *Annual Report of the BIPM Time Section* are divided by a factor 10, so the upper limit of weight appears to be 100).
* If clock H_i shows an abnormal frequency step, then p_i is set to 0. The objective criterion set to safeguard EAL against possible abnormal behaviour of some clocks is detailed in [1].

Clocks contributing to TAI are supposed to be free-running and uncorrelated. The mode of frequency prediction and the weighting procedure used in the TAI computation are identical for any clock contributing to the ensemble, whatever its manufacturer, age, or even type.

1.2. Composition of the clock ensemble used for TAI computation

In the past, the generation of TAI was entirely based on data from commercial caesium clocks and primary frequency standards. The hydrogen masers maintained in timing laboratories were used for local studies requiring excellent short-term stability, but insensitive to long-term drift. Over the last decade, the development of auto-tuning modes for the cavities has led to a large reduction in the drift of active auto-tuned hydrogen masers, with consequent improvement in their middle-term and long-term stability. Data from hydrogen masers has thus been progressively introduced into the clock ensemble with the aims of improving the stability of EAL and, since the number of participating clocks is increasing, its reliability. It was considered at the time that the residual frequency drifts of contributing hydrogen masers might cause frequency drift of EAL and thus threaten the accuracy of TAI: this problem, however, could be overcome outside the main algorithm by the use of frequency steering corrections.

The first hydrogen masers were introduced into the EAL computation for the interval September-October 1984. All 3 of the hydrogen masers involved very quickly reached the upper limit of weight, as shown in Fig. 1.

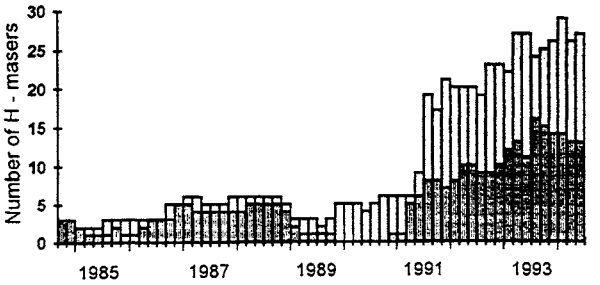


Figure 1: Histogram showing the number of hydrogen masers reaching the maximum weight compared to the total number of hydrogen masers involved in EAL computation from September-October 1984 to May-June 1994.

The total number of hydrogen masers remained at about 5, with nearly none of them reaching the maximum weight in 1990, until July 1991 when the complete set of data from hydrogen masers received regularly at the BIPM was introduced into the computation. Since then, their total number has increased very slowly, reaching 27 at mid-1994, with nearly half of them at the upper weight. For sake of comparison, the number of the newly designed commercial HP 5071A clocks increased from several units at the beginning of 1992 to 74 units at mid-1994.

Figure 2 shows the variation with time of the total relative contribution of hydrogen masers in the EAL computation.

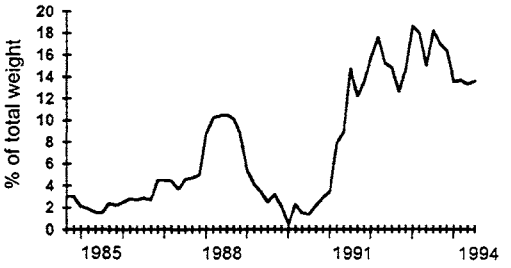


Figure 2: Relative contribution, expressed in percent of the total weight, of hydrogen masers in EAL computation from September-October 1984 to May-June 1994.

The relative contribution of hydrogen masers remained below 3% until October 1986, then increased slowly to 5% in the interval November-December 1987 as first Russian hydrogen masers were introduced. At the beginning of 1988, the value of the upper limit of weight was increased by a factor 5, leading to the present value of 1000. Consequently the relative contribution of hydrogen masers first increased, reaching 10% during several consecutive intervals of 1988, then decreased to a few percent in 1989 and 1990. Since July 1991, their global relative contribution has increased reaching a mean value of 17% in 1993.

The effect of the introduction of hydrogen masers into the TAI computation is investigated in Section 2 of this paper. In Section 3 a refined procedure for treating hydrogen maser data is proposed, and tested on real data.

2. Effect of hydrogen maser data on TAI computation

The effect of introducing hydrogen masers into the TAI computation is first investigated by comparison of the free time scale EAL with another free time scale, ESMH (Echelle Sans Maser à Hydrogène), obtained with the same algorithm ALGOS and the same clock data, but excluding the data from the hydrogen masers. An interpretation of the results is given in a second section where the statistical properties of the hydrogen maser data reported to the BIPM are studied.

2.1. Comparison between EAL and ESMH

The free atomic time scale ESMH relies upon data collected by the BIPM from commercial caesium clocks and primary frequency standards. Although the main entry of hydrogen maser data occurred in July 1991, ESMH has been generated through the algorithm ALGOS from January 1988, date of the introduction of the weighting procedure detailed in Section 1.1, that increased the relative contribution of hydrogen masers in EAL.

The stability of EAL and ESMH are characterized by curves showing the variation with averaging time τ of the Allan standard deviation $\sigma_y(\tau)$ given in Fig. 3. This results from application of the N-cornered hat technique to data obtained in comparisons between EAL, or ESMH, and the best independent time scales (maintained at the NIST, the PTB, the VNIIFTRI, the USNO and the LPTF) for the period January 1988 to June 1994.

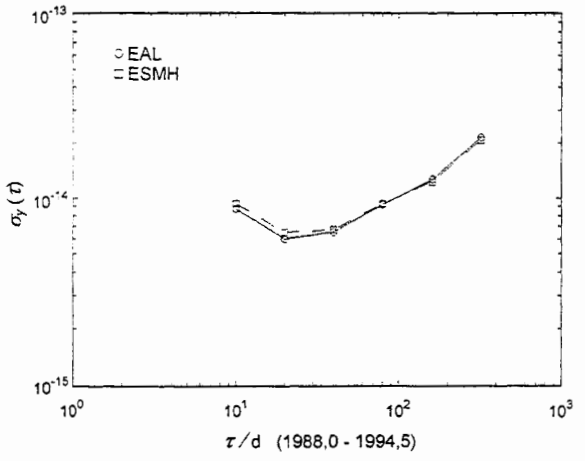


Figure 3: Stability of EAL and ESMH.

Figure 3. shows that the stability of EAL and ESMH is comparable for all averaging times, especially in the long-term where a frequency drift is visible for both scales. The value of this frequency drift, estimated through a quadratic fit applied to the time differences, collected from January 1988 to June 1994, between EAL and the primary frequency standard PTB CS2 is $-2{,}2\times10^{-17}$ d^{-1}.

Using the same procedure, illustrated in Fig. 4, the frequency drift difference between EAL and ESMH is about -3×10^{-18} d^{-1}, which indicates that ESMH has a slightly smaller drift, in absolute value, than EAL relative to PTB CS2.

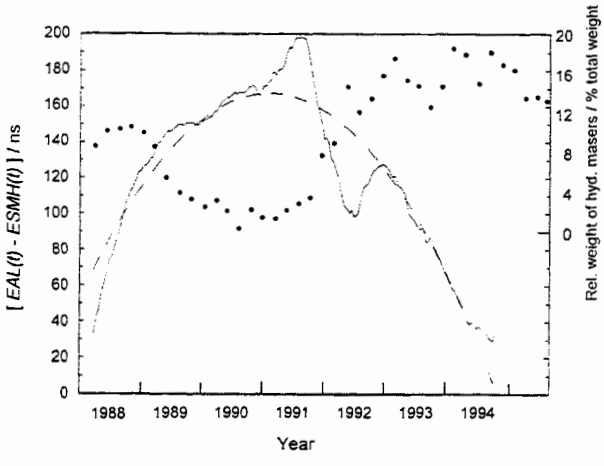


Figure 4: Determination of the frequency drift difference between EAL and ESMH.

An interesting feature of Fig. 4 is the fluctuation of the time difference between EAL and ESMH observed in 1991 and 1992, a period in which the total relative contribution of hydrogen masers to EAL increased from 2% to 16%. This important addition of hydrogen masers probably caused a local instability in EAL, which disappeared in the following year.

Another important point is the accuracy of the TAI scale interval. As noted in the introduction, TAI is derived from EAL by frequency steering. Thirteen such corrections, most of them of amplitude $0{,}5\times10^{-14}$, were applied during the period of computation of the time scale ESMH [3]. One can imagine another time scale, TESMH, deduced from the free time scale ESMH with the same frequency steering corrections as those applied to EAL for obtaining TAI. The departures, d and d', of the scale units of TAI and TESMH from the SI second, as produced by the primary frequency standard PTB CS2 on the rotating geoid, are shown in Fig. 5 for the period 1988-1994. One can associate with each point an error bar which corresponds to the uncertainty (1 σ) of PTB CS2, which is estimated to be $1{,}5\times10^{-14}$. The changes in the values of d and d', due to frequency steering, are also shown in Fig. 5.: they take the form of arrows, each corresponding to a change of $0{,}5\times10^{-14}$ s, in absolute value, or $0{,}75\times10^{-14}$ s for one of them.

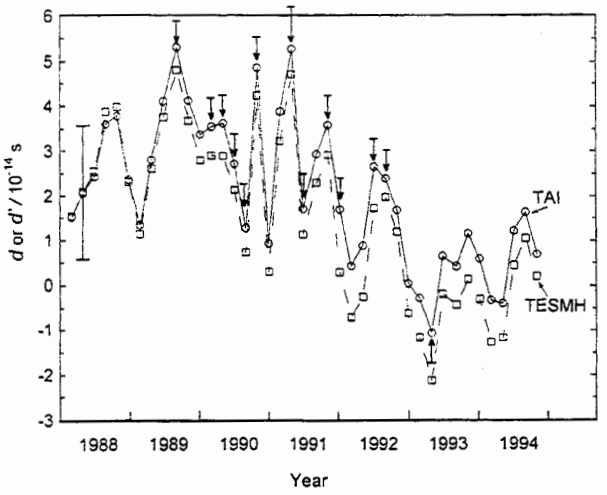


Figure 5: Departure d and d' of the scale units of TAI and TESMH from the SI second as produced by PTB CS2 on the rotating geoid.

It follows from this graph that the decisions, made in real-time, about whether or not to apply a frequency correction would be likely similar if based on ESMH rather than EAL. This is inevitable given the very small frequency drift between EAL and ESMH. We can nevertheless note that TESMH would have been slightly more accurate than TAI for the year 1994.

To conclude, the introduction of hydrogen maser data into the TAI computation did not bring an important modification of the quality of the resulting time scale. The interpretation of this phenomenon requires statistical analysis of the individual hydrogen maser data received at the BIPM.

2.2. Statistical analysis of hydrogen maser data

The data treated here is composed of series of 10 day time differences, $x_i(t)$, between EAL and the readings of the 10 hydrogen masers which operated continuously for at least two years in the period 1991-1994. These 10 hydrogen masers are maintained in 5 different timing centres and, in the following, are denoted HM_i (i = 1, ... 10). The corresponding log-log curves of the Allan standard deviation $\sigma_{yi}(\tau)$ make it possible to pick out possible frequency drifts D_{iEAL}, characterized by a straight line of slope +1, and, where this is detected, to determine its approximate amplitude from the equation:

$$|D_{iEAL}| = \sqrt{2}\sigma_y(\tau)/\tau\,, \qquad (7)$$

valid in the range of τ values corresponding to the slope +1. When a frequency drift exists, a parabolic pattern is observable in the graph of the $x_i(t)$ versus t. The sign of D_{iEAL} is then deduced from the sense of the concavity of the parabola.

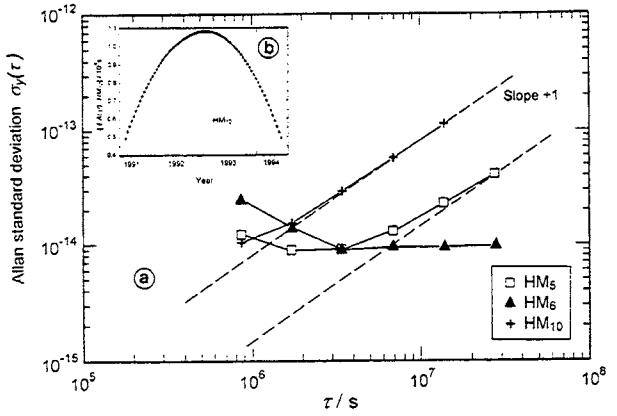


Figure 6: Determination of a possible frequency drift relative to EAL of the hydrogen masers HM_5, HM_6 and HM_{10}:
6.a. $\sigma_{yi}(\tau)$ curves,
6.b. concavity of the $[EAL(t) - HM_{10}(t)]$ curve.

The log-log graphs of $\sigma_{yi}(\tau)$ versus τ are given for several hydrogen masers in Fig. 6.a. The characteristic straight line of slope +1 is observable for one of them, HM_{10}: in the other cases no frequency drift is observed. An upper limit to this drift can be estimated by considering that the value $\sigma_{yi}(\tau)$ corresponding to the largest value of τ belongs to such a straight line of slope +1, as shown for HM5. The results of this study are given in Table 1.

HM_i	D_{iEAL} detected	$\lvert D_{iEAL} \rvert$ / d^{-1}	Sign of D_{iEAL}
HM_1	No	$\leq 1{,}4\times10^{-16}$	+
HM_2	No	$\leq 1{,}4\times10^{-16}$	+
HM_3	$80 \leq \tau$	7×10^{-16}	-
HM_4	No	$\leq 1{,}7\times10^{-16}$	+
HM_5	$160 \leq \tau$	$2{,}6\times10^{-16}$	+
HM_6	No	*	*
HM_7	No	*	*
HM_8	No	$\leq 1{,}7\times10^{-16}$	+
HM_9	No	$\leq 0{,}9\times10^{-16}$	-
HM_{10}	$40 \leq \tau$	$9{,}4\times10^{-16}$	-

Table 1: Estimation of the amplitude of the frequency drift D_{iEAL} of hydrogen maser HM_i relative to EAL deduced from the analysis of the $\sigma_{yi}(\tau)$ curves (Fig. 6.a.) obtained with the time differences $x_i(t)$ between EAL and the 10 hydrogen masers with at least 2 years of continuous operation in the period 1991-1994. The sign of D_{iEAL} is determined from the sense of the concavity of the curves $x_i(t)$ (Fig. 6.b.).
** means that no information on D_{iEAL} may be obtained from the $\sigma_{yi}(\tau)$ curves.*

It follows that only three hydrogen masers, HM_3, HM_5 and HM_{10} had a detectable frequency drift relative to EAL in the four years under study. For the basic computation interval of TAI, T = 60 d, the predominant frequency noise of HM_3 and HM_5 is random walk of frequency, like most caesium clocks, with a sufficient low level that their contribution to EAL was nearly always the maximum authorized. Hydrogen maser HM_{10} is already in a regime of frequency drift at τ = 60 d, and could thus have brought a non-negligible frequency drift to EAL, of order $-1{,}4\times10^{-17}$ d^{-1} (to be compared to the value $-2{,}2\times10^{-17}$ d^{-1} given in 2.1), if weighted at the upper relative weight of about 1,5%. Fortunately, the weight of HM_{10} in EAL computation was always very low in the period under study. This is so because the calculation of the weights, described in (6), uses the variance $\sigma_i^2(6,T)$ over six consecutive intervals, each of duration T. One can show that if a given hydrogen maser HM_i presents a pure frequency drift D_{iEAL} for an averaging time of duration T, the quantity $\sigma_i(6,T)$ is larger than the limiting value, 3,16 ns/day, as soon as $|D_{iEAL}|$ is larger than $3{,}26\times10^{-16}$ d^{-1}, thus preventing HM_i from receiving the maximum weight. The weighting procedure in ALGOS thus deweights clocks presenting large frequency drifts relative to EAL. The introduction of the hydrogen maser data into the EAL clock ensemble was thus carried out safely.

While we can feel secure about the introduction of hydrogen masers into the EAL clock ensemble in the period 1991-1994, it is important to emphasize several different points about the statistical analysis of the data they provide:

* The preceding study does not separate the part of the frequency drift coming from the time scale itself and from that due to a given hydrogen maser. To complete this point, the N-cornered hat technique can be applied to the values $x_i(t)$ in order to obtain absolute values of the Allan standard deviation of each hydrogen maser. This requires some caution: the involved elements should be statistically independent, and should correspond to the same period of continuous operation. A test has been done with an ensemble of four elements: 3 hydrogen masers maintained in 3 different timing centres, HM_3, HM_8, and HM_{10}, and EAL. The results obtained for the 3 hydrogen masers are shown in Fig. 7., and yield drift estimates similar to those obtained in the preceding section.

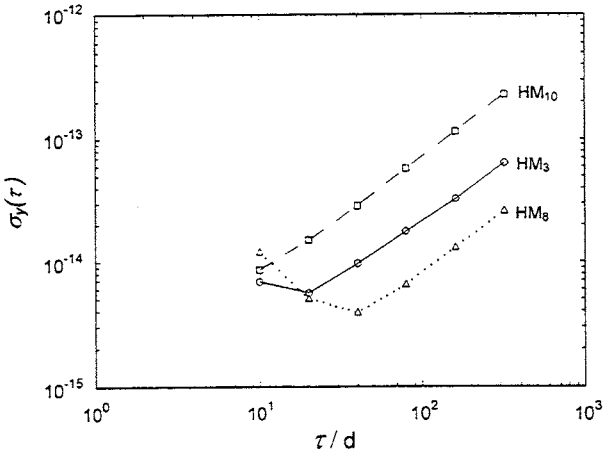


Figure 7: Allan standard deviations of HM_3, HM_8 and HM_{10}.

* Data showing the variation with the averaging time τ of the Allan standard deviation is full of information, as it indicates the regime of frequency fluctuations that the hydrogen maser experiences for the duration T of the basic interval of computation of the time scale. This knowledge is fundamental for optimum frequency prediction, and so, essential for improving the stability of the time scale. If an estimation of frequency drift is useful for this purpose, several procedures are available. Here we present an estimation based directly on the values of the Allan standard deviation: one can also treat the frequencies estimated over the period T by a linear fit, or an average of first differences.
* To estimate the frequency drift of a given hydrogen maser it is necessary to observe it for a long period. This was the case in the examples described in Section 2 where the effect of introducing hydrogen masers is examined after the event. For a time scale computation required in near-real time, the duration of the available data may not be sufficient. The safe use of hydrogen maser data then requires a long period of observation before introduction, certainly longer than the two 60 day intervals used in the EAL computation for caesium clocks.

Given these considerations, and even although the use of hydrogen maser data did not threaten EAL stability in past years, we have designed and tested a new version of the algorithm ALGOS which takes into account possible frequency drifts. This is detailed in Section 3.

3. Better use of hydrogen masers in TAI computation

The basis hypothesis of the present ALGOS is that caesium clock and hydrogen maser data may be treated by similar procedures. A new version of the algorithm ALGOS has been developed at the BIPM which applies a different mode of frequency prediction and a different observation period before introduction to different clock types:

* Hydrogen masers are observed for 6 consecutive 60 day intervals while caesium clocks are observed for two such intervals.
* The predicted frequency $y_{ip}(t)$ for date t included in the interval $[t_0, t_0 + T]$ is chosen to be the frequency $y_i(t_0)$, computed for the previous 60 day interval $[t_0 - T, t_0]$, modified by the addition of a shift of frequency deduced from the estimated drift of the hydrogen maser. This is written as:

$$y_{ip}(t) = y_i(t_0) + D_i(t) \times T, \qquad (8)$$

where $D_i(t)$ is the frequency drift of hydrogen maser HM_i, relative to EAL, estimated over a 1 year period ending at date t. Here the quantity $D_i(t)$ estimated over the past year of operation is distinguished from the quantity $D_{iEAL}(t)$ of Section 2 estimated after the event by observation of the whole life of the hydrogen maser.

This modified algorithm was tested on real clock data for the period January 1988-June 1994. To do this, the quantity $D_i(t)$ was estimated, for all participating hydrogen masers, by computation of the average of the first differences of the 6 consecutive 60 day frequencies observed during the year ending at date t. The stability of the resulting time scale E, obtained by application of the N-cornered hat technique as in Section 2.1. for ESMH, is very close to that of EAL, as shown in Fig. 8. This is inevitable since, as we have already shown, the introduction of hydrogen maser data with no special restriction on its use did not degrade the stability of EAL.

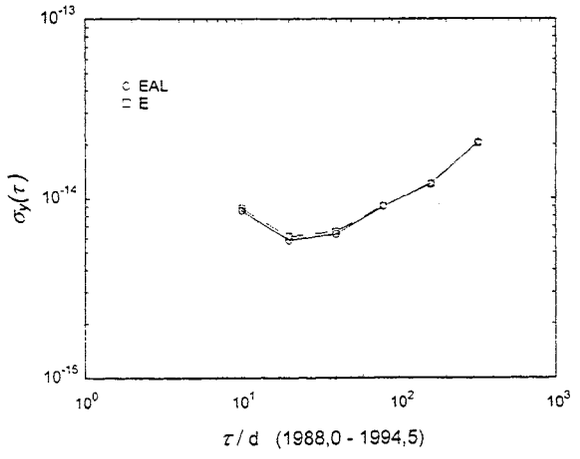


Figure 8: Stability of EAL and E.

It is possible to compare, after the event, the predicted frequency $y_{ip}(t)$ with the actual one $y_i(t)$. It follows that (8) associated with the estimation of D_i over 1 year is optimum if HM_i is in a pure regime of frequency drift or in a pure regime of random walk frequency modulation (this latter case yields a very small value of D_i so that (8) is reduced to (5) thus giving the optimum prediction). If the fluctuations of HM_i frequency are driven by both a drift and a random walk noise, the drift is not estimated precisely and fluctuates considerably from date t to date $t + T$, sometimes yielding bad frequency predictions. It may thus be necessary to make a better estimate of the drift, for example by using a much longer period of observation before introducing the maser to the ensemble. Another step in the design of the algorithm would be to refine the procedure for weighting hydrogen masers. Indeed, if the frequency drift of a hydrogen maser is determined safely and is used in predicting its frequency, the optimum use of the resulting data calls for it to be attributed a high weight. This is not the case if the weight is deduced it from (6).

4. Conclusions

Study of the impact of the introduction of hydrogen maser data to the TAI computation in the period 1988-1994 shows that it did not degrade the global stability of EAL. If this data had not been used, the fictitious atomic time obtained after frequency steering identical to that applied to EAL would have been as accurate as TAI. Statistical analysis of the behaviour of the hydrogen masers reported to the BIPM shows that a single unit experienced a large frequency drift, relative to EAL, during the period under study. The algorithm correctly handled this data through its weighting procedure, which assigns a low weight to any drifting clock.

The quality of timing data used to generate EAL is rapidly evolving thanks to the wide use of accurate time transfer methods and very stable commercial caesium clocks. This will greatly help to improve the stability of EAL. The frequency drift relative to EAL of hydrogen masers may then become the predominant trend over averaging times close to the basic computation time of EAL, but should not be treated by a simple deweighting. The introduction of a specific weighting procedure and mode of frequency prediction for hydrogen masers is already under test at the BIPM. To be effective, these require good estimates of frequency drifts, but this may be possible if a sufficiently long period of observation is used.

References

1. Tavella P. and Thomas C., Comparative Study of Time Scale Algorithms, *Metrologia*, 1991, **28**, 57-63.

2. Thomas C., Wolf P. and Tavella P., Time Scales, *Monographie du BIPM*, 1994, 94/1, 52 pages.

3. Table 5 of the *Annual Report of the BIPM Time Section*, 1993, **6**, p 35.

SESSION 13

RESONATORS III

Chairman : R.J. BESSON

LANGASITE PLANO-CONVEXE RESONATORS USING LATERAL OR PERPENDICULAR FIELD EXCITATION.

J.Detaint*, B.Capelle, A.Zarka**, D.Cochet-Muchy***.**

* FRANCE TELECOM CNET/PAB 92225 Bagneux France.
** L.M.C.P. Université Paris VI 75252 Paris France.
***CRISMATEC Z.I. Mayencins 38610 Gières France.

Abstract:

A study of plano-convexe resonators using Y cut Langasite resonators excited by a conventional thickness field or by an in-plane (lateral) field was made to obtain devices with very high Q factors. In a first part of the paper, the propertie of the thickness shear modes of the Y rotated cuts of LGS and the energy trapping properties of these modes are given. In a second part the experimental results obtained for a variety of designs using the thickness excitation or the lateral excitation are presented. Very high Q factors were measured for several overtones of the slow shear temperature compensated (C) mode [Q.f>.75 10^{13}], and for the quasi-longitudinal (A) mode [Q.f#1.8 10^{13}]. The largest values of the Q factor were observed in each case for the 9th overtone. The fact that much higher Q factors were measured for the antiresonance of the same modes is an indication of the possibility of still larger values at the resonance frequencies. A comparison of the observed and computed mode shape was made with a good agrement for the first overtones, and some differences for the 7th and the 9th ones. The same conclusions were also drawn of a comparision of the experimental and computed resonance frequencies and values of the elements of the equivalent scheme. A calculation of the relative influence of acoustic and dielectric losses on the obtained Q factors has indicated that Y cut LGS has extremely reduced dielectric losses and has shown a important difference between quartz and LGS. We conclude to the large interest of this material for low phase noise frequency generation.

1. INTRODUCTION:

The evolution of the electronic systems and almost of the telecommunication systems towards higher frequencies and baud rates has led to the need to find new solutions for the frequency generation and filtering devices. Particularly, the achievement of the extremely large information transfer rates possible on optical fibres (information highways) will require the use of frequency sources having much improved short term stabilities whereas the development of high baud rates radiocommunication systems will demand filters with much greater bandwidthes.

The present solutions in these domains rely most often, on quartz devices, but, although these are gaining years after years in frequency stability; it can exist better possibilities in using new materials with lower intrinsic acoustic losses and higher coupling coefficients [1].

The evolutions of the applications conduct also to favour the materials that allow the obtainment of very high frequency devices by some intrinsic property or some particular ability to withstand technological operations leading to higher operating frequencies [2]. Another evolution of the electronics equipments lead to research an important reduction of the dimensions of the devices.

In the recent years several new, thermally compensated piezoelectric materials permitting to obtain filters with larger bandwidths and oscillators with larger shifts or larger frequency stability (mostly short term) than quartz were developed. It was successively observed that berlinite ($AlPO_4$), lithium tetraborate ($Li_2B_4O_7$), gallium phosphate ($GaPO_4$), and langasite ($La_3Ga_5SiO_{14}$), can fulfil these conditions while equalling or surpassing quartz for other very important properties such as the crystalline perfection, the acoustic losses or the propagation velocities. Most of these new materials belong to the crystal class of quartz. Two of them are close structural analogues of quartz.

The lanthanum silico-gallate ($La_3Si_3GaO_{14}$, langasite or LGS), and several of its modifications (substitution of one or several of the elements of the formula by other elements) or structural analogs were studied in Russia since the beginning of the eighties [3][4]. Recently, several important advances were reported: the growth of crystals having large dimensions [5][6][7][8], the measurement of several sets of constants, of several physical properties [9][10][11] and the first systematic studies of devices [8][11][12][13]. Langasite and its analogs belong to the symmetry class of quartz.

In this paper we report investigations made to further precise the possibilities of using langasite to obtain the extremely high Q factors required to enhance significantly the phase noise characteristics of the crystal oscillators.

I PIEZOELECTRIC PROPERTIES OF LANGASITE.

Langasite is a material which have very attractive piezoelectric properties (more intense piezoelectric properties than quartz, temperature compensation for several cuts, low acoustic losses, etc..). It possesses also very interesting properties in optics [laser crystal (Nd...doped), photorefractive (Pr)], and for dielectric applications [extremely low dielectric losses with a zero temperature coefficient of the permitivity in several directions].

I.1 Basic properties of the Y rotated cuts of LGS:

Four quantities which determinate the most important properties of the thickness shear resonators or filters will be recalled [12] in this paragraph: the phase velocities (or the related frequency constants NFa=V/2), the coupling coefficients, the temperature coefficients of the resonance frequencies and the acoustic attenuations. The three first were computed using a one dimensional model of the piezoelectric resonators [14] with the constants of reference [10]. Acoustic attenuations were measured in the GHz range by G.Mansfeld [11]. We have also previously obtained preliminary estimations of acoustic losses from resonators Q factor measurements [12].

On figure 1 the variations of the phase velocity of the three thickness modes are represented for the Y rotated cuts. The two shear modes B and C have velocities somehow slower than for quartz. The velocity of the longitudidal mode is quite high relatively to the moderate shear velocities.

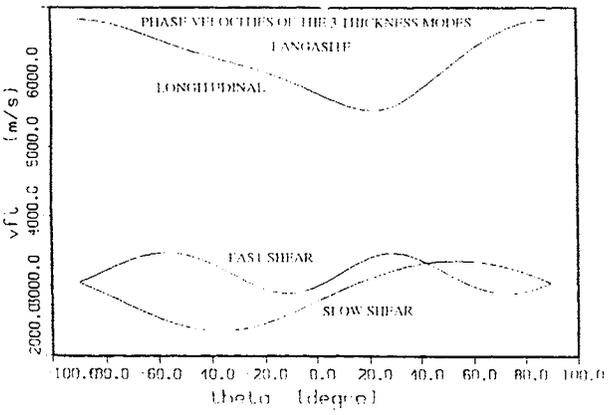


Figure 1: Phase velocities of the thickness modes.

The coupling coefficient of the only piezoelectric transverse mode has the same angular variations as for the other materials of crystal class 32. As for quartz and its analogous, the piezoelectrically active mode is depending on the rotation angle, either the fast shear (B mode) or the slow shear (C).

The temperature coefficients of the resonance frequencies of the piezo-active mode has a zero value for two Y rotated orientations (figure 3a). The compensated cut situated near the Y one corresponds nearly to the largest possible value of the coupling coefficient in

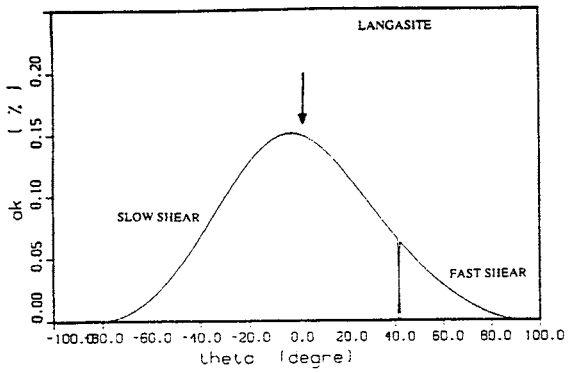


Figure 2: Coupling coefficient of the piezo-electrically excited thickness shear.

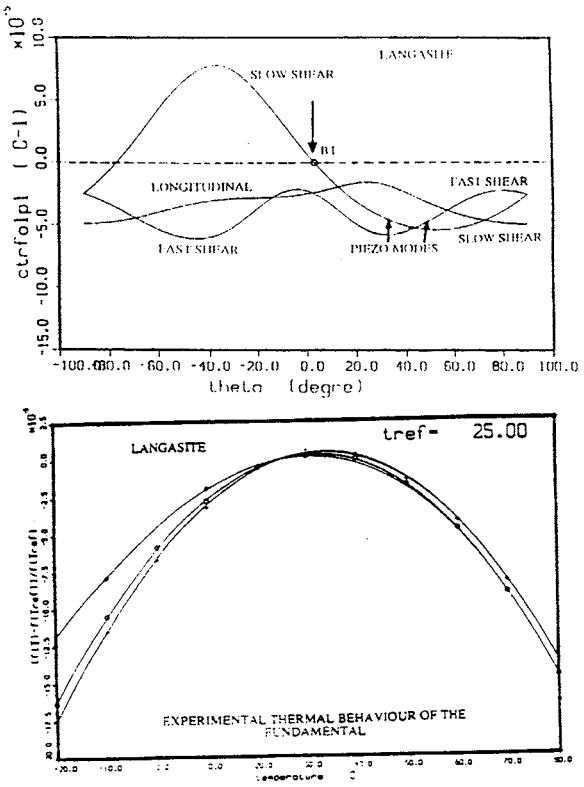


Figure 3: Temperature coefficients of Fr for the thickness modes (3a) and thermal behaviour of the main compensated cut (3b).

langasite. The angular position of this cut seems to depend on the exact composition of the material. We have observed that for Nd doped crystals, it corresponds to a zero rotation angle. The second order temperature coefficient is moderate (figure 3b). In table 1, we have summarized the properties of the Y cut of LGS.

Table 1: Properties of the Y cut.

Density=5.75e3, ε_{22}=18.8, ε_{33}=49.3.
Mode C: k=16.4% V=2790m/s, TCFa#0. $TCFa^{(2)}$=-63e-9, α=1.05 dB/(s.GHz^2)**
Mode B: V=2995m/s, TCFa#-2.e-5
Mode A: V=5780m/s, TCFa#-2.5e-5, α=.75 dB/(μs.GHz^2)**

(**)[11]

I.2 Energy trapping in the Y rotated cut of langasite.

A study of the energy trapping properties in the Y rotated cut of LGS was made using models based upon the equations given by H.F. Tiersten and co-authors [15][16][17].

On figure 4 we have represented the angular variations of the square root of the ratio of the two first coefficients of the Tiersten equations governing the lateral variations of the main displacement (u_1) of the different overtones of thickness shear langasite resonators (thickness excitation).

The solutions of these equations, are very dependant on this ratio [18] which, when it is real, is one of the most important factor of the lateral variations of the mode. When it is imaginary no conventional energy trapping can occur. For the Y rotated cuts of class 32 materials, the M_n coefficients can be real positive or negative numbers or complex numbers. For the plano-convexe resonators, in which the energy trapping results mostly of the contouring, the lateral anisotropy of the mode is governed by $\sqrt{\gamma}=\sqrt{\sqrt{M_n/P_n}}$ [16][19].

The results displayed in figure 4 indicate that γ is always a real number, greater than unity for the fundamental mode. This means that the vibration modes excited by circular electrodes are elongated in the direction of the two fold axis x. The lateral anisotropy is quite large for this mode of the compensated cut (#Y+3 or Y) of langasite as compared to the cases of AT and SC quartz.

We can also observe in figure 4, from the values of the γ coefficients computed for the different overtones in the vicinity of the compensated cut (situated near the Y one), that there are several angular regions where no energy trapping exists for a given overtone. Particularly, energy trapping disappears for the third, the 7th, and the 11th overtone at small positive angles, situated close to the Y cut. For these overtones, this implies fast variations and hence noticeable angular or "geometric" sensitivities for the anisotropy of the mode.

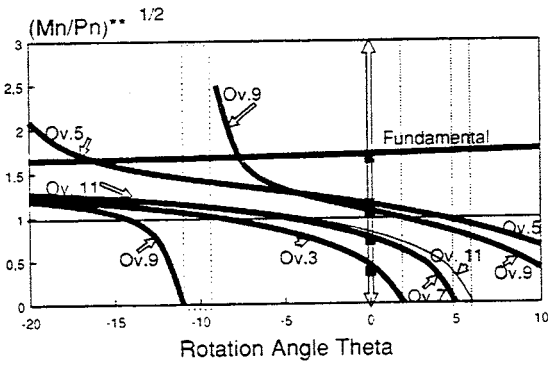


Figure 4: Energy trapping near the main compensated Y rotated cut.

According to the computed values of γ, the 3rd overtone is elongated in the z direction while they predict a more isotropic behaviour for the 5th and the 7th overtones which are elongated in the x direction. On figure 4, we have shown by square dots the experimental observations of the mode shapes that were made using X-ray topography (see below § II.2).

I.3 Energy trapping in the plano-convexe resonators with lateral field excitation.

Since there is absolutely no theoretical differences between the two cases (thickness or lateral excitation) when all the electrodes are shorted, the eigen modes of the lateral field resonators are also the solutions of the homogeneous forms of the Tiersten equations. The only differences that are to be considered are in the geometric distribution of the electroded and unelectroded regions across both type of resonators. This can be taken into account in several ways, in the resolution of the equations [19], but due to the very small influence of the mass loading in plano-convexe resonators (particularly for LGS which has a large density), this has only a practical effect on the values of the eigen frequencies (effect of the order of a fraction of a mass loading which is typically of about 0.1%). Further more, this small effect can be considered as negligible for small gaps, and to the contrary can be very easily taken into account for large gaps (electrical effect and mass loading #0.). In the intermediate cases a satisfactory approximation can be found.

To obtained the forced solutions, and to calculate the equivalent scheme of such resonators the appropriate inhomogeneous term (electric forcing term) can be determined using a transformation of the electrical potential made so to have with the transformed potential

V' the boundary condition V'=0. This can be done using the expressions of the electrical fields derived in the one dimensional theory of lateral field resonator resonators [20].

For a lateral excitation by a z field, the model predict that the 4th and the 8th overtones of the C mode are not trapped while the 2nd and the 6th are trapped. We have observed experimentally a trapped behaviour for the 4th and an untrapped behaviour for the 8th.

II. LGS PLANO-CONVEXE RESONATORS WITH THICKNESS EXCITATION

II.1 Experimental results.

Plano-convexe Y-cut resonators using 15x15mm square plates cuts in non-doped crystals elaborated by Crismatec were made using different designs (radii of curvature, electrode parameters,...). Some experiments were also made with Nd doped plates (10.mm diameter). Most of the studied resonators were made to have a thickness of the same order of magnitude as 10MHz 5th overtone quartz. Some samples nearly twice thicker or twice thinner were also experimented.

II.1.1 Experimental Results.

On **table 2**, the results of measurements made on a Y cut plano-convexe LGS resonator with thickness excitation and a curvature radius of 250mm are displayed together with the design parameters. High Q factors were measured for the resonance frequencies of the 5th and almost the 7th and the 9th overtone. It was also observed that the electrical impedance level corresponding to these resonances are much lower than what is observed for the corresponding overtone of AT quartz. This fact results of the larger values of the dielectric constant and of the coupling coefficient. It was also observed that these resonator withstands extremely high excitation levels without showing any evidence of non-linear behaviour (typically +10 to +15dBm source power, 1/5 to 1/10 of that being dissipated in the resonator).

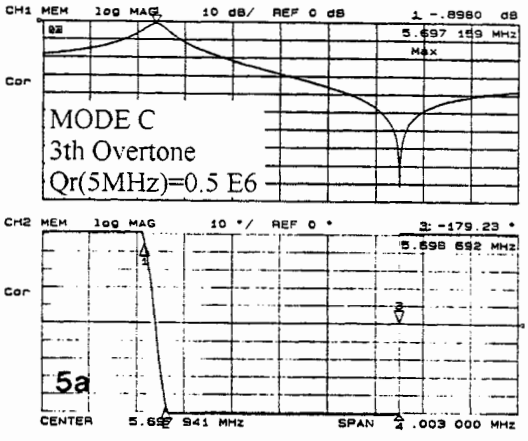
Table 2: Characteristics of a plano-convexe resonator with a curvature radius of 250mm.

Plano-convexe Thickness field
Y cut 2h=0.725mm Rc=250mm 2Re=7mm

Ov.	$Fr_{(MHz)}$	$Q_{(Fr)}$	$Q_{(5MHZ)}$	$L_{(H)}$	$R_{(Ohm)}$
1	1.918	0.064	0.024	0.045	8.56
3	5.698	0.439	0.500 *1.356 (Fa)	0.132	10.89
5	9.497	0.536	1.019 *0.995(Fa)	0.154	17.15
7	13.289	0.605	1.608	0.193	26.73
9	17.097	0.387	1.325	0.209	45.0

* measurement at Fa

The response of the fundamental mode and the 3rd, 5th, 7th overtones are displayed in figure 5. The Q.f product measured at the resonance of the 7th overtone (Q.f=.785 10^{15} for this resonator) is of the same order of magnitude as the values obtained using the HBAR technique by G.D. Mansfeld whose measurements constitute an other important indication of the possibilities to obtain extremely high Q factors from LGS devices.



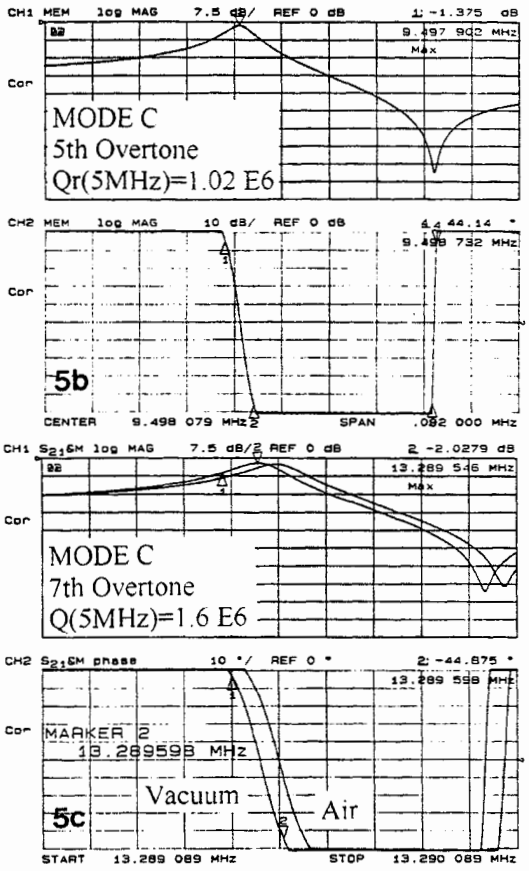


Figure 5: Electrical response of the different overtones of a plano-convexe resonator with Rc=250mm.

In **table 3**, the results of the measurements made on a Y cut plano-convexe LGS resonator with a radius of curvature of 200mm are displayed together with the design parameters. For this resonator, which was the best obtained with this curvature, high Q factors were measured for the resonance frequencies of several overtones. The same conclusion as for the previous resonator can be drawn from these results. Still higher Q factors were measured at the antiresonances of some overtones. The mode shapes corresponding to the different overtones were observed by X-Ray topography (see below § III). The measurements of **table 3** were obtained at a source power of 15 dBm. No distortion of the amplitude response was observed even at higher excitation power (up to +20dBm).

Table 3: Measurements obtained for a resonator with Rc=200mm.

Thickness excitation LGS Y cut
2h=0.737mm 2Re=6mm Rc=200mm R=2.4 e-3

Ov.	$Fr_{(MHz)}$	$Q_{(Fr)}$	$Q_{(5MHZ)}$	$L_{(H)}$	$R_{(Ohm)}$
1	1.887	0.081	0.030	0.064	9.32
3	5.590	0.527	0.590	0.149	9.97
5	9.320	0.600 *1.471	1.230 *2.743	0.177	15.7
7	13.039	0.602 *0.772	1.570 *2.013	0.218	29.7

(*) Q factors measured at the antiresonance

II.1.2 Very high Q factors at the antiresonance. On figure 5 and in table 2 and 3 we can observe that several modes have the unusual property of having much greater Q factors at the antiresonance than at the corresponding resonance frequencies. Still higher values of the Q.f product were measured at the antiresonance of other resonators. An example corresponding to observations made for the 5th overtone of a thicker plano-convexe resonator is given in **figure 6b**.

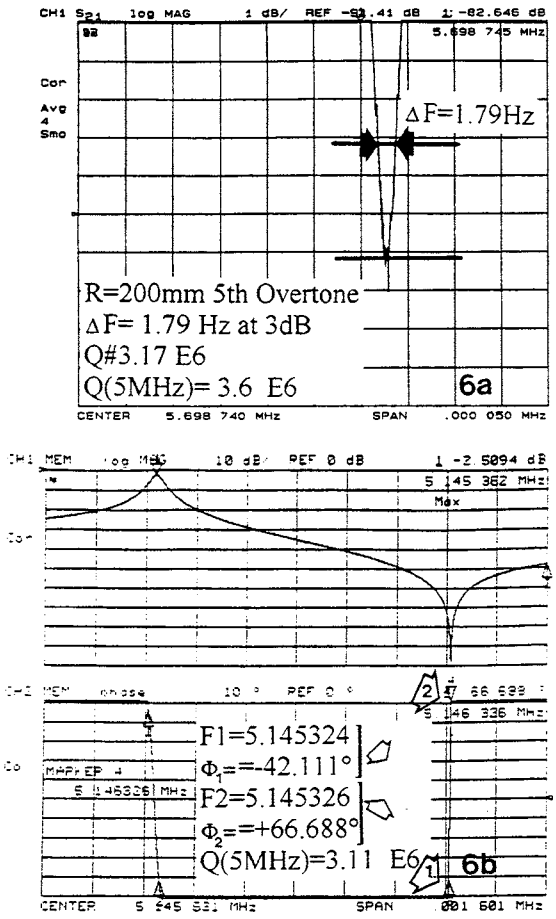


Figure 6: 3dB(6a) and phase measurement (6b) of very high Q factors at Fa.

Although these measurements are much less accurate than at the resonance frequencies and require some particular cares (quartz holder, coaxial hardware with very reduced dielectric losses etc...), it was verified that very large values of the Q.f product can be observed for the antiresonances of the 3rd and 5th overtones of most of the measured resonators. This fact which can not be predicted by the usual equivalent scheme, seems to indicate that an unusual mechanism of losses exist at (or is predominant only near) the resonance, in such resonators. This fact is also a first indication that much higher values could be obtained at the resonance if the proper solutions could be found. In figure 6 we have represented the observations made for two resonators (6a measurement at 3dB for a resonator with Rc=200mm, 6b phase measurement for a resonator with Rc=175) Values largely over 3.10^6 at 5 MHz ($Q.f>1.5\ 10^{13}$) were measured in several instances.

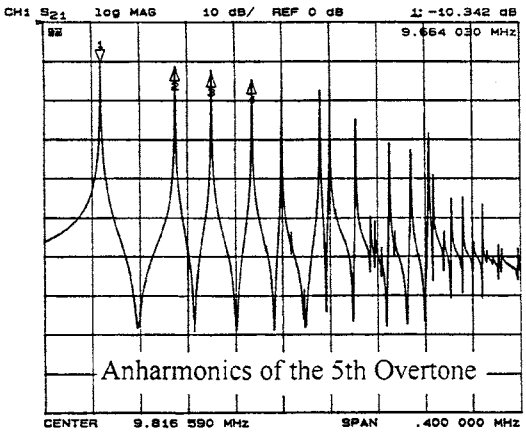


Figure 7: Anharmonics of the 5th overtone in a near 5MHz 5th overtone LGS resonator.

II.1.3 Anharmonic response. On figure 7 we can observe that a much larger number of un-attenuated anharmonic modes propagate in these resonators (5th overtone of a Rc=200mm) than in a quartz having a similar design. This results of the fact that the "dimension" of these modes are much smaller than for quartz resonators and also it is a consequence of the larger coupling coefficient (lower impedance level) which leads to much smaller attenuation in a 50 Ohm measurement system.

II.1.4 Influence of the curvature radius.

A systematic study of the influence of the curvature on the Q factor was made for resonator having a thickness in the range .7 to .8mm, in all cases plates obtained from undoped crystals were used except for the 300mm radius of curvature where Nd doped material was used. The results are given for the 3rd, 5th and 7th overtones in figure 8 where it can be observed that the best results are obtained for Rc=200mm and Rc=250mm. The results obtained for Rc=300 may be some how pessimistic due to the smaller diameters (10mm) of the plates used.

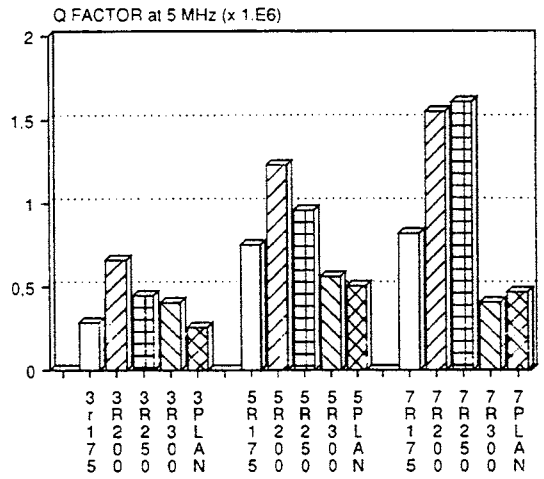


Figure 8: Influence of the curvature radius on the Q factors of different overtones.

During this study, for certain curvatures, it was noticed that an interference can exist between the 9th overtone of the C mode and the weak anharmonics of a high overtone of another thickness mode (A or B) which are excited due to the non uniform field existing in plano-convexe resonators. We have previously observed, by X-ray topography, that a similar situation exists for the 7th overtone of certain plano-convexe AT quartz resonators. When there is no interference, very high Q factors can be obtained for the 9th overtone.

II.2 Comparision with computed predictions.

II.2.1 Comparison of computed and experimental mode shapes for langasite.

On figure 9, the mode shape experimentally observed for the different overtones of the Y cut plano-convexe langasite resonator of table 3 (plate thickness =.756mm, radius of curvature 200. mm, round electrodes 6.mm diameter, mass loading =1.04 %) are compared with the computed ones. The computations were made using methods of resolution of the Tiersten equations that was previously reported [19]. The constants of references [9] and [10] were used, they lead to some how different results. On figure 8, the represented mode shapes were computed using the most recently published constants [10].

For the fundamental, the third and, to a more limited extend for the fifth overtones the agreement between the experimental mode shapes, and the computed ones is good. As expected from the theory, the anisotropy of the fundamental mode is very important and as predicted by the calculations, the direction of maximum extension of the 3rd overtone (z) is different of those existing for the fundamental and the fifth overtone (x). The agreement is much less satisfactory for the 7th and the 9th overtones. The differences observed result probably of an insufficient agreement between the material constants used for the computations and the 'real' constants of the high purity material used for the experiments. These difference are very similar to those that we have observed for plane resonators made using a Nd doped crystal [12] and the hypothesis made in this latter case of a difference in the values of the constants due to the Nd dopant cannot hold in the present case. We must also notice again the fact that the calculation of the coefficients of the Tiersten equation involves the evaluation of weighted differences of several constants, and hence these coefficients can be quite sensitive to small imprecisions in some

of the material constants. This hypothesis is accredited by the difference found between the Mn coefficients computed from several sets of constants and also by the limited accuracy, as compared with quartz, of the predicted frequencies for the different overtones and for their anharmonics.

On the whole, the computed and observed mode shapes, are very different of those known for quartz or berlinite plano-convexe resonators (different anisotropy and much smaller spatial extension at equal power). The smaller "dimensions" of the modes are predicted by the theories of the plano-convexe resonators. In the theory of H.F.Tiersten [16], it results of the much greater values of the coefficients α_n, β_n in the expressions of the Gaussian terms of the Hermitto-Gaussian main displacement (u_1).

On figure 9, we can also observe several indications of the existence of a strong interaction between the dislocations and the modes (see particularly the deformation of the elliptical theoretical shape for the 5th and the 9th overtone). Further results concerning the crystal quality and the mode shapes observed by synchrotron X-ray topography for other resonators can be found in reference [21].

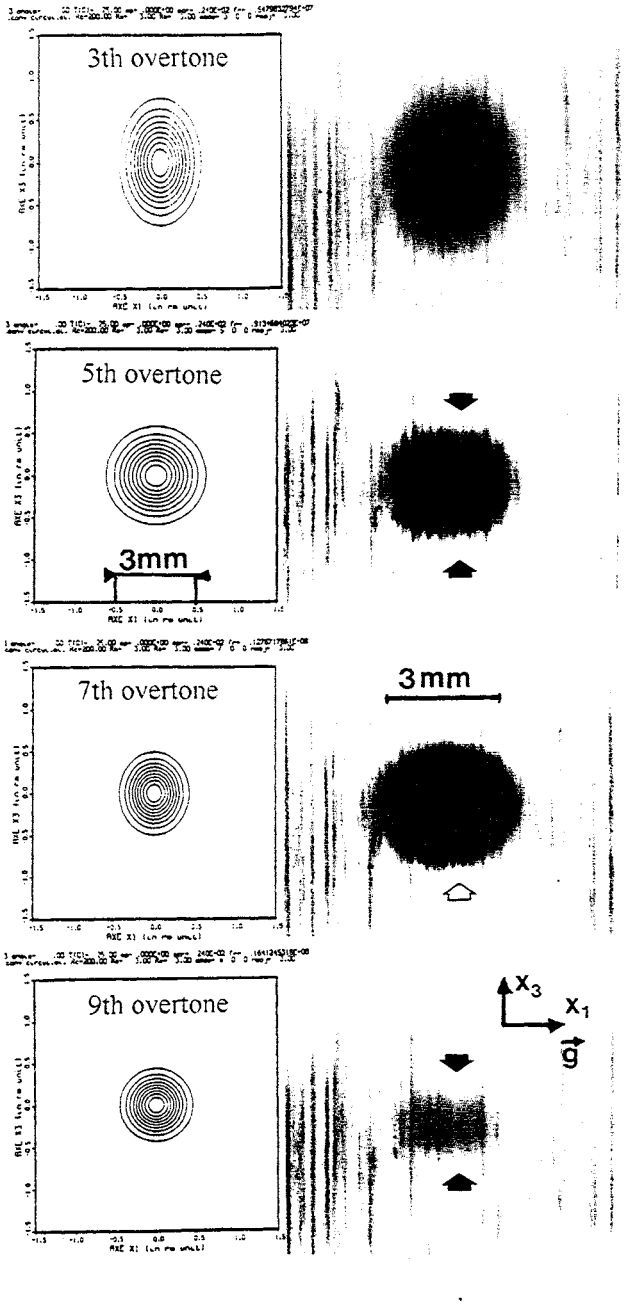


Figure 9: Comparison of experimental results with computed ones.

II.2.2 Comparison of resonance frequencies and equivalent schemes

In table 4 we compare the experimental resonance frequencies and inductances with computed ones in the case of the Rc=250mm resonator of table 2. In this case the constants of reference [9] were used. The agreement is reasonably good between the experimental and the computed frequencies. We can however observe that the slight differences (#3%) in the values of the frequencies are about one order of magnitude larger than the experimental uncertainties on the measured values of the thickness and of the mass loading that were used for the calculations. For the inductances the agreement is very good for the fundamental, the 3rd and the 5th overtone; it is only quite good for the 7th overtone. This latter difference result much probably of the already observed difference between computed and experimental mode shapes.

Table 4: Comparison of experimental and computed resonance frequencies and inductances.

	EXPERIMENT		COMPUTED (*)	
Ov.	$Fr_{(MHz)}$	$L_{(H)}$	$Fr_{(MHz)}$	$L_{(H)}$
1	1.918	0.045	1.922	0.043
3	5.698	0.132	5.707	0.138
5	9.497	0.154	9.514	0.145
7	13.289	0.193	13.300	0.251
9	17.097	0.209	17.099	0.270

Again, the very small values of the inductances, as compared with the cases of AT or SC quartz, must be noticed.

A comparison of the experimental and computed frequencies of the anharmonics was made for the different overtones of resonators having different designs. It has shown that the frequency distribution of the anharmonics is generally not very accurately predicted. This seems to indicate that the accuracy of one or several of the constants contained in the expression of the coefficients of the Tiesten equation [Mn (C'_{11},C'_{12}....) and Pn (C'_{55})] is probably somehow lower than those of C'_{66}. An example is given in table 5 for the case corresponding to figure 7.

Table 5: Frequency distribution of the anharmonics

Mode:	5.0.0	502	520	504	522	540
Experimental	9664.kHz	+62kHz	+93	+127	+152	+183
Computed	9698.kHz	+72kHz	+82	+144	+153	+163

III PLANO-CONVEXE RESONATORS WITH LATERAL FIELD EXCITATION.

III.1 Lateral field excitation in the x direction.

The low impedance level predicted and observed for conventional plano-convexe LGS resonators indicate very interesting possibilities for lateral field excitation. In table 6, some results obtained for a resonator using a lateral excitation by a field directed in the x direction are presented. As indicated by several authors and particularly by A.Ballato for quartz and other materials (See the references given in [22]), the x field excite the longitudinal modes and the fast shear modes. We can notice that very high Q factors are obtained for several overtones of the longitudinal (A) mode whereas the overtones of the fast shear (B) present much higher acoustic losses with this design. It has been observed that several anharmonics of the even overtones of the C mode that are excited due to their antisymmetry relatively to the collecting electrodes have also interesting properties. The impedance levels obtained with this design are high but can be some how reduced using the gap parameter.

Table 6: Characteristics of some of the modes excited by a lateral field in the x direction in a plano convex resonator.

LATERAL FIELD EXCITATION

Y cut 2h=1.746mm Rc=150mm

E field in X direction Gap=1.5mm

Mode	$Fr_{(MHz)}$	$Q_{(Fr)}$ $_{(10^{-6})}$	$Q_{(5MHZ)}$ $_{(10^{-6})}$	$L_{(H)}$	$R_{(Ohm)}$
A5	8.284	0.681	1.296	4.86	371.
A7	11.577	1.154	2.672	6.06	381.
A9	14.869	0.656	1.952	4.69	668.
B5	4.333	0.176	0.153	5.13	791.
B7	6.023	0.256	0.309	8.74	1077.
C4 410	3.192	0.724	0.462	30.95	857.
C6 610	4.769	1.040	0.993	12.95	370

VERY HIGH Q FACTORS in LGS RESONATORS
A mode, Y cut, L.E, X field, 2h=1.736 Rc=150mm

Overtone	Fr (MHz)	L (H)	Q at Fr ^{*1E6}	Q at 5MHz ^{*1E6}
5	8.327	6.05	1.050	1.749
7	11.636	7.71	1.473	3.429
9	14.946	7.67	1.212	3.625

*Same magnitude as alpha=.75dB/us (Q=3.3E6 at 5MHz)

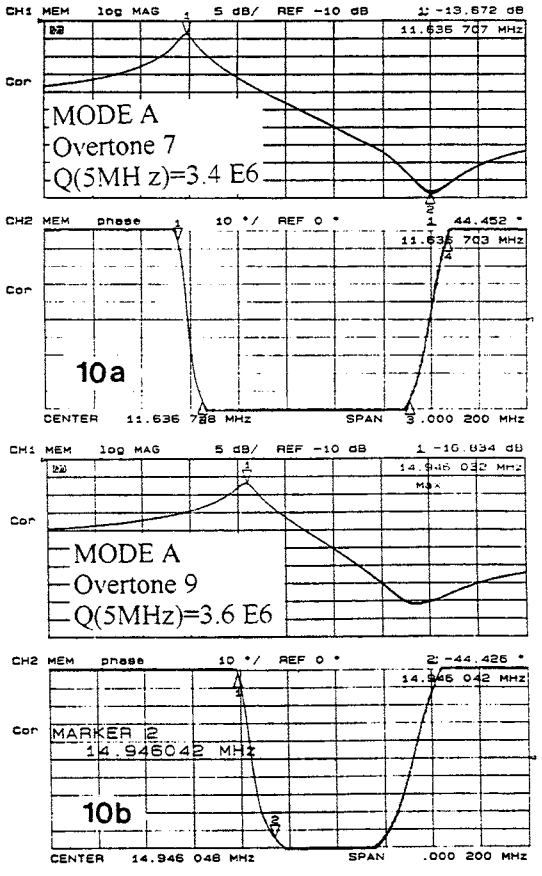


Figure 10: Characteristisc of a very high Q resonator (lateral field excitation) and response curves of the 7th and the 9th overtones of this resonator.

Much higher Q factors were observed with the overtones of the A mode for a resonator having a similar design, but a reduced gap. In figures 10a and 10b, we can observe outstanding Q factors for several overtones of the A mode (best Q.f product =1.84 10^{13} for the 9th). This latter value is higher than those observed for the Y cut of quartz, it is also moderately higher than the values measured by G.Mansfeld (1.5 10^{13}). Again higher values were observed for the antiresonances (up to about 2.25 10^{13}).

III.2 Lateral field excitation in the z direction.

For LGS, the dielectric constant in the Z direction ε_{33} has a much larger value than those relative to the x direction ε_{11} so that using lateral excitation in the Z direction, leads to reduced impedance level (the inductance is a linear function of $1/\varepsilon$).

On figures 11a.b,c, we can observe that then the 4th and the six overtone of the C mode so excited, have impedance levels very similar to that of 5th overtone quartz using thickness excitation. The Q factor of the 4th overtone is 1 million and those of the 6th overtone is .75 million. The response curve of figure 11c indicates that the distribution of the anharmonics of the 6th overtone is very similar to what is known for the odd overtones excited by a thickness field (this constitute an experimental confirmation of the considerations about the eigen modes of L.E. resonators made in § I.3).

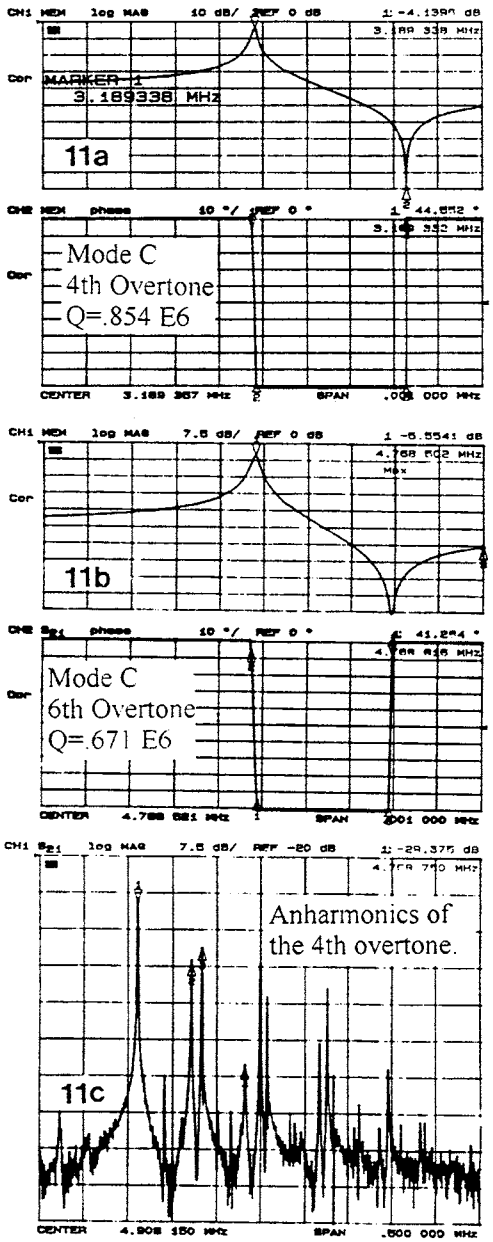


Figure 11 Responses curve and characteristics of even overtones of the C mode.

IV DISCUSSION.

IV.1 Influence of the dielectric losses in the high Q resonators.

In an attempt to determine why the Q factors measured at some antiresonance frequencies are much higher than those observed at the resonance frequencies, we have made an evaluation of the influence of the dielectric losses on the Q factors of the langasite resonators. A much greater influence than for quartz is expected since the coupling coefficient is much larger. So the electric energy part of the stored energy is much higher, and the influence of dielectric losses can be much higher. To illustrate this question let us consider the result of the one dimensional theory of the resonators with which the acoustic attenuation factors are defined and measured. In this theory [14], the acoustic velocities of the three thickness modes are computed from the stiffened elastic constants $\overline{C}^{(i)}$ (i=1,3), which are the eigen values of the piezoelectric Christoffel matrix:

$\overline{C}^{(i)}$ eigen-value of Γ_{kl}

The elements of the Chistoffel matrix are functions of the elastic, dielectric and piezoelectric constants and of the orientation of the normal to the plate. So that, in the general case the $\overline{C}^{(i)}$ are also functions of the same quantities. However, for orientations possessing certain symmetries a much simpler expression is found with the form: $\overline{C}^{(i)} = C^{(i)}(1.+k^2)$For example for the Y rotated cuts of class 32 crystals:

$C^{(3)} = C'_{66}$, where C'_{66} is the rotated C_{66} constant.

The coupling coefficient k is defined by:

$k^2 = \frac{e'^2_{26}}{\overline{C}\,\varepsilon'}$ so that $V_i = \sqrt{\frac{\overline{C}_i}{\rho}}$

As usually, we include the dielectric and acoustic (viscous) losses in replacing the real constants by imaginary ones:

$\varepsilon = \varepsilon' - j\varepsilon''$

with:

$Q_{capa} = \frac{\varepsilon'}{\varepsilon''}(\omega) = \frac{1.}{\mathrm{tg}(\delta)}$

$C^* = C' + jC''$ with$Q_{meca} = \frac{C'}{C''}(\omega) = \frac{1.}{\eta\omega} = \frac{\omega_r}{2.\alpha}$

where α is the acoustic attenuation (dB/µs).

In including the complexe values of k.C and ε in the expression of the impedance:

$Z = \frac{1}{j\omega C_0}\left\{1 - \sum_{i=1}^{3}\frac{k_i^2\,\mathrm{tg}(\theta_i)}{\theta_i}\right\}$

It is possible to compute the Q factors of fr and fa provided the dielectric and acoustic losses are known.

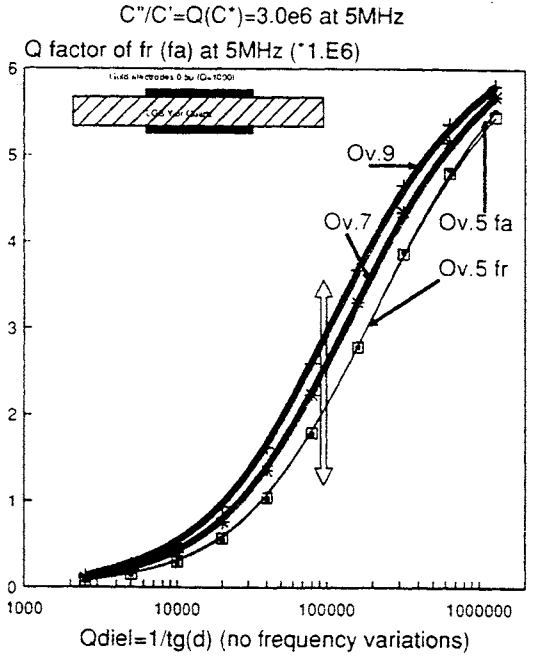


Figure 12: Influence of the overtone rank and dielectric losses on the Q factors.

Due to the fact that the values of the acoustic attenuation known for LGS and quartz may have been obtained without accounting for the dielectric losses, we have considered in our calculations a range of dielectric and elastic losses. To account also for a possible effect of the acoustic losses in the electrodes we have made these calculations using a one dimensional model of composite (layered) structures, which is equivalent for each layer to the previously considered model and ensure the required continuity conditions between the layers. The calculations were made for Y LGS 5Mhz 5th overtone resonators and for AT quartz 5MHz 5th overtone resonators electroded by gold layers (generally .5µm thick). No frequency dependence was considered for the dielectric losses.

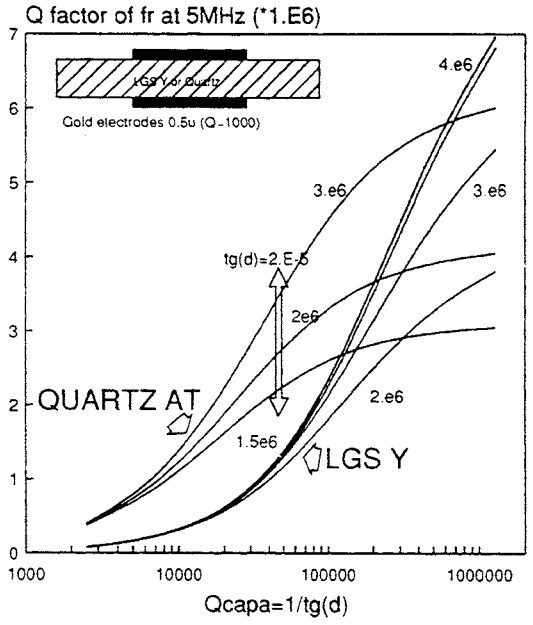


Figure 13: Comparision of the Q factors for AT quartz and LGS resonators as a function of acoustic and dielectric dissipations.

On the figures 12 and 13, we have represented the computed Q factors of the resonators as a function of the dielectric losses. In figure 12, we consider the influence of the overtone rank (5, 7, 9) in the LGS resonators. The computed Q factors of the resonance frequencies are drawn as a function of $1/\mathrm{tg}(\delta)$ for $Q_{meca}=3.10^6$ at 5MHz, the Q factor of the antiresonance frequency is drawn for the 5th overtone. In this case, it was considered that $1/\mathrm{tg}(\delta)$ has no frequency variations in the limited frquency range corresponding to the increase fron 5 to 9 of the overtone rank. On figure 12 we can observe that:

- **The dielectric losses have a large influence on the Q factors**
- There is only a very small difference between the Q factors of Fr and of Fa
- The Q factors vary greatly with the overtone rank.

On figure 13, we have considered for the 5th overtone of LGS and Quartz more hypothesis for the value of Q_{meca}. Several very important facts are to be noticed:

- To give the same Q factor as quartz, LGS must have much reduced dielectric losses (by at least a factor 4).
- With a value of $1/\mathrm{tg}(\delta)=5.10^6$ for quartz (one of the most commonly admitted value for this frequency range for alpha quartz and for pure silica) a Q.f product of $1.5\ 10^{13}$ is achieved for $2.10^6<Q_{meca}<3.10^6$.
- The curves seem to indicate that **the Q factors of actual high Q devices made either with AT quartz or Y langasite, could be limited by the dielectric losses.**

From this point of view, a detailed comparison between the cases of the C modes excited by lateral X or Z field would be very interesting, if the two corresponding dielectric loss tangents were known very accurately (The Q factors seems to be somehow greater for an x field).

A more attentive examination of the expression of k and Z would have revealed that a small "compensation" can exist between the dielectric and acoustic losses and that the influence of the dielectric losses increases rapidly with the coupling coefficient (nearly as k^2). ***This imply that the materials with large k, for which high Q factor devices were observed, must have extremely reduced dielectric losses (lithium tantalate, lithium niobate, LGS, etc...).***

These calculations made for a one dimensional resonator are expected to give a correct order of magnitude of what happens for energy trapping resonators, but a better accuracy could probably be obtained if they can be corrected for the actual mode shape and electric field distribution. They indicate also that it could be very difficult to have separate measurements, in the same frequency region, of the dielectric losses and of the acoustic losses for a piezoelectric material.

IV.2 Material perfection, technological aspects.

Several of the LGS samples used in this study have shown a very reduced density of dislocations [13][21] so that this parameter is probably not a limiting factor for the Q factors. But several indications were gained during this study of the fact that LGS could contain point defects and that this material could be further improved by reducing their densities. Several observations probably related to the latter, have shown that X-ray irradiation and medium temperature annealing of the resonators can increase somehow the Q factors. As for any very high Q devices it was found that the polishing and etching processes were very critical [12] [23].

IV.3 Design of very high Q LGS devices.

The vibration modes of LGS plano-convexe resonators are very different of those existing in quartz devices, it is quite probable that another mechanism of energy dissipation exist for these device (difference between the Q factors of Fr and Fa). Furthermore the anisotropy of the most interesting vibration modes (7th and 9th overtone of thickness shear) is very different of those known for AT and SC quartz, so that it is probably still possible to find much better designs for very high Q LGS devices. The solution of this question will probably need to get a better insight in the losses mechanisms (actual dielectric loss tangent, etc...), and a systematic investigation of other types of resonator geometries and modes. Depending on the exact nature of the losses, the use of new types of modes such as, the presently some how exotic one represented in figure 14, could be interesting. This kind of mode that could be classified as an acoustic, energy trapping whispering gallery mode (here an anharmonics of the 5th overtone of a resonator with a particular geometry) could present several advantage for high Q applications as does their electromagnetic counterparts.

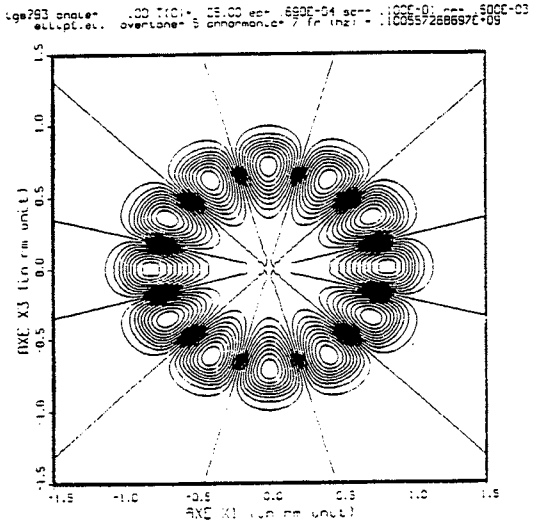


Figure 14: Exemple of a type of mode which can present some interest to obtain extremely high Q factors.

V CONCLUSION.

The present study has shown that LGS plano convex resonators display extremely high Q factors. There is several indications of the possibility of still much larger values for the resonance frequencies.

The larger coupling coefficients of LGS allows to use overtones with much higher ranks than for quartz. This fact, which overcompensate the somehow slower phase velocities, is very favourable to extend, toward higher frequencies, the application domain of plano-convexe resonators. It can also permit to obtain much larger shifts for the oscillators.

Several points must be further studied such as the question of the exact losses mechanisms and the optimal design of very high Q devices. The accuracy of the material constants can also, much probably, still be improved. The material available is already of very good quality, a possibility of increasing further its quality for the applications to devices with extreme performances seems however possible. This material allows substitutions and dopping that can also lead to still larger qualities.

On the whole, this material appears as very promising to obtain extremely high Q factor devices for applications requiring very low phase noise or extreme short term stability. Its great potential for other applications such as bulk wave filters and surface wave devices were previously demonstrated.

Acknowledgements: The authors acknowledge the assistance of MM. G. LeThuault and G. Caumont for the elaboration of the studied samples and the helpful contributions of M Joly and A.Jeanne-Michaud.

References:

[1] J.Detaint, J.Schwartzel: Proc. 8th European Frequency and Time Forum Münich March 1994 p.175-189.
[2] A.Ballato, J.G.Gualtieri, Proc.1994 IEEE Frequency Control Symposium. p.22-34 (1994).
[3] A.A. Kaminski, B.V. Mill, G.G.Khodzhabagyan: J. Phys. Stat. Sol. V.A80 p.387 (1993).
[4] A.A.Kaminski, I.M.Sylvestrova, S.E.Sarkisov, G.A. Denisenenko; J. Phys. Stat. Sol. V.A80 p.607 (1993).
[5] A.N. Gotalskaya, D.I. Drezin, V.V.Bezdelkin, V.N.Stassevich; Proc. 1993 IEEE Int. Frequency Control Symposium p. 339-346 (1993).
[6] I.M.Silvestrova, V.V.Bezdelkin, P.A.Senyushenkov, Yu.A.Pisarevsky. Proc. 1993 IEEE Int. Frequency Control Symposium p. 348-350. (1993).
[7] D.Cochet-Muchy J. de Physique IV vol 4 n°2 p.C2-33,45;(1994)
[8] M.F.Dubovik, I.A.Andreyev, Yu.S.Shmaly; Proc.1994 IEEE Frequency Control Symposium p.43-47 (1994).
[9] S.A.Sakharov, I.M.Larionov, A.V.Medvedev; Proc. 1992 IEEE Int. Frequency Control Symposium p.713-723 1992.
[10] V.V. Bezdelkin, E.E. Antonova; J. de Physique IV vol 4 n°2 p.C2-139,156 (1994).
[11] G.D.Mansfeld; Proc. 1994 IEEE Frequency Control Symposium p.35-39 (1994)
[12] J.Détaint, J.Schwartzel, A.Zarka, B.Capelle, J.P.Denis, E.Philippot; Proc. 1994 IEEE Frequency Control Symposium p.58-71 (1994).
[13] B.Capelle, A.Zarka, J.Détaint, J.Schwartzel, A.Ibanez, E.Philippot, J.P.Denis; Proc. 1994 IEEE Frequency Control Symposium p.48-57 (1994).
[14] A.Glowinski; Annales Telecom. v.27 n°3-4 p.147-158 (1972)
[15] H.F. Tiersten; J. Acoust. Soc. Amer. v.59 n°4 p.879-883 (1976).
[16] H.F.Tiersten, R.C.Smythe; J. Acoust Soc. Amer. v.65 n°6 p.1455-1460 (1979).
[17] D.S Stevens, H.F.Tiersten; J. Acoust. Soc. Amer. v.79 n°6 p.1811-1825 (1986).
[18] J.Detaint, J.Schwartzel, C.Joly, E.Philippot; Proc. 41st Annual Frequency Control Symposium p.314-324 (1987).
[19] J.Detaint, H.Carru, J.Schwartzel, B.Capelle, A.Zarka; Proc. 42rd Annual Frequency Control Symposium p.19-28 (1988).
[20] T.Yamada, N.Niizeki; J.Appl. Phys. v.41 n°9 p.3604-3609 (1970).
[21] A.Zarka, B.Capelle, J.Détaint, D.Cochet-Muchy; Proc.of the 7th European Frequency and Time Forum (these proceedings) (1995).
[22] A.Ballato, E.R Hatch, M.Mizan, B.H.T.Chai, R.S.Tilton, T.J.Lukaszek; Proc 38th Annual Frequency Control Symposium p.206-224 (1984)
[23] S.Laffey, M.Hendrickson, J.Vig, Polishing and Etching Langasite Crystals; Proc.1994 IEEE Frequency Control Symposium p. (1994).

ORIGINAL MODE SELECTION FOR WHISPERING GALLERY MODE RESONATOR

V. Giordano, W. Daniau, G. Martin, F. Lardet-Vieudrin, M. Olivier, O. di Monaco, M. Chaubet*

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 Edouard Belin, 31055 Toulouse Cedex France

ABSTRACT :

Whispering Gallery (W.G.) mode resonators are very attractive to design high spectral purity microwave oscillators. In the W.G. mode configuration the electromagnetic field is confined in a sapphire dielectric rod. The use of high purity sapphire monocristal which exhibits very low dielectric losses allows to obtain Q-factor better than 10^6 at 10 GHz [1, 2]. At the LPMO, we have undertaken the development of a 10 GHz phase noise measurement set-up with a reference source based on a W.G. mode resonator.

One of the major problem with such type of resonator is its high density of modes. $W.G._{m,\,o,\,o}$ mode typically choosen as reference frequency could be found very close to other high order modes. High stability oscillator operation should be difficult to obtain due to the presence of these undesired modes. We have implemented an original mode selection which suppresses the spurious modes over a range of 1 GHz around the main mode. The mode selection uses direct metallic film deposition on the sapphire rod. The metallic film drawing was choosen to modify the boundary conditions for the undesired modes without affecting the main resonance.

In this paper, we present experiments demonstrating the efficiency of our mode selection method and preliminary results concerning the development of our W.G. mode reference source.

1 - INTRODUCTION

Important progress have been reported by several groups on the development of high purity X-band signal source based on Whispering Gallery (WG) mode sapphire resonator [1, 2, 3]. The use of high order mode (azimutal number $m \geq 5$) of sapphire disk resonator allows to obtain very high Q-factor which is essentially limited by the dielectric losses in the sapphire crystal it self. Q-factor of the order of 2.10^5 at 300K and better than 10^7 at 77 K have been published.

At the LPMO, where are develop phase noise measurement equipements this new technology seems to us very promising. Such signal source could constitute an ideal reference operating directly in the microwave domain. It could replace avantageously the standard low frequency source (10 MHz quartz oscillator) whose phase noise performances are degrated by the multiplication process.

We have undertaken the development of a reference source based on a WG mode resonator. In this paper, we summarize our primilinary results obtained to date.

2 - DESIGN AND MODELISATAION OF W.G. RESONATOR

The figure 1 shows a prototype of WG mode resonator. A sapphire disk is placed in the center of a copper cavity. Typical sapphire disk dimensions for X-Band are : $\varnothing \simeq 35$mm and 10mm thick. Two magnetic probes insure the coupling of the resonator. Coupling coefficients could be simply varied between 0.01 to 0.5 by adjusting the position of the probes. W.G. modes correspond to hybrid modes with a high value of the azimutal number m ($m > 4$). Such modes are characterized by a efficient confinement of the electromagnetic field in the sapphire rod. Then the Q-factor is limited by the intrinsic low dielectric losses of the sapphire. Two types of mode could be excited ; ie quasi-TM : $WGH_{m,n,l}$ and quasi-TE : $WGE_{m,n,l}$. In this paper, we will focus our attention on the $WGH_{m,o,o}$ modes.

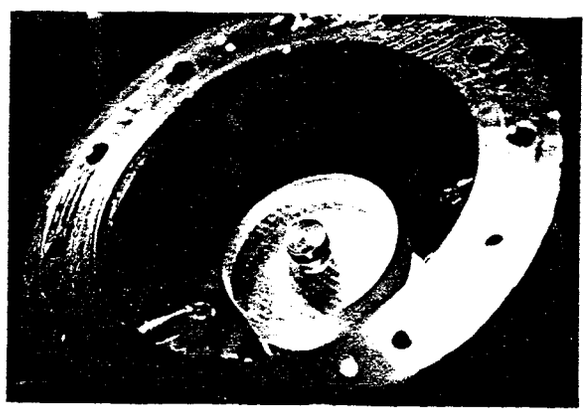


Fig. 1 - *A view of the W.G. mode resonator*

We use the approximate theory developped by Ivanov [4] to determine the resonance frequencies and Q-factor of the Whispering Gallery mode resonator. In this method, the resonator is divided in partial homogeneous regions in which electromagnetic fields are expressed in terms of single eigenmodes. Conditions of continuity on the boundaries of each partial region allows to derive a set of charateristics equations. From these equations, resonance frequencies and then relative amplitude of E and H components are determined. Measured values of sapphire loss tangent and surface resistivity of the containing metallic cavity are used to calculate the Q-factor of the resonator.

As we will see later, in spite of its relative simplicity this method gives results which agree satisfactory with the experimental observations.

It should be noted that more sophisticated methods exist as Radial Mode Matching [3] or finite element method [5]. These methods have to be implemented in a second approach if more accurate determination of resonance frequencies are needed or if a complicated design of the resonator are used.

To determine with a good accuracy the resonance frequencies and Q-factor the know ledge of several material characteristics are needed ; ie sapphire permittivity tangent, and surface resistance of containing cavity.

Sapphire crystal being anisotropic, the permittivity is tensorial. As the crystal C-axis is choosen identical to the disk axis z, we have only two different components : ε_t and ε_z are determined by measuring respectively the resonance frequency of the TE_{011} and TM_{011} modes of a Courtney type resonator [6]. The table I gives the results of our measurement for two sapphire of different origin S_1 and S_2.

	S_1	S_2	AF 997
ε_t	9,428 ± 0,002	9,428 ± 0,002	9,30 ± 0,01
ε_z	11,42 ± 0,01	11,32 ± 0,01	9,33 ± 0,02
tan δ	1.6 (± 0,2).10^{-5}	0,5 (±0,1).10^{-5}	

Tab. I - *Dielectric properties of sapphire and Aluminia at 300K*

It should be noted that ε_t is determined with a better accuracy than ε_z. This is du to the fact that the TM_{011} resonant frequency depends on more experimental parameters : ε_t, resonator dimensions and air gap between the dielectric resonator and the copper plates. To shenghten our confidence in this method we use the same experimental procedure to measure in the both directions the permittivity of an isotropic material (Aluminia AF 997). We can see on the last coulumm of the Table I that the measurement agree quite well with the prediction ; ie $\varepsilon_t = \varepsilon_z$.

Sapphire has very low dielectric losses, tan $\delta < 10^{-4}$. Q-factor of the Courtney resonator is then principally limited by the surface resistance of the copper plates. Simple measurement of this Q-factor is not sufficient to determine accuratly the dielectric loss of sapphire rod. Following the method of Kobayashi we use two sapphire rods A and B with the same diameter but with different lenght $l_B = 2l_A$ [7]. Resonators A and B are excited respectively on the

TE_{011} and TE_{012} modes. Measurement of the Q-factors of these two modes allows a precise determination of tan δ.

To determine the surface resistance Rs of the copper cavity, we remove the sapphire rod and measure the Q-factor of the TM_{011} mode of the empty cavity. Assuming homogeneous surface resistance and by applying a well established formula [8], we obtained Rs = 0,11 Ω at 3 GHz.

The previous measurements are used to compute resonance frequencies and Q-factors.

Results on our computations confronted with the experimental observations at 300 K are summarized in table II and on figure 2 and 3.

m	S_1		S_2	
	fexp	fcal	fexp	fcal
2	7,263	7,402		
3	8,658	8,628	6,182	6,235
4	9,824	9,700	7,070	7,014
5	10,878	10,751	7,855	7,778
6	11,918	11,810	8,612	8,546
7	12,962	12,884	9,381	9,330
8	14,015	13,970	10,142	10,121
9			10,911	10,918

Tab. 2 - *Resonance frequencies in GHz of WGHm,o,o modes for the sapphire resonators S_1 and S_2*
$\varnothing_1$ = 26mm, l_1 = 6.5mm
$\varnothing_2$ = 32mm, l_2 = 9mm

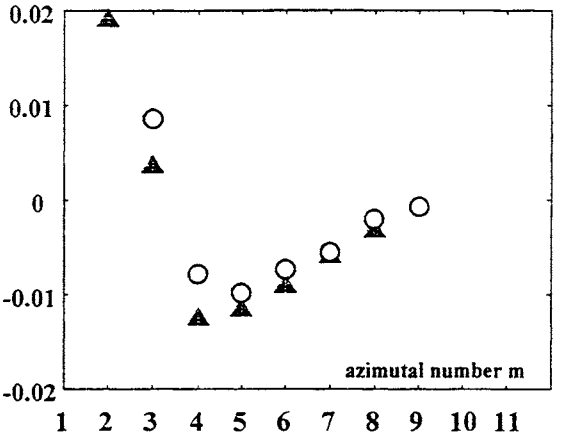


Fig. 2 - *Relative frequency error for $S_1(\Delta)$ and $S_2(O)$*

Unloaded Q-factor

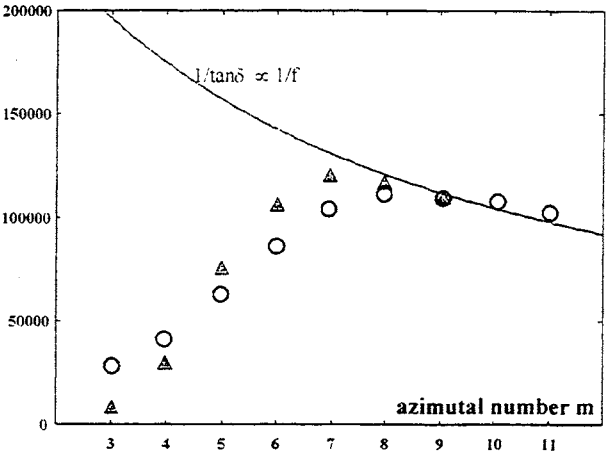


Fig. 3 - *Experimental Q-factor (Δ) at 300 K*
Calculated Q-factor (O) at 300K
$(\tan \delta)^{-1}$ at 300K

We note the good agreement between the calculated resonance frequencies and the measured ones. The precision of the model increases as the azimutal number. With m<4, clamping structure has non negligible influence on f and Q-factor. With m≥7, the electromagnetic field is well confined in the sapphire disk. The resonance frequencies are determined with a relative precision better than 0.5% and Q-factor is very close to $(\tan \delta)^{-1}$.

MODE SELECTION

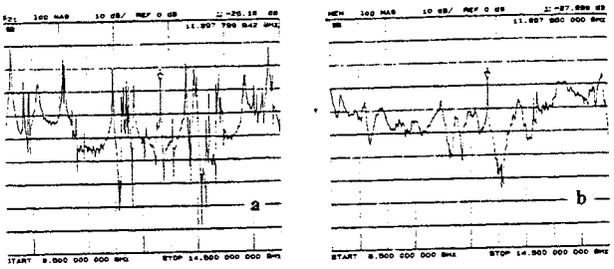


Fig. 4 - *a - WGH resonator spectrum*
b - WGH resonator spectrum with mode selection

The figure 4a shows a typical spectrum recorded with a WG resonator in its copper cavity. One of the major problem with a such type of resonator is its high density of modes. The $WG_{6,0,0}$ mode choosen here is found very close to a number of spurious modes. These underisable resonances correspond to other high order modes (as $WG_{m,n,l}$ with n and $l \neq 0$) or to the modes of the empty cavity perturbed by the sapphire disk. Stable operation could be difficult to obtain in such conditions. Modal selection have to be implemented in order to reduce the influence of the spurious modes.

In the WG mode resonator, the electromagnetic field components have periodical variations along the azimutal direction. Our idea is to place an active microwave material (absorber or metallic surface) at the nodes position (2xm nodes for $WG_{m,o,o}$). This technique is analog to the one reported in the reference [9] to enhance the performance of cavity power combiners. The originality of our technique is to deposit directly on the sapphire the mode selector. The figure 4b demonstrates the efficiency of our mode selector. It was recorded by placing on the sapphire disk a simple printed circuit made of a thin Aluminium film on a polymer substrat. The drawing of the metallic thin film was a 2xm branches star. This first approach allowed an optimization of the metallic drawing without using heavy thin film deposition technologies.

After several runs, the dimensions of the metallic lines were experimentally determined to obtain a good spurious mode selection and a small change in the operating mode. The Aluminium film was deposited directly on the sapphire disk. The figure 5 shows one of our resonators designed to operate on the $WG_{7,o,o}$ modes.

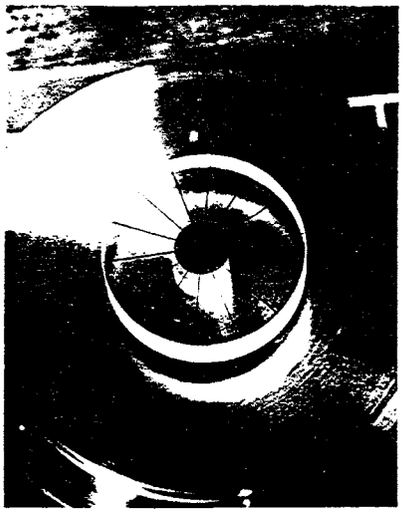


Fig. 5 - *Metallic lines of the WG mode resonator*

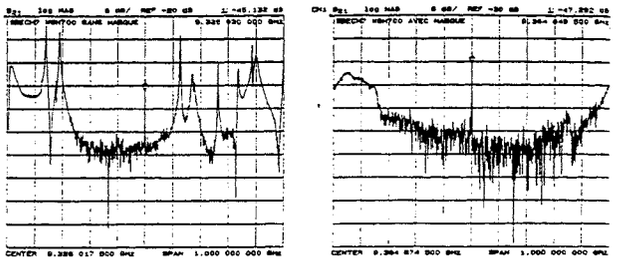


Fig. 6 - *Efficiency of the mode selection*

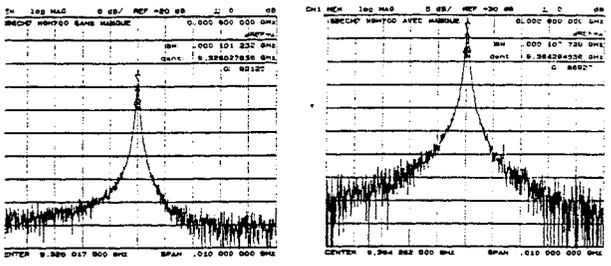


Fig. 7 - *Effect on the Q-factor*

Figures 6 and 7 show the effect of the metallic lines respectively on the spurious modes and on the Q-factor of the $WGH_{7,0,0}$ mode. Very low coupling ($\beta < 0,1$) was used to insure a direct determination of the unloaded Q-factor. We observe a small degradation of Q (about 5%). We plan to test other metallic line dimensions and other materials in order to improve this primilinary result.

CONCLUSION

In this paper we summarized our first results obtained with Whispering Gallery mode resonators. These works are a first step in the development of a high spectral purity microwave source at the LPMO.

We have implemented an original mode selection technique which satisfactory works for room temperature WG mode resonator. Further experiments are needed to improve the Q-factor limitation.

ACKNOWLEDGMENTS

We are grateful to CNES and BNM for their financial supports. The authors are indebted to P. Berthelot and J.C. Bauduy for their technical contributions to the elaboration of the resonators.

REFERENCES

[1] "Microwave frequency discriminator with a cooled sapphire resonator for ultra-low phase noise". D.G. Santiago and Q.J. Dick. Proceedings of the 1992 IEEE Frequency Control Symposium

[2] "Low noise microwave oscillators based on high-Q-temperature stabilized sapphire resonators". M.E. Tobar, E.N. Ivanov, R.A. Woode, J.H. Searls. Proceedings of the 1994 IEEE Frequency Control Symposium, pp. 433-440.

[3] "Microwave Oscillators Incorporating Cryogenic Sapphire Dielectric Resonators". R.C. Taber and C.A. Flory. IEEE Trans on UFFC, Vol 42, n°1, Jan 1995, pp. 111-119.

[4] "Approximate Approach to the Design of shielded Dielectric Disk Resonators with Whispering Gallery Modes". E.N. Ivanov, D.G. Blair, V.I. Kalinichev. IEEE Trans on MTT, Vol 41, n°4, April 1993, pp. 632-638.

[5] "Study of Whispering Gallery Modes in Anisotropic Single-Cristal Dielectric Resonators". J.Krupka, D. Cros, M. Aubourg and P. Guillon. IEEE Trans on MTT, Vol 42, n°1, Janv. 1994, p. 56-61.

[6] "Analysis and Evaluation of a Method of Measuring the complex Permittivity and Permeability of Microwave Insulators". W.E. Courtney. IEEE Trans on MTT, Vol 18, n°8, Aug. 1970, pp. 476-485.

[7] "Microwave Measurement of Dielectric Properties of Low-Loss Materials by the Dielectric Rod Resonator Method". Y. Kobayashi and M. Katoh. IEEE Trans on MTT, Vol 33, n°7, July 1985, pp. 586-592.

[8] "Foundations for Microwave Engineering". R.E. Collin, Mc. Graw-Hill, pp. 327-329.

[9] "TM_{ono} and TM_{m10} Mode Oversized Cylindrical Cavity Power Combiners". S. Nogi, K. Fukui. IEEE Trans on MTT, Vol 35, n°9, Sept. 1987, pp. 835-842.

FINITE ELEMENT MODELLING OF INFRARED-SENSITIVE RESONATORS USED AS THERMAL SENSORS

M.R.HAMROUR, B.DULMET AND S.GALLIOU

Laboratoire de Chronométrie, Electronique et Piezoelectricité
Ecole Nationale Superieure de Mecanique et des Microtechniques
Chemin de l'Epitaphe, 25030 Besançon Cedex - FRANCE

ABSTRACT

It is known that quartz absorbs IR radiations, especially in the 5-7.7 µm band. So, it is interesting to investigate the usability of quartz resonators as thermal sensors. In this paper, we present a reasonable approach to model the sensitivity of quasi-thickness vibrations to a quasi-static bias created by an incident infra-red beam.
Both permanent and transient thermal cases are considered. The temperature maps are computed with the help of a Finite Element Method package. The corresponding thermoelastic displacement gradients and stress fields are computed in a second step. Thereafter, the final frequency shifts are obtained in the framework of a general perturbation procedure for small dynamic fields superimposed on a quasi-static bias previously proposed by Tiersten and coworkers. We especially point out the respective contributions of thermally-induced stress and direct sensitivity to non homogeneous temperature maps in the final frequency shift. It appears that they are not the same, depending on the character of the bias (either transient or permanent state) and on the considered mode.

INTRODUCTION

In this work, we consider that the IR sensitivity of the resonant frequency of a quasi-thickness mode in a quartz plate of arbitrary orientation is the cumulative result of two effects :

- The first one can be defined as the temperature sensitivity at zero stress. To estimate it, it is necessary to get the actual temperature map in the whole volume of the resonator.

If the temperature distribution in the vibrating area is not homogeneous, a first frequency shift can be computed with help of a pertubation volume integral [1] from the knowledge of the temperature map and either from the temperature derivatives of fundamental elastic constants, as defined in the reference paper by Sinha and Tiersten [2], or from the temperature derivatives of so called effective elastic constants referred to the Lagrangian configuration [3].

- The second effect is the thermally-induced stress sensitivity. To estimate it, it is necessary to find the quasi-static thermoelastic stresses that are induced by the previously calculated temperature maps.

The total frequency shift can be computed with help of the previously quoted pertubation method and requires the knowledge of third order non linear elastic constants.

The final frequency shift is obtained on summing the two effects, indeed. In this approach, there is absolutely no difference between the cases of permanent and transient thermal states, except that temperature maps may strongly differ.
As a matter of fact, we propose to distinguish the direct temperature sensitivity from the thermoelastic stress sensitivity to keep the thinking of the problem on a more or less intuitive ground : although both effects are closely mingled in a general thermoelastic theory, we found interesting to isolate two kinds of terms in a general expression which agrees with Tiersten's theory. This is made in the first section of the paper.

The maps of temperatures and thermoelastic stresses have been obtained with help of SYSTUS finite element program, commonly used at LCEP for general purposes.
After that, the pertubation volume integral has been roughly estimated knowing any quantity of interest in the central region where the resonant modes exhibit their maximum amplitude. A first modelling has been made for a previously presented experiment [4], concerning an SC-cut plano-convex disk resonator with annular electrodes. Corresponding results are presented in the second section in which we have attempted to realistically model the actual experiment.
The third section proposes a modelling of a simple design consisting of a rectangular SC-cut plate in a cantilever mounting.
In this case, the temperature maps might have been analytically obtained. We have performed the analysis for several orientations of the rectangular plate around its own normal. And we have checked that the maximum frequency sensitivity with respect to induced stresses remains extremely low. This modelling is essentially proposed to illustrate how it is possible to investigate a whole bunch of identical designs with different crystalline orientations in a single automated analysis, provided that the FEM package provides with tools to automatically control a succession of runs. The lack of an integrated software tool to transparently chain the evaluation of static maps to the subsequent evaluation of the pertubation integral was partially overcome with help of Maple V mathematics program.

I. THERMOELASTIC APPROACH

It has been shown in a reference paper [5] that linearized equations for an infinitesimally small dynamic field superimposed on a bias can be written as :

$$\left.\begin{aligned} \tilde{K}_{L\gamma,L} &= \rho_0 \ddot{u}_\gamma \\ \tilde{K}_{L\gamma} &= G_{L\gamma Mv}.u_{v,M} \end{aligned}\right\} \quad (1)$$

where $\tilde{K}_{L\gamma}$ is the dynamic increment of the asymetric Piola-Kirchhoff stress tensor ($\tilde{K}_{L\gamma} \neq \tilde{K}_{\gamma L}$).

The G coefficients are called "effective" elastic coefficients because they do depend on the bias. In this kind of theory, all fields of interest are mapped onto the coordinates of a known reference state which is defined as the stress free "natural" static state of the body prior applying the bias to it. It has been demonstrated that the effective elastic coefficients can be expressed in terms of fundamental elastic constants and of the static strains / stresses [5] The following expression is valid if we retain only linear terms in the bias itself :

$$\left.\begin{aligned} G_{L\gamma Mv} &= {}_2C_{L\gamma Mv} + {}_2\hat{C}_{L\gamma Mv} \\ \hat{C}_{L\gamma Mv} &= T^1_{LM}\delta_{\gamma v} + {}_3C_{L\gamma MvAB}E^1_{AB} \\ &\quad + {}_2C_{L\gamma MK}w_{v,K} + {}_2C_{LKMv}w_{\gamma,K} \end{aligned}\right\} \quad (2)$$

where T^1 and E^1 respectively denote the static stress and strain tensors associated to the bias, while w is the static displacement. It is worth observing that G reduces to ${}_2C$ when the bias vanishes.

In a purely elastic problem, ${}_2C$ and ${}_3C$ tensors are true constants in a given reference frame. In a thermoelastic problem, ${}_2C$ and ${}_3C$ depend on temperature and we must consider the additional formula :

$$T^1_{LM} = C_{LMRS}(E^1_{RS} - \alpha_{RS}\Delta\Theta) \quad (3)$$

if we stick to a first order expansion in terms of temperature. Here, α indicates thermal expansion tensor. Retaining only the first order expansion coefficient α is consistent with the fact that the expansion (2) retains only linear terms in terms of static strains and stresses. By direct substitution of (3) into (2), we obtain the specific expression of the variation of the effective elastic constants due to a thermoelastic bias :

$$\begin{aligned} \Delta\left(G_{L\gamma M\varepsilon}\right) = \Delta\left({}_2C_{L\gamma M\varepsilon}\right) &+ C^0_{LMRS}\cdot(E^1_{RS} - \alpha_{RS}\Delta\Theta)\cdot\delta_{\gamma\varepsilon} \\ &+ {}_3C^0_{L\gamma M\varepsilon AB}\cdot E^1_{AB} + {}_2C^0_{L\gamma KM}\cdot w_{\varepsilon,K} \\ &+ {}_2C^0_{LKM\gamma}w_{\varepsilon,K} \end{aligned} \quad (4)$$

where the nought superscript indicates that we take the value of the corresponding quantity in the natural state, and δ stands for Kronecker's symbol. All quantities without this superscript are proportional to $\Delta\Theta$

In the present paper, we solve heat diffusion equation with help of SYSTUS finite element program. For the considered class of problems, we can expect non homogeneous temperature distributions in the resonator. Very classically, the finite element program can also be used to get associated thermoeleastic stresses, strains, and displacements, in a subsequent run. We can then determine static quantities E_{RS}, $\Delta\Theta$ and w appearing in Eq (4).

In addition, $\Delta\left({}_2C_{L\gamma M\varepsilon}\right) \equiv T^{(1)}_2 C_{L\gamma M\varepsilon} \times C^o_{L\gamma M\varepsilon} \times \Delta\Theta$ (without summing over repeated indexes), where ${}^{(1)}{}_2C$ is exactly the set of first order temperature derivatives of fundamental elastic constants in the sense of Sinha and Tiersten. By substitution of every known quantity into (4), we get the local value of ΔG, which will be obviously non homogeneous for the considered class of problems. Hence, a straightforward manner to get the corresponding frequency shift has been also established by Tiersten in well known papers, such as [1]. We can do it consistently by using Eq.(4) in the above mentioned form. Nevertheless, before calling the perturbation procedure, we found more enlightening to rearrange somewhat the various terms appearing in Eq. 4.

In this view, we introduce the following quantities :

$$\left.\begin{aligned} \eta^1_{RS} &= E^1_{RS} - \alpha_{RS}\,\Delta\Theta \\ v_{\varepsilon S} &= w_{\varepsilon,S} - \delta_{\varepsilon R}\,\alpha_{RS}\,\Delta\Theta \end{aligned}\right\} \quad (5)$$

η^1 is the part of the static strain which creates a non-zero stress, according to (3). It is also directly given by SYSTUS after the following process has been performed: the program first calculates the whole static displacement and the static stresses. Then, it computes the corresponding strains upon inverting Hooke's anisotropic law.

Correspondingly, v has the same physical meaning for displacement gradients instead of strains and we might as well consider the perfectly symmetric expressions :

$$\left.\begin{aligned} \eta^1_{RS} &\approx \tfrac{1}{2}\left(\delta_{\varepsilon R}v_{\varepsilon S} + \delta_{S\sigma}v_{\sigma R}\right) + \ldots \\ E^1_{RS} &\approx \tfrac{1}{2}\left(\delta_{\varepsilon R}w_{\varepsilon,S} + \delta_{S\sigma}w_{\sigma,R}\right) + \ldots \end{aligned}\right\} \quad (6)$$

if we abandon the well known intrinsically non linear part of the strain coming from the products of displacement gradients.

Then, we can split ΔG into two separate parts with distinct physical meanings :

$$\left.\begin{aligned} \Delta\left(G_{L\gamma M\varepsilon}\right) &= {}_1\Delta + {}_2\Delta \\ {}_1\Delta &= \Delta\left({}_2C_{L\gamma M\varepsilon}\right) - \left[{}_3C^0_{L\gamma M\varepsilon AB}\alpha_{AB} \right. \\ &\quad \left. + \left({}_2C^0_{L\gamma MS}\delta_{\varepsilon R} + {}_2C^0_{LSM\varepsilon}\delta_{\gamma R}\right)\alpha_{RS}\right]\Delta\Theta \\ {}_2\Delta &= {}_3C^0_{L\gamma M\varepsilon AB}\eta^1_{AB} + {}_2C^0_{L\gamma MS}v_{\varepsilon S} + {}_2C^0_{LSM\varepsilon}v_{\gamma S} \end{aligned}\right\} \quad (7)$$

where we took advantage of the symmetry of α to factor α_{RS} in ${}_1\Delta$. Here, ${}_1\Delta$ is identical to the ΔG variation

$\Delta G_{L\gamma M\varepsilon}=T^1G_{L\gamma M\varepsilon}\times G_{L\gamma M\varepsilon}\times\Delta\Theta$ where T^1G is exactly the set of first order temperature derivatives of "effective" elastic constants referred to the natural state, in the sense of Ref[2]. The interest of using T^nG temperature coefficients lies in the fact that they can be directly computed up to the third order from the knowledge of frequency-temperature characteristics of resonators submitted to slow and supposedly homogeneous temperature variations.

Now, we can compute the global frequency shift according to Tiersten's perturbation formula :

$$\omega-\omega_0\approx\frac{1}{2\omega_0}\frac{\int_{V_0}Q_jQ_l\Delta(G_{2j2l})(u^0_{,2})^2dV}{\int_{V_0}\rho_0\left(u^0\right)^2dV} \qquad (8)$$

where Q is the normalized vector holding the polarizations of the considered thickness-mode (A or B or C) and u is the high frequency dynamic displacement. We can extend the computation of ${}_1\Delta$ in (8) with the third order formula :

$${}_1\Delta=\sum_{n=1}^{3}T^{(n)}G_{L\gamma M\varepsilon}\,\Delta\Theta^n \qquad (9)$$

instead of (7b), the explicit extension of which up to the third order would require high order fundamental constants, not available at this time. For symmetrical reasons, higher order expansion of ${}_2\Delta$ is not possible. Since direct temperature sensitivity can be relatively high and the correponding cubic behaviour is very neat for well known temperature-compensated modes and orientations, while the stress-induced sentivity is known to be generally small in compare with the former, we state that it does make sense to mix third order expansion for ${}_1\Delta$ and first order expansion for ${}_2\Delta$ at least from the practical point of view, if not the theoretical one, indeed.

II. FIRST ANALYSIS

First, let us briefly recall the experiment and results previously described in Ref[4]. The sensitive device is a standard plano-convex thickness-shear resonator, operating at 10 MHz on its C mode, third overtone. One of its major surfaces supports a deposited ring-shaped electrode whereas the other major surface is coated with a simple circular electrode. Ni-Cr strips support the resonator providing a low thermal conductance (cf Fig.1). The metal holder of the device is pierced in such a way that the face with the annular electrode is exposed to the incident beam. Quartz is a good absorber between 5 µm and 7.7 µm, and still absorbs up to 70% of the incident energy for wavelength upper than 7.7 µm, up to 14 µm. So, the uncoated central area of the quartz resonator works in a direct absorption mode whereas the goldened part acts as a reflector (in the infra-red band, gold reflects 98% of the incident energy) .
Experiments have been performed under vacuum in order to guarantee a high-quality vibration of the quartz resonator and to cancel parasitic absorptions of the air. Figs2-3 show measurements obtained for the B-mode and the C-mode respectively, when the resonator is submitted to a pulse of I.R. flux : the emitted beam of a large spectral irradiance source is chopped and collimated towards the uncoated central area of the resonator that is where the vibration is trapped (the plano-convex zinc selenide lens used in this experiment is a band-pass within [2µm,12µm]). In fact, the measured frequency is the output frequency of a mixer of which both inputs are the output frequency of the oscillator associated to the detector-quartz in one hand, and the output frequency of the oscillator using a reference-quartz in the other hand. The latter is obviously identical to the detector one and operates in the same operating conditions excepted for the incident flux. This beat technique must cancel the ambient temperature effects. The operating temperature is about 27°C where the static frequency -temperature coefficients of the B-mode and C-mode are reduced to the first order ones : $\Delta f_B=a_B.\Delta\theta$ and $\Delta f_C=a_C.\Delta\theta$ with $a_B\approx300$ Hz/°C and $a_C\approx10$ Hz/°C.

To model this actual experiment, we followed the theoretical approach described in the introductory section of this paper. Since the purely thermal analysis occuring in the first step of the method only requires one degree of freedom by node, it is possible to perform very refined meshes. Nevertheless, since the results of this kind of analysis should be reused in the thermoelastic FEM analysis of the second step, which requires three degrees of freedom by node (the three components of mechanical displacement), it is in fact necessary to use the same mesh for both FEM steps. We present the mesh that we used at Fig 4.

As a matter of fact, we simplified the evaluation of the perturbation integral (8) occuring in the final step of the proposed approach : since we have obtained relatively homogeneous temperature and thermoelastic stress maps in the central region, we took advantage of it to reduce the computation to a simple unidimensional integration along the thickness in this region, where trapped modes exhibit their maximum amplitude.

In this way, we obtained the total frequency shifts simulations of Figs 5-6, which are in accordance with corresponding experimental results of Figs 2-3 when quartz temperature increases under the incident flux effect. Nevertheless, the observed behavior during the decreasing period of the temperature when the incident flux is turned off, was not expected from the simulation. This defect might be due to a non perfect compensation of the external perturbations by the reference oscillator.

In addition, we have found that the direct temperature sensitivity coming from ${}_1\Delta$ plays the most significant part in the total frequency shift. The part of the total frequency shift created coming from ${}_2\Delta$ is at maximum 10-15% in the case of the C mode, which is thermally compensated, and hence would never be the preferred mode to detect IR flux. Predicted influence of ${}_2\Delta$ terms is nearly negligible (less than 1%) in the case of B mode.

III. SECOND ANALYSIS

In the previous case, obtained temperature maps are only slightly dependent on anisotropy of thermal conductivity. Since the incident thermal flux is applied on a circular region, we obtained isolines of temperature wich are almost circular in the central region. Hence, the thermally-induced stresses in this region are roughly isotropic with respect to the direction of observation around the plate normal. Thus, it is not surprising to get a small effect of thermoelastic sresses on the global frequency shift since SC-cut is known to minimize the effect of a nearly isotropic contribution of plane stresses on the frequency. Thus, we found interesting (from the academic point of view) to model a configuration where the static stresses are neatly oriented in the SC-cut plane. The most simple stucture is given in Fig.7. It consists of a square plate clamped at one edge and heated on the opposite edge. Then we get a roughly constant thermal gradient from the heated edge to the clamped one, and we found enlightening to model the part of the global frequency shift which comes from thermally-induced stresses as a function of azimuthal orientation of the square plate in the SC-cut plane.

We followed the same thermoelastic approach applied on the plano-convex structure. In this view, we performed a single automated analysis program.

We found that the thermally-induced effect due to thermoelastic bias, for an infrared flux about 130 µW, reaches its maximum and minimum respectively, when the azimuthal angle is about 70° and 0°, as is shown on Fig.8.

We plan to use such integrated approach to investigate other cut orientations and sensor configurations.

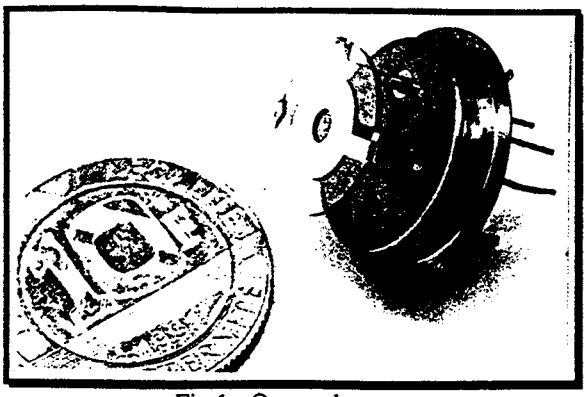


Fig.1 - Quartz detector

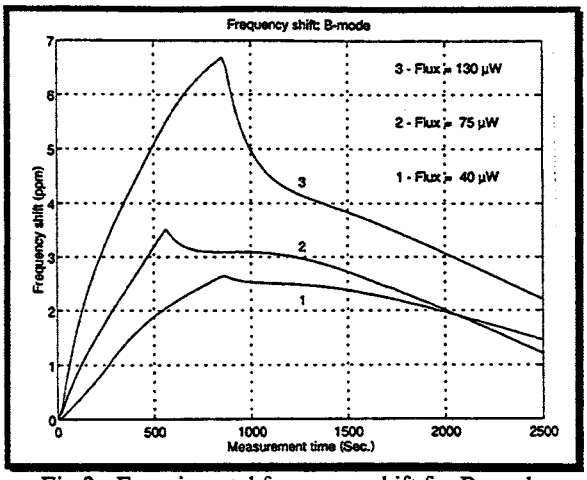


Fig.2 - Experimental frequency shift for B-mode.

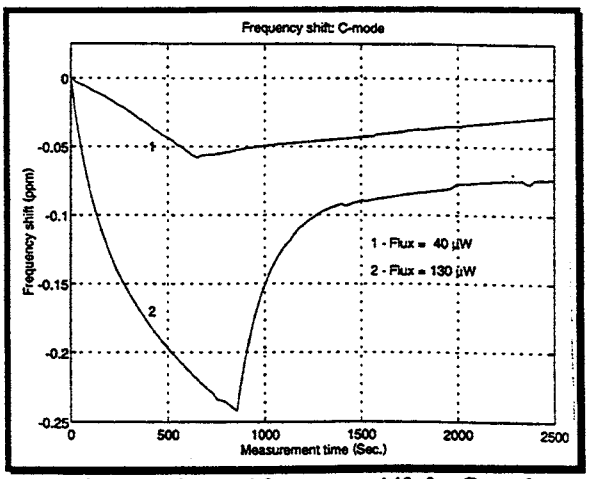


Fig.3 - Experimental frequency shift for C-mode.

REFERENCES

[1] - H.F.TIERSTEN.
"Perturbation theory for linear electrostatic equations for small field superposed on a bias", J. Acoust. Soc. Am.64(3) (1978).

[2] - B.K. SINHA AND H.F TIERSTEN.
"First temperature derivatives of the fundamental elastic constants of quartz", J.Appl. Phys. 50 (4), April 1979, pp 2732-2739.

[3] - B. DULMET AND R.BOURQUIN.
"Influence of the definition of material, constants in the study of crystal plates thermal behavior", Proc. of 3rd EFTF, pp 305-312 - (1989).

[4] - M.R.HAMROUR AND S. GALLIOU.
"Analysis of the infrared sensitivity of a quartz resonator - Application as a thermal sensor ", IEEE 1994 Ultrasonics Symposium.

[5] - J.C. BAUMHAUER AND H.F. TIERSTEN.
"Non linear Electrostatic Equations for Small Ficlds Superimposed on a Bias", J. Acoust. Soc. Am. 54,1017 (1973) .

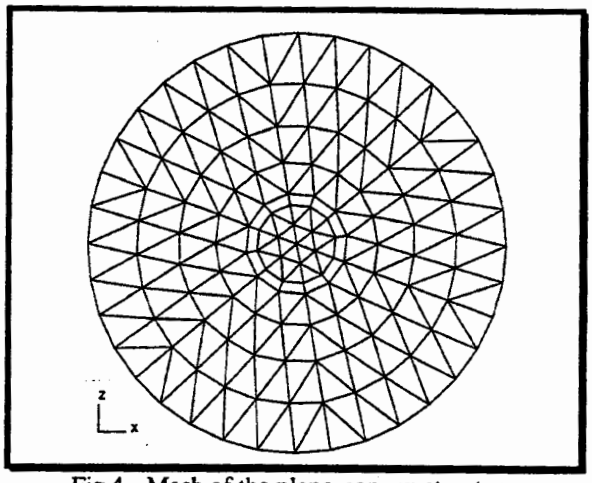


Fig.4 - Mesh of the plano-convex structure.

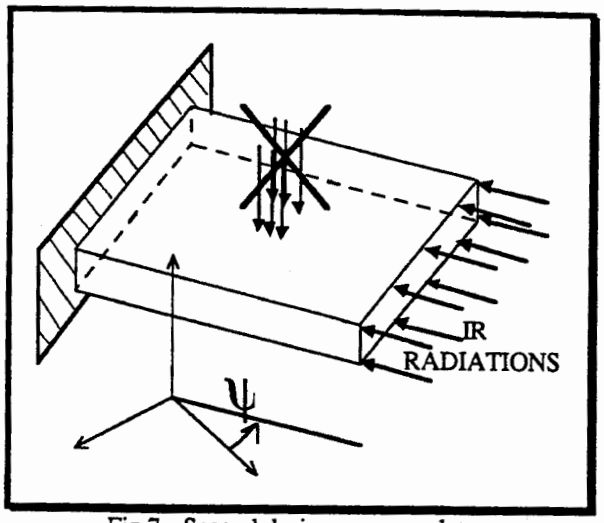


Fig.7 - Second design : square plate.

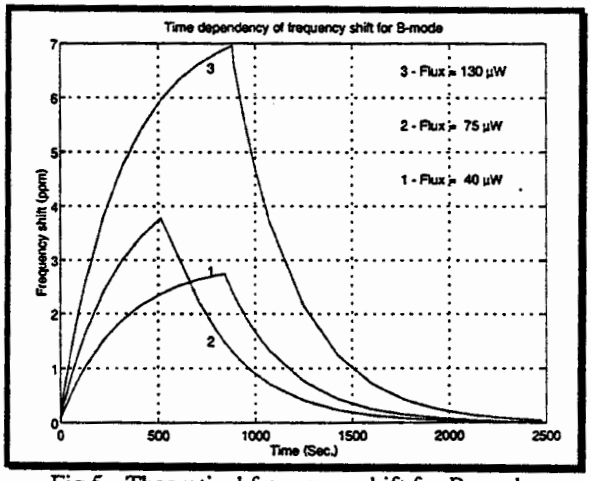


Fig.5 - Theoretical frequency shift for B-mode.

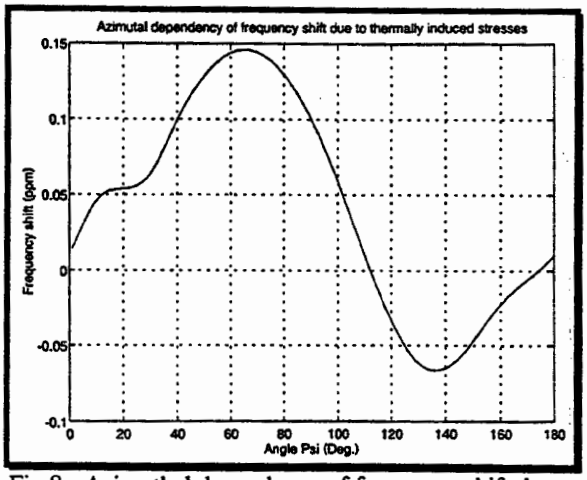


Fig.8 - Azimuthal dependency of frequency shift due to thermally induced stresses

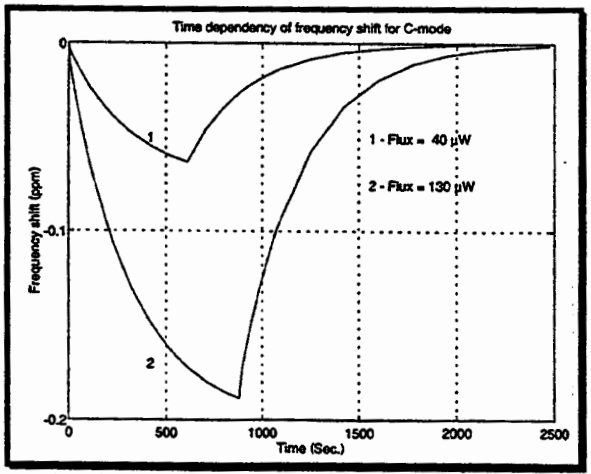


Fig.6 - Theoretical frequency shift for C-mode.

FREQUENCY SPECTRUM AND FREQUENCY-TEMPERATURE CHARACTERISTICS OF THIRD OVERTONE AT AND BT-CUT QUARTZ RESONATORS

J. Zelenka

Technical University in Liberec, Electrical Engineering Department,
Hálkova 6, 461 17 Liberec, Czech Republic. Tel: 42-48-25441, Fax: 42-48-27383

ABSTRACT

The linear elastic equations for a small field superposed on a thermally induced deformation by steady and uniform temperature changes are used for the expression of the influence of the stiffness and inertia of electrodes on the frequency-temperature characteristic of third harmonic of the thickness-shear vibration of Y-cut quartz resonators. The two dimensional equations of motion derived by Lee and Nikodem and based on the trigonometric series expansion are used for the description of the vibration of fully plated strip bounded in digonal axis direction.

1. INTRODUCTION

The influence of electrodes deposited on the surface of quartz plates on the frequency-temperature of the plate was studied by Tiersten and Sinha [1], Stevens and Tiersten [2] and Sherman [3] during last eight years. In two previous papers [4] and [5] the results of the study of the elastic stiffness and inertia of electrodes on the frequency-temperature characteristic of the fundamental thickness shear mode of the vibration of AT-cut quartz plates were also described. The two dimensional piezoelectric plate equations based on the Mindlin's power series expansion method were used by the derivation of the equations for the plated crystal plate. The linear field equations for small vibrations superposed on thermally induced deformations described by Lee and Yong [6] were used by the derivation of the influence of the electrodes on the frequency-temperature characteristics.

In the present paper the linear elastic equations for small field superposed on thermally induced deformation by steady and uniform temperature changes are used for the expression of the influence of the stifness and inertia of electrodes on the frequency-temperature characteristic of third harmonic of the thickness-shear vibration of Y-cut quartz resonators. The two dimensional equations of motion derived by Lee and Nikodem [7], [8] and based on the trigonometric series expansion are used for the description of the vibration of fully plated strip bounded in digonal axis direction. The coupled modes of vibration of Y-cut quartz plates including flexure, face-shear, first and third harmonic thickness-shear in the direction of the digonal axis, second thickness-shear in the direction perpendicular to the digonal axis and second thickness stretch in the direction of thickness of the plate are considered by the solving of the problem. The Mindlin's [9] reduction of the number of stress-equation of motion is used by the derivation.

2. EQUATIONS OF INCREMENTAL MOTION FOR THE UNPLATED CRYSTAL PLATES

We consider Y-cut quartz strip bounded in the thickness direction by free faces at $x_2 = \pm b$ and in the width direction by free edges at $x_1 = \pm a$ (x_1 axis is digonal axis). When we insert the two dimensional equations of motion derived by Lee, Syngellakis and Hou [10] into the equations derived by Lee and Yong [6] and described the small amplitude waves superposed on homogenious thermal strain we receive the incremental stress-strain-temperature equations of motion. When the piezoelectric properties of the plate are neglected we obtain

$$\beta_{ik} t^{(n)}_{jk,j} - \frac{n\pi}{2b}\beta_{ik}\bar{t}^{(n)}_{2k} + \beta_{ik} F^{(n)}_k = (1+\delta_{n0})\varrho \ddot{u}^{(n)}_i, \qquad (1)$$

where the components of the n-th order of the incremental stress $t^{(n)}_{ij}$ and $\bar{t}^{(n)}_{ij}$ and the incremental face traction $F^{(n)}_j$ are

$$t^{(n)}_{ij} = (1+\delta_{n0}) D_{ijkl} e^{(n)}_{kl} + \sum_{m=1}^{\infty} B_{mn} D_{ijkl} \bar{e}^{(n)}_{kl},$$

$$\bar{t}^{(n)}_{ij} = D_{ijkl}\bar{e}^{(n)}_{kl} + \sum_{m=0}^{\infty} B_{nm} D_{ijkl} e^{(n)}_{kl},$$

$$F^{(n)}_j = t_{2j}(b) - (-1)^n t_{2j}(-b). \qquad (2)$$

The n-th order incremental strain components $e^{(n)}_{ij}$ and $\bar{e}^{(n)}_{ij}$ are defined by

$$e^{(n)}_{ij} = \frac{1}{2}(\beta_{kj} u^{(n)}_{k,i} + \beta_{ki} u^{(n)}_{k,j}),$$

$$\bar{e}^{(n)}_{ij} = \frac{n\pi}{4b}(\delta_{2i}\beta_{jk} u^{(n)}_k + \delta_{2j}\beta_{ik} u^{(n)}_k). \qquad (3)$$

In the relations (1), (2) and (3) were introduced

$$\beta_{ik} = \delta_{ik} + \alpha^{\Theta}_{ik},$$

$$D_{ijkl} = C_{ijkl} + D^{(1)}_{ijkl}\Theta + D^{(2)}_{ijkl}\Theta^2 + D^{(3)}_{ijkl}\Theta^3, \qquad (4)$$

where the linear thermal expansion coefficients α^{Θ}_{ij} are defined by

$$\alpha^{\Theta}_{ij} = \alpha^{(1)}_{ij}\Theta + \alpha^{(2)}_{ij}\Theta^2 + \alpha^{(3)}_{ij}\Theta^3 \qquad (5)$$

and $\alpha^{(n)}_{ij}$ are nth order expansion temperature coefficients.

The temperature derivatives $D^{(n)}_{ijkl}$ are given by

$$D^{(1)}_{ijkl} = C^{(1)}_{ijkl} + C_{ijklmn}\alpha^{(1)}_{mn},$$

$$D^{(2)}_{ijkl} = \frac{1}{2}\tilde{C}^{(2)}_{ijkl} + C_{ijklmn}\alpha^{(2)}_{mn},$$

$$D^{(3)}_{ijkl} = \frac{1}{6}\tilde{C}^{(3)}_{ijkl} + C_{ijklmn}\alpha^{(3)}_{mn} \qquad (6)$$

and $C^{(1)}_{ijkl}$, $\tilde{C}^{(2)}_{ijkl}$, $\tilde{C}^{(2)}_{ijkl}$ are first-order temperature derivative, effective second- and third-order temperature derivatives of elastic stiffnesses C_{ijkl}.

Following Mindlin's [10] reduction procedure we will consider the modes of vibration which are to be straight-crested along x_3 and antisymmetric with respect to both x_1 and x_2 axis. From the expansion of the displacement u_j $(j = 1,2,3)$ in series of trigonometric functions of the thickness-coordinate, x_2,

$$u_j = \sum_{n=0}^{\infty} u_j^{(n)}(x_1, x_3, t)\cos(\frac{n\pi}{2}(1 - \frac{x_2}{b})) \qquad (7)$$

we take only

$$\begin{aligned}
u_1 &= [(u_1^{(1)}\cos(\frac{\pi}{2}(1-\frac{x_2}{b})) + u_1^{(3)}\cos(\frac{3\pi}{2}(1-\frac{x_2}{b}))]\ e^{i\omega t},\\
u_2 &= [u_2^{(0)} + u_2^{(2)}\cos(\pi(1-\frac{x_2}{b}))]\ e^{i\omega t},\\
u_3 &= [u_3^{(0)} + u_3^{(2)}\cos(\pi(1-\frac{x_2}{b}))]\ e^{i\omega t}, \qquad (8)
\end{aligned}$$

in which the displacement $u_j^{(n)}$ depends only on x_1.

From the incremental stress-strain-temperature equation of motion (1) remain

$$\begin{aligned}
&\beta_{22}t^{(0)}_{12,1} + \frac{1}{b}\beta_{22}F_2^{(0)} + \frac{1}{b}\beta_{23}F_3^{(0)} + 2\varrho\omega^2 u_2^{(0)} = 0,\\
&\beta_{23}t^{(0)}_{13,1} + \frac{1}{b}\beta_{23}F_2^{(0)} + \frac{1}{b}\beta_{33}F_3^{(0)} + 2\varrho\omega^2 u_3^{(0)} = 0,\\
&\beta_{11}t^{(1)}_{11,1} - \frac{\pi}{2b}\beta_{11}\bar{t}^{(1)}_{12} + \frac{1}{b}\beta_{11}F_1^{(1)} + \varrho\omega^2 u_1^{(1)} = 0,\\
&\beta_{22}t^{(2)}_{12,1} - \frac{\pi}{b}\beta_{22}\bar{t}^{(2)}_{22} + \frac{1}{b}\beta_{22}F_2^{(2)} + \frac{1}{b}\beta_{23}F_3^{(2)} + \varrho\omega^2 u_2^{(2)} = 0,\\
&\beta_{23}t^{(2)}_{13,1} - \frac{\pi}{b}\beta_{23}\bar{t}^{(2)}_{23} + \frac{1}{b}\beta_{23}F_2^{(2)} + \frac{1}{b}\beta_{33}F_3^{(2)} + \varrho\omega^2 u_3^{(2)} = 0,\\
&\beta_{11}t^{(3)}_{11,1} - \frac{3\pi}{2b}\beta_{11}\bar{t}^{(3)}_{12} + \frac{1}{b}\beta_{11}F_1^{(3)} + \varrho\omega^2 u_1^{(3)} = 0 \qquad (9)
\end{aligned}$$

and the only non-zero components of incremental strain are, from (3) and (8) and for $\beta_{12} = \beta_{21} = \beta_{13} = \beta_{31} = 0$

$$\begin{aligned}
e^{(0)}_{12} &= \frac{1}{2}(\beta_{22}u^{(0)}_{2,1} + \beta_{23}u^{(0)}_{3,1}),\\
e^{(0)}_{13} &= \frac{1}{2}(\beta_{23}u^{(0)}_{2,1} + \beta_{33}u^{(0)}_{3,1}),\\
e^{(1)}_{11} &= \beta_{11}u^{(1)}_{1,1},\\
e^{(2)}_{12} &= \frac{1}{2}(\beta_{22}u^{(2)}_{2,1} + \beta_{23}u^{(2)}_{3,1}),\\
e^{(2)}_{13} &= \frac{1}{2}(\beta_{23}u^{(2)}_{2,1} + \beta_{33}u^{(2)}_{3,1}),\\
e^{(3)}_{11} &= \beta_{11}u^{(3)}_{1,1},\\
\bar{e}^{(1)}_{12} &= \frac{\pi}{4b}\beta_{11}u_1^{(1)},\\
\bar{e}^{(2)}_{22} &= \frac{\pi}{b}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}),\\
\bar{e}^{(2)}_{23} &= \frac{\pi}{2b}(\beta_{32}u_2^{(2)} + \beta_{33}u_3^{(2)})'\\
\bar{e}^{(3)}_{12} &= \frac{3\pi}{4b}\beta_{11}u_1^{(3)}. \qquad (10)
\end{aligned}$$

After introducing a low frequency correction factor κ_1 and a high frequency correction factor κ_2 and using the abbreviated notation of the elastic stiffnesses $D_{\lambda\mu}$ the surving incremental stress-displacement relation are obtained from (2) and (10)

$$\begin{aligned}
t^{(0)}_{12} &= 2[D_{56}(\beta_{23}u^{(0)}_{2,1} + \beta_{33}u^{(0)}_{3,1}) + D_{66}(\beta_{22}u^{(0)}_{2,1} + \beta_{32}u^{(0)}_{3,1})] +\\
&+ \frac{2}{b}D_{56}\beta_{11}(\kappa_1 u_1^{(1)} + u_1^{(3)}),\\
t^{(0)}_{13} &= 2[D_{55}(\beta_{23}u^{(0)}_{2,1} + \beta_{33}u^{(0)}_{3,1}) + D_{56}(\beta_{22}u^{(0)}_{2,1} + \beta_{32}u^{(0)}_{3,1})] +\\
&+ \frac{2}{b}D_{56}\beta_{11}(\kappa_1 u_1^{(1)} + u_1^{(3)}),\\
t^{(1)}_{11} &= D_{11}\beta_{11}u^{(1)}_{1,1} + \frac{4}{3b}(D_{12}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}) +\\
&+ D_{14}(\beta_{23}u_2^{(2)} + \beta_{33}u_3^{(2)}),\\
t^{(2)}_{12} &= D_{56}(\beta_{23}u^{(2)}_{2,1} + \beta_{33}u^{(2)}_{3,1}) + D_{66}(\beta_{22}u^{(2)}_{2,1} + \beta_{32}u^{(2)}_{3,1}) -\\
&- D_{66}\beta_{11}(\frac{2}{3b}u_1^{(1)} - \frac{18}{5b}u_1^{(3)}),\\
t^{(2)}_{13} &= D_{55}(\beta_{23}u^{(2)}_{2,1} + \beta_{33}u^{(2)}_{3,1}) + D_{56}(\beta_{22}u^{(2)}_{2,1} + \beta_{32}u^{(2)}_{3,1}) -\\
&- D_{56}\beta_{11}(\frac{2}{3b}u_1^{(1)} - \frac{18}{5b}u_1^{(3)}),\\
t^{(3)}_{11} &= D_{11}\beta_{11}u^{(3)}_{1,1} - \frac{8}{5b}[D_{12}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}) +\\
&+ D_{14}(\beta_{23}u_2^{(2)} + \beta_{33}u_3^{(2)})],\\
\bar{t}^{(1)}_{12} &= \kappa_1\frac{4}{\pi}[D_{56}(\beta_{23}u^{(0)}_{2,1} + \beta_{33}u^{(0)}_{3,1}) + D_{66}(\beta_{22}u^{(0)}_{2,1} + \beta_{32}u^{(0)}_{3,1})] -\\
&- \frac{4}{3\pi}[D_{56}(\beta_{23}u^{(2)}_{2,1} + \beta_{33}u^{(2)}_{3,1}) + D_{66}(\beta_{22}u^{(2)}_{2,1} + \beta_{32}u^{(2)}_{3,1})] +\\
&+ \frac{\pi}{2b}D_{66}\beta_{11}u_1^{(1)},\\
\bar{t}^{(2)}_{22} &= \frac{8}{\pi}D_{12}\beta_{11}(\frac{1}{3}u^{(1)}_{1,1} - \frac{1}{5}u^{(3)}_{1,1}) + \frac{\pi}{b}[D_{22}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}) +\\
&+ D_{24}(\beta_{23}u_2^{(2)} + \beta_{33}u_3^{(2)})],\\
\bar{t}^{(2)}_{23} &= \frac{8}{\pi}D_{14}\beta_{11}(\frac{1}{3}u^{(1)}_{1,1} - \frac{1}{5}u^{(3)}_{1,1}) + \frac{\pi}{b}[D_{24}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}) +\\
&+ D_{44}(\beta_{23}u_2^{(2)} + \beta_{33}u_3^{(2)})],\\
\bar{t}^{(3)}_{12} &= \frac{4}{3\pi}[D_{56}(\beta_{23}u^{(0)}_{2,1} + \beta_{33}u^{(0)}_{3,1}) + D_{66}(\beta_{22}u^{(0)}_{2,1} + \beta_{32}u^{(0)}_{3,1}) +\\
&+ \frac{12}{5\pi}[D_{56}(\beta_{23}u^{(2)}_{2,1} + \beta_{33}u^{(2)}_{3,1}) + D_{66}(\beta_{22}u^{(2)}_{2,1} + \beta_{32}u^{(2)}_{3,1})] +\\
&+ \frac{3\pi}{2b}D_{66}\beta_{11}u_1^{(3)}, \qquad (11)
\end{aligned}$$

where

$$\kappa_1 = \frac{\pi}{\sqrt{8}}, \qquad \kappa_2 = 0.901.$$

The inctremental-temperature equations of motion of the unplated Y-cut quartz resonators are obtained by substitution of the incremental-stress-displacement relations (11) into the incremental-stress-strain-temperature equations of motion (9).

The edge conditions at $x_1 = \pm a$ are

$$t^{(0)}_{12} = t^{(0)}_{13} = t^{(1)}_{11} = t^{(2)}_{12} = t^{(2)}_{13} = t^{(3)}_{11}. \qquad (12)$$

3. DISPERSION RELATIONS

For the straight-crested waves propagating in the x_1 direction, omitting the factor $e^{i\omega t}$ we let

$$\begin{aligned}
u_2^{(0)} &= A_2^{(0)}\sin\xi x_1, \qquad u_2^{(2)} = A_2^{(2)}\sin\xi x_1,\\
u_3^{(0)} &= A_3^{(0)}\sin\xi x_1, \qquad u_3^{(2)} = A_3^{(2)}\sin\xi x_1,\\
u_1^{(1)} &= A_1^{(1)}\cos\xi x_1, \qquad u_1^{(3)} = A_1^{(3)}\cos\xi x_1, \qquad (13)
\end{aligned}$$

After substitution of (13) in (11) and the result in (9) we obtain a set of six simultaneously, linear algebraic equations in the six amplitudes $A_j^{(n)}$

$$\begin{bmatrix}
h_{11} & h_{12} & h_{13} & 0 & 0 & h_{16}\\
h_{12} & h_{22} & h_{23} & 0 & 0 & h_{26}\\
h_{13} & h_{23} & h_{33} & h_{34} & h_{35} & 0\\
0 & 0 & h_{34} & h_{44} & h_{45} & h_{46}\\
0 & 0 & h_{35} & h_{45} & h_{55} & h_{56}\\
h_{16} & h_{26} & 0 & h_{46} & h_{56} & h_{66}
\end{bmatrix}
\begin{bmatrix}
A_2^{(0)}\\ A_3^{(0)}\\ A_1^{(1)}\\ A_2^{(2)}\\ A_3^{(2)}\\ A_1^{(3)}
\end{bmatrix} = 0 \qquad (14)$$

The dimensionless and real coefficients h_{pq} are

$$
\begin{aligned}
h_{11} &= 2(\bar{D}_{55}\beta_4^2 + 2\bar{D}_{56}\beta_2\beta_4 + \bar{D}_{66}\beta_2^2)z^2 - \frac{2}{\kappa_2}\Omega^2,\\
h_{12} &= 2(\bar{D}_{55}\beta_3\beta_4 + \bar{D}_{5}6(\beta_2\beta_3 + \bar{D}_{66}\beta_2\beta_4)z^2,\\
h_{13} &= \kappa_1\frac{4}{\pi}\beta_1(\bar{D}_{56}\beta_4 + \bar{D}_{66}\beta_2)z^2,\\
h_{16} &= \frac{4}{\pi}\beta_1(\bar{D}_{56}\beta_4 + \bar{D}_{66}\beta_2)z^2,\\
h_{22} &= 2(\bar{D}_{55}\beta_3^2 + 2\bar{D}_{56}\beta_3\beta_4 + \bar{D}_{66}\beta_1\beta_4)z^2 - 2\Omega^2,\\
h_{23} &= \kappa_1\frac{4}{\pi}\beta_1(\bar{D}_{56}\beta_3 + \bar{D}_{66}\beta_4)z^2,\\
h_{26} &= \frac{4}{\pi}\beta_1(\bar{D}_{56}\beta_3 + \bar{D}_{66}\beta_4)z^2,\\
h_{33} &= (z^2\bar{D}_{11}\beta_1^2 + \bar{D}_{66}\beta_1^2 - \Omega^2)z^2,\\
h_{34} &= -\frac{4}{3\pi}[(\bar{D}_{56} + 4\bar{D}_{14})\beta_1\beta_4 + (\bar{D}_{66} + 4\bar{D}_{12})\beta_1\beta_2]z^2,\\
h_{35} &= -\frac{4}{3\pi}[(\bar{D}_{56} + 4\bar{D}_{14})\beta_1\beta_3 + (\bar{D}_{66} + 4\bar{D}_{12})\beta_1\beta_4]z^2,\\
h_{44} &= (\bar{D}_{66}\beta_2^2 + 2\bar{D}_{56}\beta_2\beta_4 + \bar{D}_{55}\beta_4^2)z^2 +\\
&+ 4(\bar{D}_{22}\beta_2^2 + 2\bar{D}_{24}\beta_2\beta_4 + \bar{D}_{44}\beta_4^2) - \Omega^2,\\
h_{45} &= (\bar{D}_{55}\beta_3\beta_4 + \bar{D}_{56}\beta_2(\beta_3 + \beta_4) + \bar{D}_{66}\beta_2\beta_4)z^2 +\\
&+ 4[\bar{D}_{22}\beta_2\beta_4 + \bar{D}_{24}(\beta_2\beta_3 + \beta_4^2) + \bar{D}_{44}\beta_3\beta_4],\\
h_{46} &= \frac{4}{5\pi}\beta_1[(4\bar{D}_{12} + 9\bar{D}_{66})\beta_2 + (4\bar{D}_{14} + 9\bar{D}_{56})\beta_4]z^2,\\
h_{55} &= (\bar{D}_{55})\beta_3^2 + 2\bar{D}_{56}\beta_3\beta_4 + \bar{D}_{6}6\beta_4^2)z^2 +\\
&+ 4(\bar{D}_{22}\beta_4^2 + \bar{D}_{24}\beta_3\beta_4 + \bar{D}_{44}\beta_3^2) - \Omega^2,\\
h_{56} &= \frac{4}{5\pi}[(4\bar{D}_{12} + 9\bar{D}_{66})\beta_1\beta_4 + (4\bar{D}_{14} + 9\bar{D}_{56})\beta_1\beta_3]z^2,\\
h_{66} &= (z^2\bar{D}_{11}\beta_1^2 + 9\bar{D}_{66}\beta_1^2 - \Omega^2)z^2,
\end{aligned}
\tag{15}
$$

where

$$
\beta_i = \beta_{ii}, \quad \beta_4 = \beta_{23} = \beta_{32},
$$

$$
z = \frac{2}{\pi}\xi b, \quad \Omega = \frac{\omega}{\bar{\omega}}, \quad \bar{\omega}^2 = \pi^2\frac{C_{66}}{4\varrho b^2}, \quad \bar{D}_{pq} = \frac{D_{pq}}{C_{66}}. \tag{16}
$$

The determinant of the coefficients of A_j^n set equal to zero

$$
| h_{pq} | = 0 \tag{17}
$$

produces the dispersion relation Ω vs z^2 from which the six branches of the dispersion curves (flexure, face-shear, first thickness-shear in x_1 direction, second thickness-shear in x_3 direction, third thickness-shear in x_1 direction and second thickness-stretch in x_2 direction) can be computed.

4. FREQUENCY SPECTRUM

For each of the roots z_n^2 from the six roots of (17) the five amplitude ratios

$$
\alpha_{1n} = \frac{A_{2n}^{(0)}}{A_{1n}^{(1)}}, \quad \alpha_{2n} = \frac{A_{3n}^{(0)}}{A_{1n}^{(1)}}, \quad \alpha_{3n} = \frac{A_{2n}^{(2)}}{A_{1n}^{(1)}}, \quad \alpha_{4n} = \frac{A_{3n}^{(2)}}{A_{1n}^{(1)}},
$$

$$
\alpha_{5n} = \frac{A_{1n}^{(3)}}{A_{1n}^{(1)}}, \tag{18}
$$

can be obtained from five of the six equations (14). The six displacements given in (13) can be replaced by the relation

$$
u_2^{(0)} = \sum_{n=1}^{6} A_{1n}^{(1)}\alpha_{1n}\sin\xi_n x_1, \quad u_2^{(2)} = \sum_{n=1}^{6} A_{1n}^{(1)}\alpha_{3n}\sin\xi_n x_1,
$$

$$
u_3^{(0)} = \sum_{n=1}^{6} A_{1n}^{(1)}\alpha_{2n}\sin\xi_n x_1, \quad u_3^{(2)} = \sum_{n=1}^{6} A_{1n}^{(1)}\alpha_{4n}\sin\xi_n x_1,
$$

$$
u_1^{(1)} = \sum_{n=1}^{6} A_{1n}^{(1)}\cos\xi_n x_1, \quad u_1^{(3)} = \sum_{n=1}^{6} A_{1n}^{(1)}\alpha_{5n}\cos\xi_n x_1, \tag{19}
$$

The displacement (19) we substitute in the incremental stress-displacement-temperature relationas (11) and the results in the edge conditions (12). We obtain the six equations

$$
\sum_{n=1}^{6} A_{1n}^{(1)} b_{mn} = 0, \quad m = 1, ..., 6, \tag{20}
$$

The frequency spectrum (Ω vs a/b) and the frequency-temperature characteristics (Ω vs Θ for the given dimensions ratio a/b) of the unelectroded strip bounded in x_1 direction can be obtained from the relation

$$
| b_{mn} | = 0. \tag{21}
$$

5. MECHANICAL EFFECT OF PLATING

Let the Y-cut quartz strip is fully covered by the metal the electrodes which are perfectly conduting, are extremely thin compared with the thickness of the crystal plate, and are isotropic in the plane. The thickness of the upper electrode lets be $2b'$ and the lower electrode $2b''$. The upper and the lower electrodes assumed to be from the same material with the mass density ϱ', the planar elastic stiffness γ_{1111}^{p} and the thermal expansion coefficient $\alpha^{\Theta'} = \alpha_{ii}^{\Theta'}$ $(i = 1, 2, 3)$ and nonzero are also only $\beta' = \beta_{ii}'$. All forces, $t_{2b}^{(0)'}$ and $t_{2b}^{(0)''}$, and moments, $t_{ab}^{(1)'}$ and $t_{ab}^{(1)''}$, per unit length associated with the elementary theory of flexure assumed to be zero for the thin metal layers.

The Mindlin's simplified extensional equations for the very thin electrode platings [11] may be expressed as the incremental equations for the upper electrode in the form

$$
\delta_{ib}\beta_{ib}' t_{ab,a}^{(0)'} + \frac{1}{b'}\beta_{is}F_s^{(0)'} = 2\varrho'\ddot{u}_i^{(0)'} \qquad a, b = 1, 3; \quad i, s = 1, 2, 3 \tag{22}
$$

where

$$
F_i^{(0)'} = t_{2i}'(b') - t_{2i}'(-b'), \tag{23}
$$

$$
t_{ab}^{(0)'} = 2\gamma_{abcd}^{\Theta}e_{cd}^{(0)'}, \tag{24}
$$

$$
e_{cd}^{(0)'} = \frac{1}{2}(\beta_{kd}'u_{k,c}^{(0)'} + \beta_{kc}'u_{k,d}^{(0)'}). \tag{25}
$$

The similar equations can be written for the lower electrode.

Let the outside surfaces of the electrodes are traction free. Then the traction conditions take the form

$$
\beta_{jk}' t_{2k}'(b') = 0, \qquad \beta_{jk}' t_{2k}''(-b'') = 0,
$$

$$
\beta_{jk}' t_{2k}'(-b') = \beta_{jk}t_{2k}(b), \qquad \beta_{jk}' t_{2k}''(b'') = \beta_{jk}t_{2k}(-b). \tag{26}
$$

and

$$
u_j(b) = u_j^{(n)}, \qquad u_j(-b) = (-1)^n u_j^{(n)}. \tag{27}
$$

The six surface loadings $F_2^{(0)}$, $F_3^{(0)}$, $F_1^{(1)}$, $F_2^{(2)}$, $F_3^{(2)}$ and $F_1^{(3)}$ which occure in the incremental stress equation of motion (9) can be obtained from (2)

$$
\begin{aligned}
F_2^{(0)} = F_2^{(2)} &= t_{22}(b) - t_{22}(-b),\\
F_3^{(0)} = F_3^{(2)} &= t_{23}(b) - t_{23}(-b),\\
F_1^{(1)} = F_1^{(3)} &= t_{12}(b) - t_{12}(-b).
\end{aligned}
\tag{28}
$$

The succesive addition and sudtraction of (23) and its double primed counterpart, respectively, we obtain

$$
\begin{aligned}
F_i^{(0)'} + F_i^{(0)''} &= -(t_{2i}'(-b) + t_{2i}''(b)),\\
F_i^{(0)'} - F_i^{(0)''} &= -(t_{2i}'(-b) - t_{2i}''(b)).
\end{aligned}
\tag{29}
$$

After substituting from (26) into (28)

$$\beta_{22}F_2^{(0)} = \beta_{22}F_2^{(2)} = \beta'_{22}(t'_{22}(-b) - t''_{22}(b'')),$$

$$\beta_{33}F_3^{(0)} = \beta_{33}F_3^{(0)} = \beta'_{33}(t'_{23}(-b) - t''_{23}(b'')),$$

$$\beta_{11}F_1^{(1)} = \beta_{11}F_1^{(3)} = \beta'_{11}(t'_{12}(-b) - t''_{12}(b'')) \qquad (30)$$

and considering (29) we have

$$\beta_{22}F_2^{(0)} = \beta_{22}F_2^{(2)} = -\beta'_{22}(F_2^{(0)'} + F_2^{(0)''}),$$

$$\beta_{33}F_3^{(0)} = \beta_{23}F_3^{(0)} = -\beta'_{33}(F_3^{(0)'} + F_3^{(0)''}),$$

$$\beta_{11}F_1^{(1)} = \beta_{11}F_1^{(3)} = -\beta'_{11}(F_1^{(0)'} - F_3^{(0)''}). \qquad (31)$$

From the sum and difference of (22) and its double primed counterpart we can express $F_i^{(n)'} \pm F_i^{(n)''}$. Substituting the result into (31), using (27) and considering that $u_j^{(0)'}(b') = u_j(b)$ and $u_j^{(0)''}(b') = u_j(-b)$ we obtain

$$\begin{aligned}
\beta_{22}F_2^{(0)} &= 2\varrho\omega^2 bRu_2^{(0)},\\
\beta_{22}F_2^{(2)} &= 2\varrho\omega^2 bRu_2^{(2)},\\
\beta_{33}F_3^{(0)} &= 2\varrho\omega^2 bRu_3^{(0)},\\
\beta_{33}F_2^{(2)} &= 2\varrho\omega^2 bRu_3^{(2)},\\
\beta_{11}F_1^{(1)} &= \beta'_{11}(b't'_{11,1} - b''t''_{11,1}) + 2\varrho\omega^2 bRu_1^{(1)},\\
\beta_{11}F_1^{(3)} &= \beta'_{11}(b't'_{11,1} - b''t''_{11,1}) + 2\varrho\omega^2 bRu_1^{(3)}, \qquad (32)
\end{aligned}$$

where

$$R = \frac{\varrho'b' + \varrho''b''}{\varrho b}. \qquad (33)$$

Substituting from (32) into (9) we received the incremental stress-strain-temperature equation of motion for the plated Y-cut quartz plate in the form

$$\beta_{22}\tau_{12,1}^{(0)} + 2\varrho\omega^2(1+R)u_2^{(0)} = 0,$$

$$\beta_{23}\tau_{13,1}^{(0)} + 2\varrho\omega^2(1+R)u_3^{(0)} = 0,$$

$$\beta_{11}\tau_{11,1}^{(1)} - \frac{\pi}{2b}\beta_{11}\bar{\tau}_{12}^{(1)} + \varrho\omega^2(1+2R)u_1^{(1)} = 0,$$

$$\beta_{22}\tau_{12,1}^{(2)} - \frac{\pi}{b}\beta_{22}\bar{\tau}_{22}^{(2)} + \varrho\omega^2(1+2R)u_2^{(2)} = 0,$$

$$\beta_{23}\tau_{13,1}^{(2)} - \frac{\pi}{b}\beta_{23}\bar{\tau}_{23}^{(2)} + \varrho\omega^2(1+2R)u_3^{(2)} = 0,$$

$$\beta_{11}\tau_{11,1}^{(3)} - \frac{3\pi}{2b}\beta_{11}\bar{\tau}_{12}^{(3)} + \varrho\omega^2(1+2R)u_1^{(3)} = 0, \qquad (34)$$

where

$$\tau_{12}^{(0)} = t_{12}^{(0)}, \quad \bar{\tau}_{12}^{(1)} = \bar{t}_{12}^{(1)}, \quad \bar{\tau}_{12}^{(3)} = \bar{t}_{12}^{(3)},$$
$$\bar{\tau}_{22}^{(1)} = \bar{t}_{22}^{(1)}, \quad \bar{\tau}_{23}^{(2)} = \bar{t}_{23}^{(2)},$$

$$\beta_{23}\tau_{13}^{(0)} = \beta_{23}t_{13}^{(0)} + \frac{b'}{b}\beta'_{33}(t_{13}^{(0)'} + t_{13}^{(0)''}),$$

$$\beta_{11}\tau_{11}^{(1)} = \beta_{11}t_{11}^{(1)} + \frac{b'}{b}\beta'_{33}(t_{11}^{(0)'} + t_{11}^{(0)''}),$$

$$\beta_{23}\tau_{13}^{(2)} = \beta_{23}t_{13}^{(2)} + \frac{b'}{b}\beta'_{33}(t_{13}^{(0)'} + t_{13}^{(0)''}),$$

$$\beta_{11}\tau_{11}^{(3)} = \beta_{11}t_{11}^{(3)} + \frac{b'}{b}\beta'_{33}(t_{11}^{(0)'} + t_{11}^{(0)''}), \qquad (35)$$

From the condition of the continuity of the mechanical displacement (27) we obtain

$$\beta_{33}e_{13}^{(0)'} = \beta'_{33}e_{13}^{(n)}, \quad n = 0, 2$$

$$\beta_{11}e_{11}^{(0)'} = \beta'_{11}e_{11}^{(m)} \quad m = 1, 3. \qquad (36)$$

Substituting from (2) and (24) into (35) we obtain

$$\begin{aligned}
\tau_{13}^{(0)} &= 2[D_{55}^*(\beta_{23}u_{2,1}^{(0)} + \beta_{33}u_{3,1}^{(0)}) + D_{56}(\beta_{22}u_{2,1}^{(0)} + \beta_{32}u_{3,1}^{(0)})] +\\
&+ \frac{2}{b}D_{56}\beta_{11}(\kappa_1 u_1^{(1)} + u_1^{(3)}),\\
\tau_{11}^{(1)} &= D_{11}^*\beta_{11}u_{1,1}^{(1)} + \frac{4}{3b}(D_{12}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}) +\\
&+ D_{14}(\beta_{23}u_2^{(2)} + \beta_{33}u_3^{(2)}),\\
\tau_{13}^{(2)} &= D_{55}^*(\beta_{23}u_{2,1}^{(2)} + \beta_{33}u_{3,1}^{(2)}) + D_{56}(\beta_{22}u_{2,1}^{(2)} + \beta_{32}u_{3,1}^{(2)}) -\\
&- D_{56}\beta_{11}(\frac{2}{3b}u_1^{(1)} - \frac{18}{5b}u_1^{(3)}),\\
\tau_{11}^{(3)} &= D_{11}^*\beta_{11}u_{1,1}^{(3)} - \frac{8}{5b}[D_{12}(\beta_{22}u_2^{(2)} + \beta_{23}u_3^{(2)}) +\\
&+ D_{14}(\beta_{23}u_2^{(2)} + \beta_{33}u_3^{(2)})], \qquad (37)
\end{aligned}$$

and

$$\tau_{12}^{(n)} = t_{12}^{(n)}, \quad \bar{\tau}_{ij}^{(n)} = \bar{t}_{ij}^{(n)}, \qquad (38)$$

where

$$D_{11}^* = D_{11} + 4\frac{b'}{b}(\frac{\beta'_{11}}{\beta_{11}})^2\gamma_{11}^{\Theta},$$

$$D_{55}^* = D_{55} + 2\frac{b'}{b}(\beta'_{33}\hat{\beta}_{23})^2\gamma_{55}^{\Theta} \qquad (39)$$

and

$$\hat{\beta}_{23} = \frac{\beta_{23}}{\beta_{22}\beta_{33} - \beta_{23}^2}.$$

6. CONCLUSION

The frequency temperature characteristics of the thickness-shear vibrations of the Y-cut quartz plates are given especially by the third and sixth equations of (14). From the comparison of h_{33} and h_{66} for the plated resonators

$$h_{33} = (z^2\bar{D}_{11}^*\beta_1^2 + \bar{D}_{66}\beta_1^2 - \Omega^2)z^2,$$

$$h_{66} = (z^2\bar{D}_{11}^*\beta_1^2 + 9\bar{D}_{66}\beta_1^2 - \Omega^2)z^2$$

follows that the influence of the D_{11}^* on the frequency-temperature characteristic of the third overtone given by h_{66} is smaller then on the frequency-temperature characteristic of the fundamental thickness-shear mode of vibration given by h_{33}.

7. REFERENCES

[1] H.F. Tiersten and B.K. Sinha, "Temperature induced frequency changes in electroded AT-cut quartz thickness-shear resonators," in *Proc. 31th AFCF*, pp. 23-28, 1977

[2] D.S. Stevens and H.F. Tiersten, "Temperature induced frequency changes in electroded AT-cut quartz trapped energy resonators," in *Proc. 34th AFCS*, pp. 384-392, 1980.

[3] J.H. Sherman, "Temperature coefficient of the frequency shift arising from electrode film stress," in *IEEE Trans. Son. Ultrason*, vol. SU-30, pp. 104-110, 1983.

[4] J. Zelenka, "The calculation of the resonant frequency temperature dependence of quartz bulk resonators," in *Proc. 8th Piezoel. Conf. PIEZO'94, Zakopane, Poland*, In print.

[5] J. Pavlovec, J. Suchanek and J. Zelenka, "The influence of the thickness of electrodes on the frequency-temperature characteristic of AT-cut quartz resonators," in *Proc. 8th Piezoel. Conf. PIEZO'94, Zakopane, Poland*, In print.

[6] P.C.Y. Lee and Y.K. Yong, "Frequency-temperature behaviour of thickness vibration of double-rotated quartz plates," in *38th AFCF*, pp. 164-175, 1984.

[7] P.C.Y. Lee and Z. Nikodem, "An approximate theory for high-frequency vibrations of elastic plates," in *Int. J. Solids Structures 8*, pp. 581-612, 1972.

SESSION 14

SYNCHRONIZATION IN SYSTEMS

Chairman : P. KARTASCHOFF

Highly Accurate Realtime Clocksynchronization with GPS Time Receivers

*Hartmut Nau, W. Klische**

DLR Oberpfaffenhofen, Institute of Radio Frequency Technology, 82230 Wessling-Germany
** K+K, Vechelde - Germany

1. Abstract

There is a need for highly accurate synchronization of clocks in an ground network with remote stations in distance of up to several 100 km. For intance monitoring the hight of airplanes from four ground stations. There fore a study was made to find out what accuracy can be achieved with the help of GPS time receivers operated in an differential mode replacing atomic clocks.

After a short introduction of the function of an GPS time receiver the error analysis for differential GPS clocksychronizations under SA (selective availability) conditions will be given.

In an experiment tests were performed with time receivers from different companies. In an first step the difference between the 1 pps output signals of two receivers at one place was observed and an accuracy below 10 ns could be achieved.

2. Principle of GPS Time Receivers

The principle of an GPS receiver for time is to compare the calculated slant range between the receiver location and the satellite position with the measured transit time of the radio signal. The clock bias is the result of the comparison by taking into account the different delay times caused by ionspheric, tropospheric and receiver delay. The receiver clock can be synchronized to the atomic clock on board the satellite. The difference between onboard time and GPS time and also to UTC (USNO) Washington D.C. is given by polynom coefficients in the data message transmitted together with the ranging signal [3]. The position of the satellite has to be evaluated from the ephemeris data transmitted by the satellite. If the location coordinates of the receiver are already known it is sufficient to track one satellite, otherwise they have to be determined by tracking four satellites. In this way a clock synchronization to GPS time or UTC (USNO) can be realized. Typical time accuracies are 100 ns (1 Sigma) given in [2].

But the level of accuracy is not available for the civil user now. Selective Availability (SA) is the name given to the method chosen to create two levels of accuracy.

With SA on the level of accuracy is reduced to 320 ns [4]. Companies producing GPS time receivers try to overcome the degradation by different techniques.
Ball Efratom is using the GPS measurements to discipline a Rubidium frequency oszillator. Accuracies below 100 ns could be achieved [5]. True Time Corporation has announced a new algorithm SatHopTM to reduce the degradation effect of SA [8].

K + K (Velchede) is developing a precise time receiver by making extensive use of phase carrier measurement [6].

If not the absolute time scale UTC has to be reproduced but only clock synchronization between distant clocks is needed, a differential method can be used. The relative time accuracy between distant stations can be much better than the absolute. If the stations simultaneously track the same GPS satellite, then the effect of certain control segment- and satellite-induced errors on the relative time accuracy are much reduced, and relative time accuracy can be as good as 10 to 20 ns.

3. Error analysis for differential GPS clock synchronization under SA conditions

Satellite Ephemeris Error

Satellite ephemeris errors have small error effects on the differential user caused by slightly different look angles to each satellite. The effect is small due to the high altitude of the GPS satellite orbits (20 000 km). The decorrelation in this case is caused by the different line-of-sight. As illustrated in Figure 1, for a separation distance of 100 km, the angular separation of the two line-of-sights is $D/r \approx 0.3°$. Figure 1 gives the situation where the satellite S at location Strue is believed to be at location Sassumed. These discrepancies can be due to inaccuracies in ephemeris or can be intentionally introduced as SA technique. For users in a distance D from each other the range error introduced is

$RE = \Delta r - \Delta r'$,

or

$RE \approx d \sin\alpha - d \sin(\alpha - \epsilon)$
$RE \approx d (\sin\alpha - \sin\alpha \cos\epsilon + \sin\epsilon \cos\alpha)$.

First order approximation for small ϵ gives

$RE \approx \epsilon\, d \cos\alpha$
with $\epsilon \leq D/r$.

This leads to $RE \leq (D \times d \times \cos\alpha)/r$

The worst case will be for $a = 0°$

For a worst case approximation we have

$$RE \leq \frac{D}{r} \cdot \Delta r$$

With SA applied a position accuracy of the satellite is assumed of d=100m. This assumption is based on the GPS-operator guaranteed User Range Accuracy of the slant range to a satellite of URA=32m (1-sigma). For a worst case consideration with a seperation of the receiving stations of D=100 km, a height of the satellite above ground of r=20 000km and a satellite position accuracy of d=100 m RE results in 0.5 m,

with $c = 3 \times 10^8$ m/s the time error is TE = ± 1.7 ns

The maximum time error introduced by the unsufficient knowledge of the ephemeris of the satellite is therefore less than ± 2 ns.

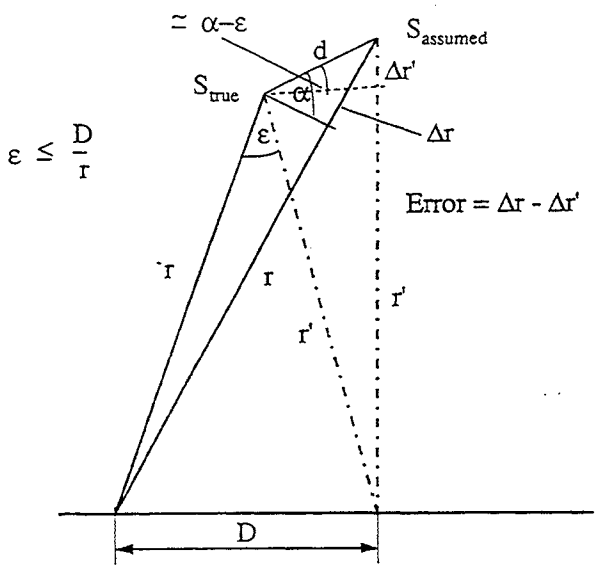


Figure 1. Degradation of accuracy with distance of receiving stations.

Selected Availability

SA has two components which degrade the accuracy for time diseemination, clock "dither" and reduced accuracy of the ephemeris parameters called "epsilon". The effect of "epsilon" on the differential method is as discussed before. Dither has no effect if all stations take measurements simultaneously.

Clock Error

A constant clock error doesn't lead to an error in differential mode.

Ionospheric Propagation Delay Errors

Ionospheric irregularities can cause residual errors because the stations' propagation paths to the satellite see a slightly different ionospheric environment. Propagation tests [ ] have indicated that the irregularities can result in propagation delays that vary by about half a meter.

Tropospheric Delay Errors

The tropospheric model accurately accounts for all but about 1 to 3 meters. With differential operation, most of these remaining errors cancel out as well, except for low angles below five degrees.

Multipath Error

Multipath errors are potentially serious effects for differential GPS. Much of the problems can be eliminated by good antenna design (choke ring) and placement. Otherwise multipath could cause differential errors of a few meters.

Receiver Noise

Receiver noise, caused primarily by measurement noise, is particularly bothersome in differential applications because the participating station receivers will have uncorrelated noise. Therefore modern receivers use carrier phase measurements for smoothing [6].

Error budget for time synchronisation in a differential mode of operation with a separation of the stations of about 100 km.

• Satellite ephemeris error • Selective availability	2.0 ns
• Clock error	-
• Ionospheric propagation delay error	2.0 ns
• Tropospheric propagation delay error	1.7 ns
• Multipath error	2.0 ns
• Receiver noise	5.0 ns
	$\sum$ ± 6.3 ns

4. Experimental Results

Two different types of receivers, two from Magnavox 4200 D and two from K + K, were tested by operating them at the same place. The results are given in Figures 2-5. For short periods of time (2 hours) the precision of the difference between the 1 pps pulses was $\delta < 10$ ns.

By averaging the measurements over one day, the results increased and the precision becomes s ≈ 40 ns (Magnavox) and s = 10 to 15 ns (K + K). This are very preliminary results which have to be proved by further investigations. Especially the results by operating two receivers at distant sites have to be proved by further experiments. The coordination problem for coordinated simultaneous-view time transfer has to be solved.

5. Summary

In several aplication synchronized clocks at distant places are needed. As a goal 10 ns should be envisaged. This can be done by using GPS time receivers operating in a differential mode using a coordinated simultaneous-view time transfer technique.

All time receivers track the same satellite at a specific time. This can be coordinated by one station (e.g. Master Station) using ephemeris broadcast data from broadcast Almanach for evaluation predictions of the satellite's visibility.

An accuracy of 10 to 20 ns with an GPS Time Receiver using pseudo noise code and carrier phase measurements to discipline a quarz oszillator clock seems to be applicable. But further investigations have to be done.

Acknowledgement: I would like to thank Mr. J. Hahn for the data evaluations.

6. References

[1] NAVSTAR GPS User Equipment-Introduction; NATO, Feb. 1991.

[2] NAVSTAR Global Positioning System (GPS); STANAG 4294 NATO, 1. Oktober 1991.

[3] G. Hübner: Performance of GPS Disciplined Rubidium Os cillators under Selective Availability (SA); Proceedings EFTF, March 1994, Munich.

[4] H. Nau: Satellitennavigationssysteme für weltweite Zeitver teilung, Uhren synchronisation und Frequenzvergleiche; Funkuhren Zeitsignale Normalfrequenzen S. 163 W. Hil berg: Sprache und Technik, Groß-Bieberau, 1993.

[5] W. Klische; Verwendung von GPS-Empföngern zur Frequenz- und Zeitbestimmung; W. Hilberg: Sprache und Technik, Groß-Bieberau, 1993.

[6] B. Breuer, J. Campbell, A. Müller: GPS-Meß- und Auswerte verfahren unter operationellen GPS-Bedingungen;

[7] B. Breuer, J. Campbell, A. Müller: GPS-Meß- und Aus werte verfahren unter operationellen GPS-Bedingungen; SPN Zeitschrift für Satellitengestützte Positionierung, Navigation und Kommunikation 3/93, 2. Jahrgang, September 1993.

[8] W. Melbourne et al.: Oscillator disciplining with Selective Availability truetime "Sathoptm" Algorithm; SPN Journal for Satellite-Based Positioning, Navigation and Communikation 4/93, 2. Jahrgang, Dezember 1993.

Quick look / raw data - measurement started 14.02.95
MAG1 - MAG2

File:	140295.UHR
Number of measurements:	130000
Time interval τ:	1 s
Stdev:	40.1 ns
Mean:	467.3 ns
Program version:	01/95

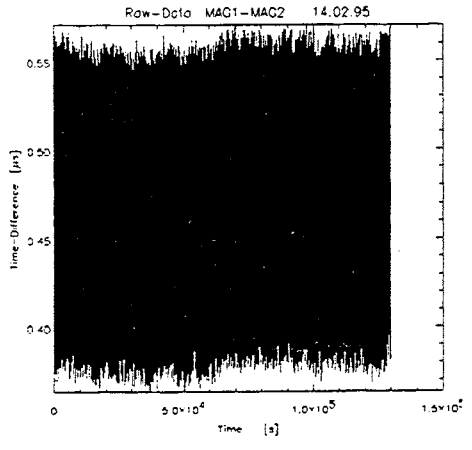


Figure 2. Time difference between two receivers (Magnavox) tracking all satelliltes in view at the same place for about 24 hours.

Quick look / raw data - measurement started 29.04.94
K&K1- K&K2

File:	290494.UHR
Number of measurements:	9909
Time interval τ:	10 s
Stdev:	13.3 ns
Mean:	100.7 ns
Program version:	01/95

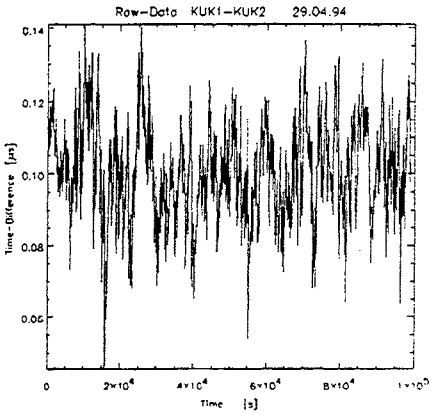
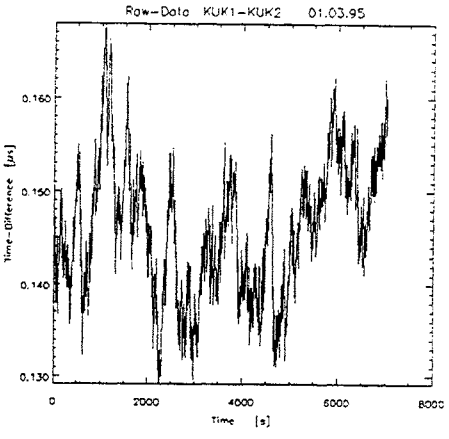


Figure 3. Time difference between two receivers (K+K) tracking all satellites in view at the same place for about 24 hours.

Institut für Hochfrequenztechnik

Quick look / raw data - measurement started 01.03.95
K&K1- K&K2

File:	010395_1.UHR
Number of measurements:	7026
Time interval τ:	1 s
Stdev:	7.2 ns
Mean:	145.9 ns
Program version:	02/95



Data processed by J. Hahn, March 2, 1995

Figure 4. Time difference between two receivers (K+K) tracking the same satellite two hours at the same place.

Development of the UTC(CRL) Time Dissemination Server for the Computer Network

Tadahiro GOTOH (tara@crl.go.jp)
Kashima Space Research Center, Communications Research Laboratory
893-1 Hirai, Kashima, Ibaraki, 314 Japan

Kuniyasu IMAMURA
Communications Research Laboratory
4-2-1 Nukui-Kita, Koganei, Tokyo, 184 Japan

Abstract

The recent rapid spread of computer networks has made computer synchronization very important. Network Time Protocol (NTP) is excellent for computer synchronization and is now widely used, but it has several problems limiting the precision and accuracy of synchronization. CRL has developed a new design philosophy and designed a new NTP server. This philosophy is used to propose a method to evaluate the network effect. This paper discusses the problems of the NTP and presents the design philosophy and design details of the new server.

1 Introduction

Since the rapid spread of computer networks such as Internet throughout the world, time synchronization of computers has become very important. With a main frame computer system where batch processing was dominant, there was a single central computer which kept the time, and so there was no need for time synchronization.

However, in a system where many computers are connected in a network, it can sometimes be problematic if each computer clock is running asynchronously. Thus, a primitive method for time synchronization was adopted, in which computers refer to each other for the time to establish synchronization in the network. However, in a wide-area network with high-speed communication channels, the number of computers connected is so great that time synchronization by this method is difficult. Thus, for time synchronization in a wide-area network such as Internet, the Network Time Protocol (NTP) was developed. The most important characteristics of the NTP are that it can compensate the network delay time and synchronize many remotely-located clocks. The NTP is now the most popular time synchronization protocol in the Internet, but it also has several problems which need to be overcome to achieve high accuracy and precision of synchronization.

The Communications Research Laboratory (CRL), the standard organization for time and frequency in Japan, has designed a new NTP server that considers these problems. This paper discusses the problems of the NTP and presents the design philosophy and design details of the new NTP server.

2 Network Time Protocol

2.1 Principle of the NTP

The NTP was developed for time synchronization among UNIX workstations in the Internet. Figure 1 shows the hierarchy of the NTP system which is made up from 16 strata. Each stratum, as a client, requests time information from the upper stratum, and as a server, supplies time information to the lower stratum.

The stratum 1 has the special function of getting time information from an external reference clock through the serial I/O port of a UNIX machine in order to synchronize its internal clock. Using a receiver for standard-frequency and time-signal emissions or a GPS receiver as the external reference clock, a stratum 1 server can synchronize itself with UTC. The NTP package already provides the program codes for WWV and DCF77. By connecting these signals with the serial I/O port, the clock of stratum 1 can easily be synchronized with UTC.

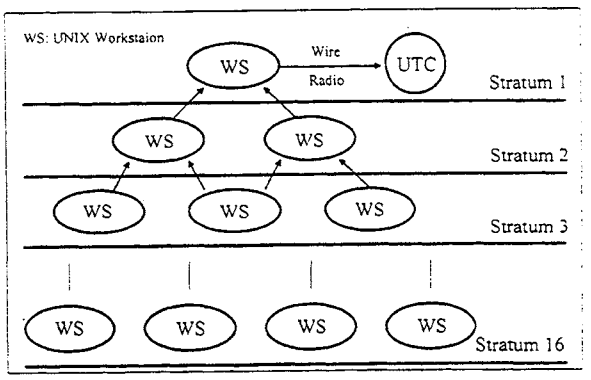


Figure 1: NTP Strata

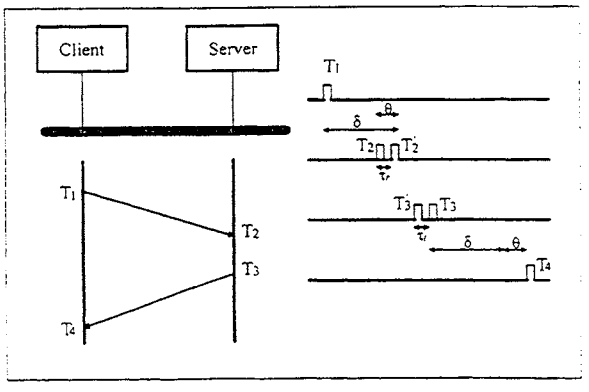


Figure 2: Calculating Delay and Offset

Figure 2 illustrates how a remotely-located clock is synchronized by the NTP. A client transmits a request packet at time T_1, and a server receives the packet at time T_2 and transmits a return packet at time T_3. The client receives the packet at time T_4. These times are recorded in NTP packets and used to compensate the network propagation delay time δ and the clock offset θ as shown in equations (1) and (2). Here, one should notice that there is some timing error τ_r due to the software de-

lay in the receiving process. This is the time difference between the receiving time T_2' recorded by the program and the time T_2 when the server physically received the NTP packet. There is also a similar timing error τ_t due to the software delay in the transmitting process. These errors result in the synchronization error.

$$\delta = (T_4 - T_1) - (T_3 - T_2) \quad (1)$$

$$\theta = \frac{(T_2 - T_1) + (T_3 - T_4)}{2} \quad (2)$$

Information on the condition of the server clock is also recorded in the NTP packet and sent back to the client. The client accesses several servers and synchronizes its internal clock with the server clock that is in the best condition, judging from the conditions of the servers and the delay times.

While most software for time synchronization sets a clock directly for the desired time, the NTP uses the UNIX system call of *adjtime* for clock synchronization. This system call does not directly set the clock, but adjusts its frequency. Thus there is no time step with the NTP, which is inevitable in the case of other software.

2.2 Problems of the NTP

Although the NTP is an excellent protocol for clock synchronization, it has several problems, particularly when a UNIX workstation is used as a server.

The first problem is the low time resolution of UNIX workstations. Although the UNIX OS allows the time resolution of 1 microsecond, the time resolution of most UNIX machines is only 1 millisecond at most, since they were not originally designed for real-time processing and precise time was not needed in the early days of network history. Therefore, even if the precise time is given, the real-time resolution is limited by that of the internal clock of the server. The second problem is that the UNIX standard I/O interface is a serial one. In many NTP systems, the stratum 1 server uses the serial I/O port for getting time information from an external clock. However, the timing of data transfer in a serial port is determined by the onboard clock of the port. The jitter of this clock degrades the timing precision and the low transfer speed of the serial I/O also limits the time resolution.

The third problem is the UNIX OS itself. The UNIX OS is a multi-task OS, which means that all processes are executed in a time-divided way. Thus, the required time for processing an NTP packet τ_r or τ_t changes depending on the load of the server and is very difficult to estimate.

The fourth problem is the asymmetry of the propagation delay times of the incoming and outgoing paths. The NTP assumes that the delay times in the incoming and outgoing paths are the same. But, in a real network, this assumption is wrong. This asymmetry may result in a large synchronization error. However, the effect of this error has not been qualitatively evaluated.

3 Design of a New NTP Server with High Precision

3.1 Design Philosophy and Outline of the Server

Considering these problems of the NTP, the following design philosophy for an NTP server is obtained.

1. An external clock of high performance should be used as the server instead of the internal clock of computer.
2. A parallel I/O interface with high I/O speed should be used instead of a serial I/O interface.
3. A single-task OS should be used instead of a multi-task OS.
4. A method must be established to evaluate the error due to the asymmetry of the propagation delay time in the network.

Based on this design philosophy, a new NTP server has been designed. Figure 3 shows the outline of the server. Instead of the clock in the server machine, a Time Code Generator (TCG) is used as a stable external reference clock, which has the time resolution of 1 microsecond. The time code is transferred through a PC/AT ISA bus which is a parallel I/O bus and has high I/O speed. An IBM compatible PC which runs under a single-task OS of MSDOS is used as the server machine, not a UNIX machine.

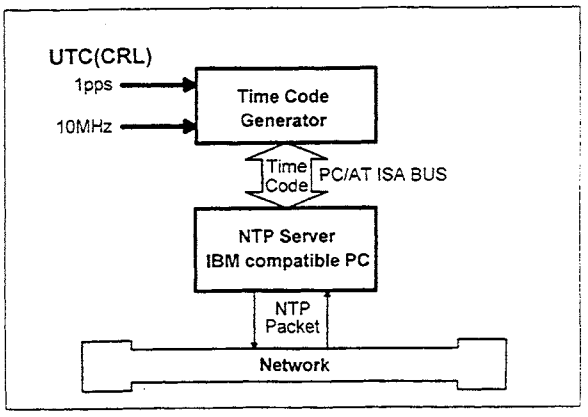


Figure 3: New NTP Master Server

3.2 Time Code Generator

Figure 4 shows a block diagram of the TCG. The TCG generates the time code of 64 bits by using two 32-bit synchronous counters; namely the second binary counter and the microsecond binary counter which are driven by the 1 pps and 10 MHz signals from UTC(CRL). The time code of the TCG is compatible with that of the UNIX system which is the number of seconds elapsed since January 1 1970, and can be easily transferred to a UNIX system. Since the time code of TCG is directly used as the timestamp for the NTP packets, the precision and resolution is not limited by the poor performance of the internal clock at all.

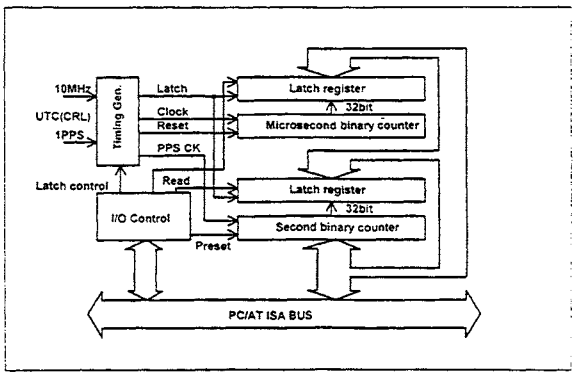


Figure 4: Time Code Generator Block Diagram

The data of the counters are output to the PC/AT ISA bus through Latch registers. In order to get the time code data of 64 bits, the computer must read the counter data four times in turn, because the PC/AT ISA bus is a 16-bit parallel bus. The four counter data are simultaneously latched at the timing of the first reading. Judging from the transfer speed of the

PC/AT ISA bus, it takes 1.5 microseconds for the computer to get all the time code data. But this delay time is constant since no interrupt process is used. Thus, the data transfer time is greatly improved compared with that of a serial I/O bus.

3.3 Software Design of the Server

The software for the newly-designed server runs under the MS-DOS which is a single-task OS. Under a single-task OS, processing time is almost constant and can be estimated from the number of program steps. Hence, it is possible to reduce the timing error due to the software delay by estimating τ_r and τ_t.

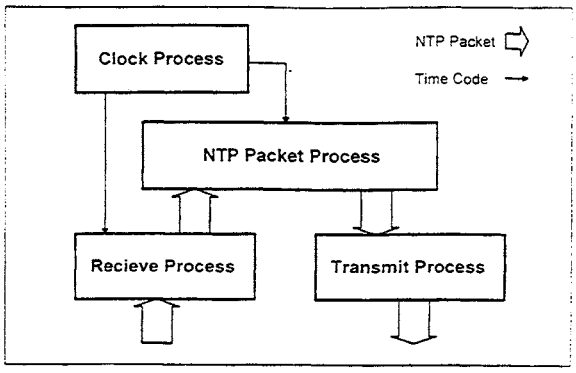


Figure 5: Implemantation Model of Software

As shown in Figure 5, the software consists of four processing units. The first unit is the receive process, which receives the packet from the network and records the receive time, whether the packet is an NTP packet or not. In the usual NTP software, the receiving time of the NTP packet is recorded after the receiving process in each layer of the TCP/IP. In this software, on the other hand, it is instantaneously recorded when an interrupt process from the network driver occurs. Thus, τ_r can be made very small. The second unit is the NTP packet process, which analyzes the received packet and generates a return packet. The third unit is the transmit process, which sends the return packet back to the network. The fourth unit is the clock process, which provides an accurate time for the other processes.

4 Evaluation of the Network Delay Error

4.1 Asymmetry of Network Paths

As already stated in Section 2.2, the asymmetry of the delay times of the incoming and outgoing paths results in a synchronization error. The following are the causes of such asymmetry.

In a wide-area network, IP routers are used for inter-network connection. For routing, packets are regenerated in IP routers. In a network with many routers, the processing time in a router may not be the same for incoming and outgoing directions.

In an Ethernet, a popular network, the probability of packet collision will increase as the number of computers connected in the network increases, since many packets flow in a single channel. When packets collide, they are sent again. This resending process changes the propagation delay time of the network path.

As the network gets bigger, there are many paths to a destination, and it cannot be assured that the incoming and outgoing paths are the same. For selection of a server, the path to the server should also be considered carefully, as well as the stratum level of the server. A stratum 2 server with a simple path may give higher precision than a stratum 1 server with complicated paths.

4.2 A Method to Evaluate the Network Effect

This asymmetry effect of a network should be evaluated in order to synchronize computers with high precision. However, it is very difficult to distinguish the asymmetry error with the software error in a UNIX system. As stated in Section 3.3, it is possible to estimate and reduce the software error in the newly-designed server. Therefore, the network effect can be evaluated, as shown in Figure 6. The NTP monitor in the figure is almost the same as the server, except that it has a function to ask for time information from a server as a client. The monitor gets time information from the server via the network and also directly from UTC. From the difference between this time information, one can evaluate the network effect. An experiment will be made on various networks such as a network with heavy traffic and many paths to the destination, or with light traffic and simple paths.

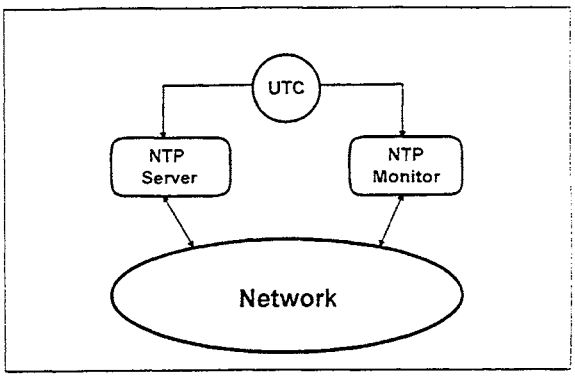


Figure 6: Measuring Network Delay Error

5 Conclusion

Although the NTP is an excellent protocol for computer synchronization, it has several problems. From the examination of these problems, a new design philosophy for an NTP server is developed. Based on this philosophy, a new NTP server is designed and a method to evaluate the network effect is proposed. This new server can achieve higher precision and accuracy of computer synchronization than a standard UNIX machine server.

6 Acknowledgment

The authors thank Hiroyasu FUKASE of Internet Initiative Japan for his discussion and kind cooperation.

References

[1] David L. Mills, "Measured Performance of the Network Time Protocol in the Internet System", RFC1128, October 1989.

[2] David L. Mills, "Network Time Protocol (Version 3) Specification, Implementation and Analysis", RFC1305, March 1992.

SDH-Based Time and Frequency Transfer System

Masami Kihara and Atushi 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

Using transmission systems based on Synchronous Digital Hierarchy (SDH), network synchronization systems can provide time and frequency signals to network nodes. This paper introduces a new network system configuration and interface for time and frequency transfer.

Time and frequency transfer uses a bi-directional path that consists of outgoing and incoming paths composed of the same medium. The frame format based on the synchronous digital hierarchy level (STM-N) is used to transfer time and frequency. A second signal, time information, and delay data are imposed on the Section Over Head (SOH) of STM-N. The new time and frequency transfer scheme can be implemented in one-way, two-way and common-view configurations. These configurations can replace conventional methods.

1. INTRODUCTION

Telecommunication networks can be used to transfer time and frequency. Reference frequency has already distributed to end nodes over the chaining in the master-slave method to synchronize digital cross-connect and switching systems. Reference time can also become a promising tool for a time-stamping in distributed network operation systems, for a fault location in network systems and for time information in broad band ISDN (B-ISDN) services.

Telecommunication networks can disseminate a national standard for frequency and time as well, not just synchronize national/international networks themselves. This time transfer capability is an alternative to conventional dissemination methods of time and frequency such as LORAN-C, GPS and time transfer methods using stationary satellites such as INTELSAT.

We present a network system configuration and interface for time and frequency transfer in global telecommunication networks based on the Synchronous Digital Hierarchy (SDH) with the standardized SDH-based interfaces recommended in ITU-T[1]-[3].

2. BASIC CONCEPT

2.1 Expected performance

Time and frequency are transferred through a transfer medium between two nodes. To transfer time, the location data of nodes is not required: however, it is required that delay and delay variation of the media can be explicitly evaluated. Therefore, time and frequency transfer in telecommunication networks uses a bi-directional path that consists of outgoing and incoming paths composed of the same medium. Difference in delay between outgoing and incoming paths allows us to determine time and frequency error.

Table 1 shows estimated performance in time and frequency transfer. The new method using section layer in the SDH is compared to conventional methods that use the path layer in existing networks. While paths in existing networks can transfer time and frequency without modifying network equipment, equipment delay, which is variable and asymmetry, affect time and frequency performance even if a delay compensation method is applied. Time and frequency transfer capabilities are limited to the order of micro second and worse than 10^{-12}, respectively. On the other hand, the new method uses the section layer in the SDH and so eliminates equipment delay. Its performance is affected only by circuit and fiber cable asymmetry. Time transfer accuracies of submicro second and frequency transfer accuracies of 10^{-14} are feasible.

Table 1 - Time and frequency transfer performance in networks

Network	Frequency	Time
Using path layer in existing networks		
Clock extraction	$10^{-9}-10^{-12}$	—
Time service without delay compensation	$10^{-9}-10^{-12}$	<0.5sec
Time service with delay compensation	$10^{-9}-10^{-12}$	<500μsec
Using section layer in SDH-based networks	$10^{-11}-10^{-14}$	<1μsec

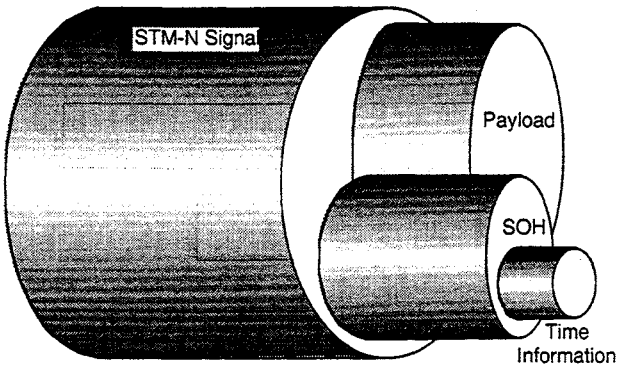


Fig. 1 - STM-N signal structure including a time and frequency signal

Table 2 - Synchronous digital hierarchy bit rates

Synchronous digital hierarchy level	Hierarchical bit rate kb/s
1 (STM-1)	155 520
4 (STM-4)	622 080
16 (STM-16)	2 488 320
64 (STM-64) *	9 953 280

* This level requires further study

2.2 SDH-based interfaces

In view of the global standardized interface and customer time applications in B-ISDN, Synchronous Transport Module - 1 (STM-1) is suitable for the information structure used to transfer network reference time. STM-N, which is defined at a multiple of 155.52Mb/s, can also be used. The frame format based on STM-N shown in Table 2 is in accordance with ITU-T Rec. G.708.

In addition, using the Section Over Head (SOH) offers another advantage. We can use the STM-N payload area for B-ISDN services at the same time when the time and frequency transfer application is executed without reducing the transmission capacity of the transmission system. Nodes receiving network reference time can separate general information in Payload and SOH information, and then extract network reference time from SOH information, as shown in Fig. 1.

Multiplex SOH (MSOH), the lower 5 bytes of 9 rows in SOH (see Fig. 2) , is appropriate for time and frequency transfer because it can be transmitted as far as the opposite network node even if devices such as repeaters are located in the path between network nodes. Promising bytes are shaded in Fig.2 (b). These bytes are not defined by ITU-T for any purpose.

A multi-frame in the bytes assigned to time and frequency transfer should be defined to transmit a second signal because the STM-N frame period is 8 kHz. The multi-frame format can identify a second period as shown in Fig. 3.

2.3 Standardization

Because SDH-based interfaces are used, the new SDH-based time and frequency transfer system can become an international standardized system. To encourage its penetration, we should make effort not to make local (national) standards. Basic physical specifications will be discussed in ITU-T and ITU-R this year[4][5]. The new transfer system should not be rapidly introduced before ITU-T and ITU-R recommend new specifications concerning them.

We also expect that most SDH systems will provide the new function

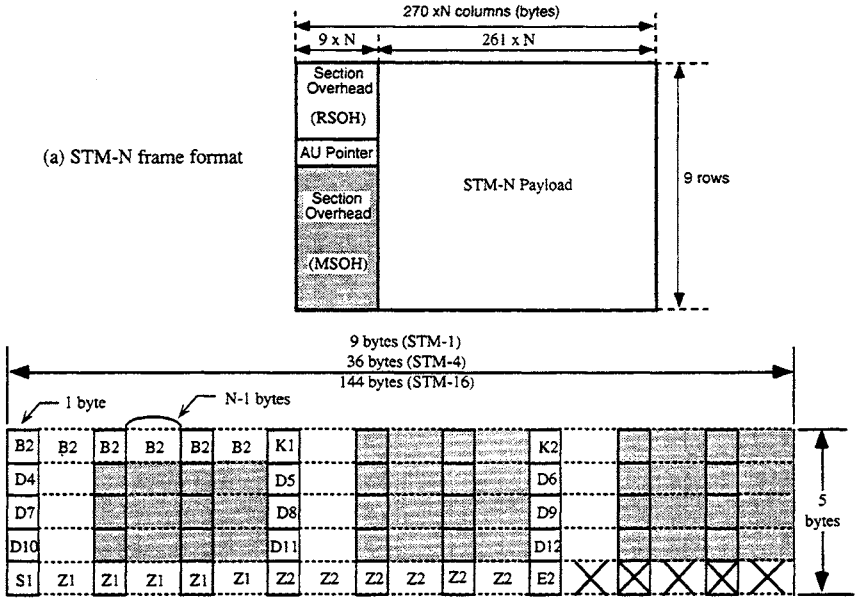


Fig. 2 - STM-N frame format and MSOH byte mapping

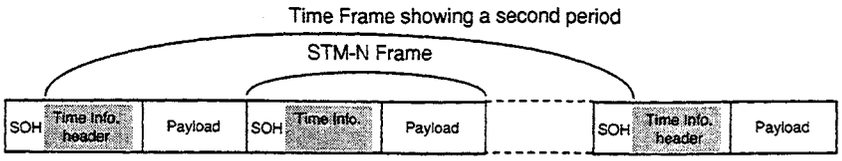
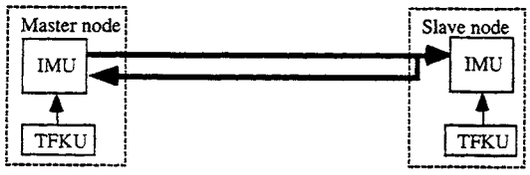
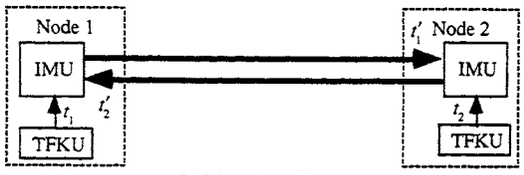


Fig. 3 - Multi frame structure



(a) Master-slave topology



(b) Mutual topology

TFKU: Time and frequency keeping unit
IMU: Interface and measurement unit

Fig. 4 - Basic topologies in time and frequency transfer

because of its value to international telecommunications. While two-way time transfer using stationary satellites such as INTELSAT can only be accomplished for limited periods due to expensive charges, the new method can continuously achieve international two-way time transfer without degrading service quality which would incur high charges.

3. TRANSFER TOPOLOGY

There are two basic topologies in time and frequency transfer as shown in Fig. 4.

1) Master-slave topology

One node always receives the time and frequency signal and so is called the slave node. The other node is called the master node; however, the name of the master node is not related to whether the master node is the origin of the time and frequency signal or not. The master node measures the delay experienced on outgoing and incoming paths over a round-trip and manages the delay data. The time and frequency signal is compensated by the delay data and sent from the master node to the slave node, or the time and frequency signal is first sent to the slave node and then compensated at the slave node by the delay data transferred from the master node. The time and frequency signal is advanced by the delay data which is calculated by the following equation:

$$T_d = T_{rd}/2 \qquad (1)$$

T_d: delay data

T_{rd}: round-trip delay

2) Mutual topology

Both nodes send and receive the time and frequency signal. Delay is caused on the outgoing and incoming paths and is measured at the two nodes by comparing the time and frequency signals generated at each node. Measurement results are exchanged. The time difference between the two nodes can be calculated from the results.

$$T_{12} = \frac{1}{2}(T_{d1} - T_{d2}) \qquad (2)$$

T_{12}: time difference between two nodes

T_{d1}: time difference measured at node 1 between the time and frequency signal sent from node 2 and the time and frequency signal generated at node 1.

T_{d2}: time difference measured at node 2 between the time and frequency signal sent from node 1 and the time and frequency signal generated at node 2

The mutual topology can be used between any two nodes regardless of their hierarchy. Each node measures the time difference between the other node rather than compensating the time and frequency signal.

4. BASIC CATEGORY

The above topologies can be implemented in three ways as shown in Fig. 5. The path defined in the topology is constructed by network elements. The network element connects two locations that are geographically separated and provides both network and node interfaces including Time and Frequency Keeping Unit (TFKU). The interface supports two signals: a time and frequency signal that includes a time marker, time information and delay data, and a data signal which includes time difference data and general data.

1) One-way configuration

Slave nodes receive the time and frequency signal generated at a reference node. The data signal may not be required. Specifications of time and frequency depend on that of the reference node.

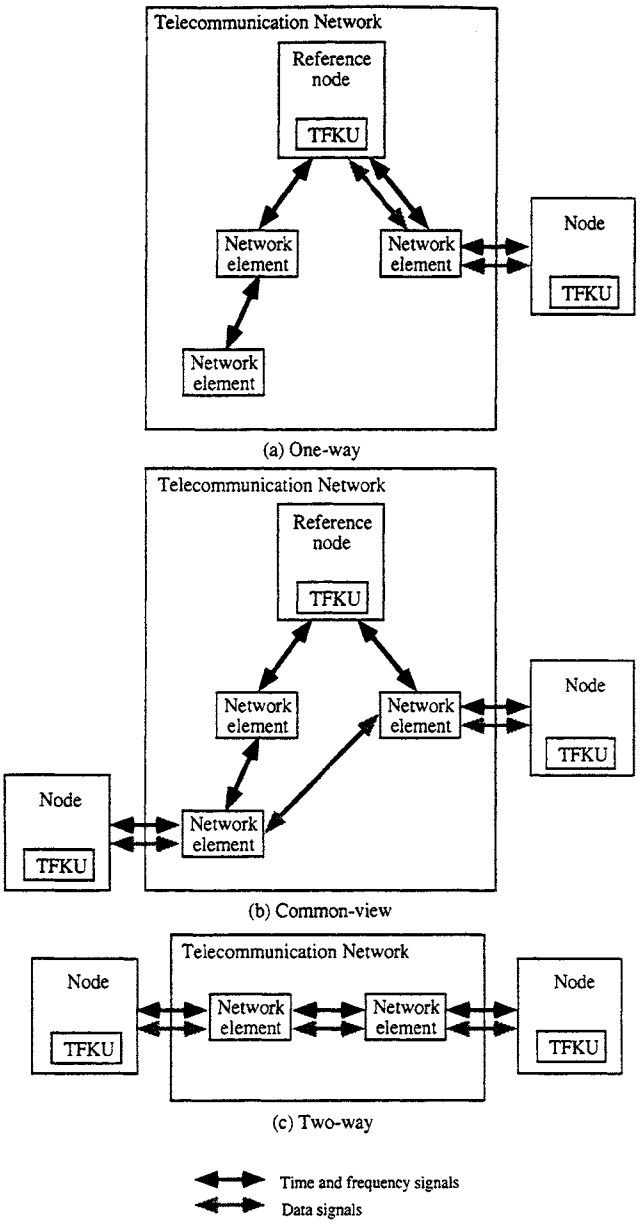


Fig. 5 - Basic category for time and frequency transfer

2) Common-view configuration

All nodes initially receive the same time and frequency signal generated at the reference node as slave nodes. Nodes then exchange data signals to calculate the time difference between them in the mutual topology. The time and frequency error of the reference node TFKU is suppressed within the nodes due to the simultaneous observation of the common source.

3) Two-way configuration

Nodes in the mutual topology can exchange time and frequency signals generated by their TFKUs. The network must provide nodes with access to the interface used by the reference node.

5. SYSTEM CONFIGURATION

Figure 6 shows the basic configuration of the time and frequency transfer system. A 155Mb/s path is established between two nodes which are connected by an STM-N transmission system. Multiplexers/ demultiplexers used in this configuration allow the set up of the time and frequency path shaded in Fig. 6.

In Fig. 6 (a), TFKU can use the time and frequency path through Clock Interfaces (IFs) and SOH Operation Units. The STM-N interface including time and frequency information is used between nodes. Interfaces between Clock IF and TKFU might be different from that of STM-N.

Figure 6 (b) shows the configuration from a network node to Customer Premises Equipment (CPE). Time and frequency information is extracted at SOH Operation Unit and Clock IF, and is imposed on STM-1 IF toward CPE. This is because sections between network nodes and between multiplexer/demultiplexer and CPE are separated, and there is no relationship between sections. While the configuration shown in Fig. 6 (b) is complicated, CPE can use both a general B-ISDN service and a time and frequency information service through one STM-1 IF.

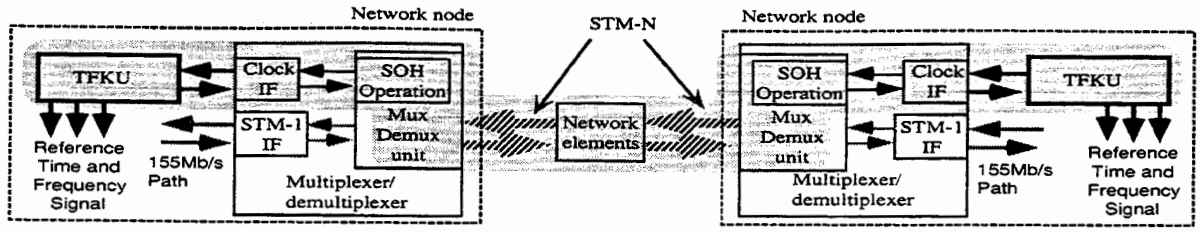
1) One-way example

If the relationship between two nodes is master-slave, the slave node closes the time and frequency path, and the master node can use the time and frequency path as a round trip path. This is a one-way configuration which is suitable for public network nodes in a regional area and for subscriber networks from the public network node to customer premises equipment.

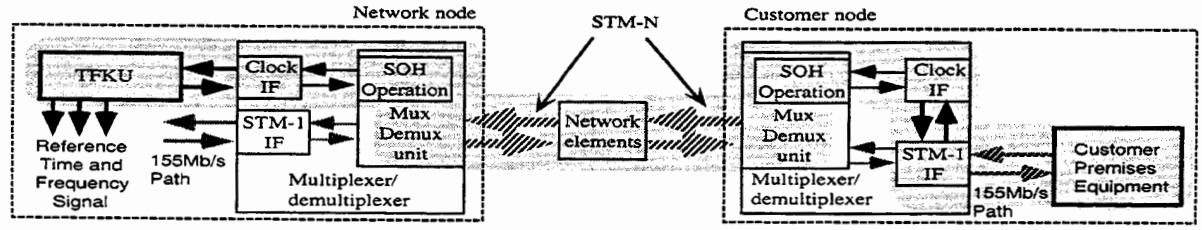
2) Common-view example

Common-view configuration can also be used as a variation of the one-way configuration. In this configuration, the time and frequency error of the reference node TFKU is suppressed within the nodes due to the simultaneous observation of the common source. All nodes initially receive the same time and frequency signal generated at the reference node as slave nodes. Nodes then exchange time data to calculate the time difference between them using the 155Mb/s path.

In the common-view example shown in Fig. 6, the time and frequency signal is generated by TFKU, called the reference clock in ITU-T, and then distributed to multiplexers/demultiplexers through the TFKU called the slave clock in ITU-T. The time and frequency signal is immediately transferred to the TFKU in connected nodes through clock IF and STM-1 IF from the slave clock. Each multiplexer/demultiplexer adds general STM-1 signals to STM-16 and drops them from STM-16, and also adds the time and frequency signal to STM-1.



(a) Configuration in a network node



Time and frequency signal paths

STM-N signal including time and frequency signals

General information

Time and frequency signal

(b) Configuration from a network node to custromer premises equipment

Fig. 5 - System configuration

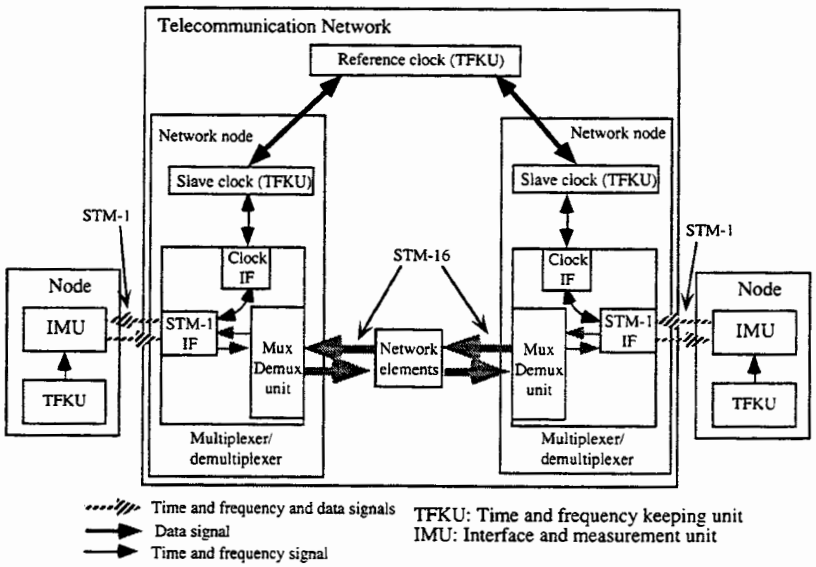


Fig. 6 - Common-view example

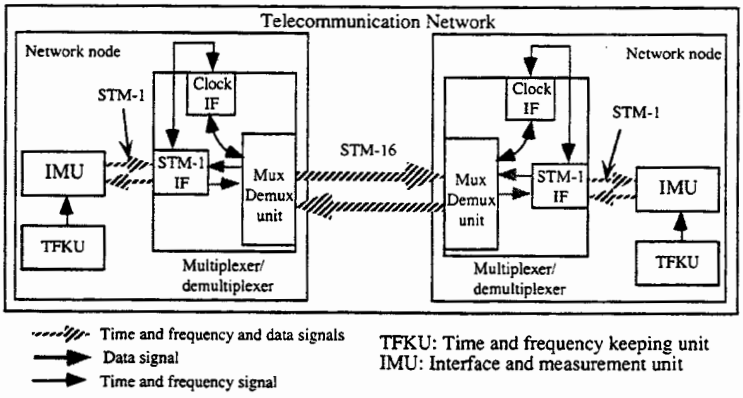


Fig. 7 - Two-way example

3) Two-way example

If there is no hierarchy between two nodes, the two-way configuration can be used. Nodes in the mutual topology can exchange time and frequency signals generated by their TFKUs. Time difference data measured at each node can be transferred to each other via the time and frequency path or the 155Mb/s path.

Figure 7 shows the two-way example. Each multiplexer/demultiplexer adds the time and frequency signal to STM-16 and drops the time and frequency signal from STM-16 other than general STM-1 signals. Information imposed on the MSOH described in 2.2 is terminated at the end of a section. This means that the MSOH information can be transmitted only between multiplexers/demultiplexers and between multiplexer/demultiplexer and the node. The MSOH information in STM-1 generally can not be transported to the MSOH in STM-16, nor can the MSOH information in STM-16 to the MSOH in STM-1. In this example, the time and frequency signal generated by TFKU is extracted in STM-1 IF, transferred to the Mux/demux unit via Clock IF, and then imposed on the MSOH in STM-16. In the other network node, the time and frequency signal is extracted at the Mux/demux unit, and transferred to STM-1 IF via Clock IF, and then imposed on MSOH in STM-1 IF again. The time and frequency signal thus can be directly transferred to the other node by this configuration.

There is no network element such as multiplexers/demultiplexers between network nodes in this example. STM-16 can be transmitted via other multiplexers/demultiplexers if the original clock of STM-16 is not changed; howerver, it is required that SOH bytes carrying time and frequency information included in STM-16 are tranferred from one section to the other section at intermediate multiplexers/ demultiplexers.

In this example, TFKU is located in the network node, as is the multiplexer/demultiplexer. However, a subscriber network can be used to connect TFKU to the multiplexer/demultiplexer if TFKU is located in a node different from the network node.

5. CONCLUSION

A time and frequency transfer system using SDH-based interfaces will be introduced as the next generation synchronization system. To realize time and frequency transfer in telecommunication networks, standardization is the most serious issue. First, carrying time information in the finite SOH bytes should be discussed in ITU-T.

ACKNOWLEDGMENTS

We are grateful to Dr. Takao Matsumoto of NTT Optical Network Systems Laboratories for his encouragement.

REFERENCES

[1]ITU-T Recommendation:"Synchronous Digital Hierarchy Bit Rates", G.707, ITU-T

[2]ITU-T Recommendation:"Network node interface for the synchronous digital hierarchy," G.708

[3]ITU-T Recommendation:"Synchronous multiplexing structure", G.709, ITU-T

[4]ITU-R Working Party 7A:"Time and frequency transfer using digital telecommunication networks", Doc. 7A/5, Nov. 1994

[5]ITU-R Working Party 7A:"Liaison statement from Working Party 7A to ITU-T Study Group 13", Doc. 7A/6, Nov. 1994

RELATIVISTIC THEORY FOR THE REALIZATION OF COORDINATE TIME SCALES AND THE SYNTONIZATION OF CLOCKS IN THE VICINITY OF THE EARTH

P. Wolf [1,2] and G. Petit [1]

[1]Bureau International des Poids et Mesures
Pavillon de Breteuil, 92312 Sèvres CEDEX, France
[2]Queen Mary and Westfield College, School of Mathematical Sciences
Mile End Road, London E1 4NS, Great Britain

Abstract: A well known prediction of Einstein's general theory of relativity states that two ideal clocks that move with a relative velocity, and are submitted to different gravitational fields will in general be observed to run at different rates. Similarly the rate of a clock with respect to the coordinate time of some space-time reference system is dependent on the velocity of the clock in that reference system and on the gravitational fields it is submitted to. For the syntonization of clocks and the realization of coordinate times (like TAI) this rate shift has to be taken into account at an accuracy level which should be below the frequency stability of the clocks in question, i.e. all terms that are larger than the instability of the clocks should be corrected for. We present a theory for the calculation of the relativistic rate shift for clocks in the vicinity of the Earth, including all terms larger than one part in 10^{18}. This, together with previous work on clock synchronization (Petit & Wolf 1994a, b), amounts to a complete relativistic theory for the realization of coordinate time scales at picosecond synchronization and 10^{-18} syntonization accuracy, which should be sufficient to accomodate future developments in time transfer and clock technology.

1. Introduction

When using the concept of syntonization in a relativistic context certain ambiguities might appear which can lead to confusion and misunderstanding. It is therefore essential to first clarify the different meanings of the expression as used in time metrology within a relativistic framework.

For the realization of coordinate time scales (like TAI) it is necessary to syntonize clocks with respect to the coordinate time in question, ie. to determine the rate of a clock A with respect to an ideal coordinate time of some space-time reference frame. For example, using a geocentric non-rotating frame with TCG as coordinate time (as defined by the IAU (1991)), in the first post-Newtonian approximation, the proper time τ_A of the clock A is related to TCG by

$$d\tau_A/dTCG = 1 - [U(\mathbf{w}) + v^2/2]/c^2 + O(c^{-4}) \qquad (1)$$

where $(cTCG, w^k)$ are coordinates in the geocentric frame with $\boldsymbol{w}$ representing the triplet w^k. The potential at the position of the clock $U(\boldsymbol{w})$ is the sum of the Earth's potential and tidal potentials of external bodies, and $v = ((dw^i/dTCG)(dw^i/dTCG))^{1/2}$ is the coordinate speed of the clock in the geocentric, non-rotating frame. Note that this rate depends entirely on the chosen reference frame. It is a coordinate quantity which cannot be obtained directly from measurement, but must be calculated theoretically using the definition of the reference frame in question with the appropriate metric equation.

When using repeated time transfers employing the convention of coordinate synchronization (Allan & Ashby 1986, Petit & Wolf 1994) for the determination of the relative rate of two clocks A and B the resulting rate predicted by theory is simply

$$d\tau_A/d\tau_B = (d\tau_A/dTCG)(dTCG/d\tau_B) \qquad (2)$$

with $d\tau/dTCG$ given in (1). This is a combination of coordinate dependent quantities and therefore entirely dependent on the chosen reference frame and the convention of synchronization.

The value of the quantity (2) is not to be confused with the result obtained by an observer measuring the relative frequency of two signals emitted by A and B, each signal being locked to the clocks in A and B. This quantity, $[f_A]_O/[f_B]_O$, is a measurable quantity which is specific to the measuring observer O. It is possible to relate the proper received frequency $[f_A]_O$ at the time of reception t_r to the proper emitted frequency $[f_A]_A$ at the time of emission t_e by the formula

$$[f_A]_O / [f_A]_A = d\tau_A/d\tau_O \times (1+\boldsymbol{k}_A \cdot \dot{\boldsymbol{w}}_O(t_r)/c)/(1+\boldsymbol{k}_A \cdot \dot{\boldsymbol{w}}_A(t_e)/c) \qquad (3)$$

where, in the formulation of the Doppler effect, the Newtonian approximation is used with $\boldsymbol{k}_A$ being the unit vector along the propagation path, $\boldsymbol{k}_A = (\boldsymbol{w}_A(t_e)-\boldsymbol{w}_O(t_r))/|\boldsymbol{w}_A(t_e)-\boldsymbol{w}_O(t_r)|$. Applying this formula to A and B, it is clear that $[f_A]_O/[f_B]_O$ differs from the value given by (2) by the Doppler terms which are dependent on the relative motion of the clocks and the observer.

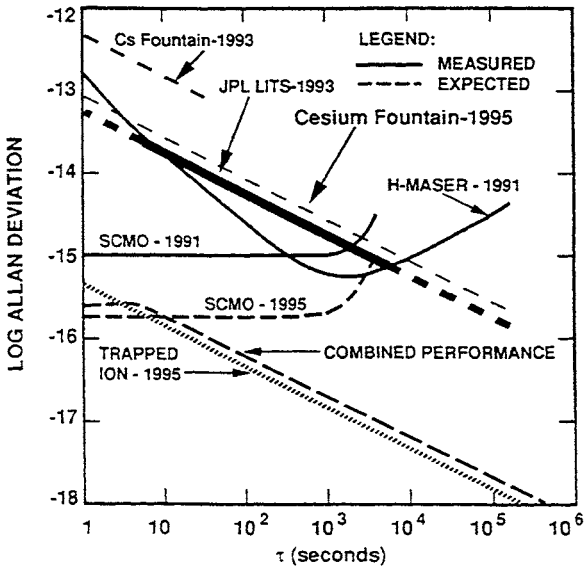


Fig. 1: Present and expected clock stabilities (from Maleki (1993)).

In sections 2 and 3 we will consider the syntonization of clocks with respect to coordinate times TCG (Geocentric Coordinate Time) and TT (Terrestial Time, the ideal form of TAI) as defined by the IAU (1991) together with the transformation relating the two. The aim is to provide expressions in the form of (1) including all terms whose magnitudes exceed current and near future clock stabilities which are estimated to reach parts in 10^{18}, as shown in figure 1 (Maleki 1993).

When determining the relative rate of two distant clocks one might be interested in time varying effects only (ie. effects that influence the observed frequency stability), which, as will be shown, can be calculated at higher accuracies than constant frequency shifts. They are discussed briefly in section 4.

2. Syntonization with respect to TCG

Using the metric given in resolution A4 of the IAU (1991) the relation between the proper time of a clock τ and TCG can be expressed as,

$$d\tau/dTCG = 1 - [U_E(\mathbf{w}) + \overline{U}(\mathbf{x}_E+\mathbf{w}) - \overline{U}(\mathbf{x}_E) - \overline{U}_{,k}(\mathbf{x}_E)w^k + v^2/2 + Q_k w^k]/c^2 + O(c^{-4}) \quad (4)$$

where coordinates in the barycentric frame are represented by $(cTCB,x^k)$ with $\mathbf{x}$ denoting the triplet x^k and the subscript E refering to the Earth's centre of mass. $U_E(\mathbf{w})$ and $\overline{U}(\mathbf{x})$ are the Newtonian gravitational potentials of the Earth and of external masses respectively, $v = ((dw^i/dTCG)(dw^i/dTCG))^{1/2}$, the coordinate speed of the clock in the geocentric, non-rotating frame and Q_k is the correction for the non-geodesic barycentric motion of the Earth.

We find that in the vicinity of the Earth the term in Q_k and terms of order c^{-4} (given explicitly in Brumberg & Kopejkin 1990 and Kopejkin 1988) amount to a few parts in 10^{19} or less. This implies that the specification of coordinate conditions (harmonic, standard post-Newtonian etc...) and the state of rotation of the frame (kinematically or dynamically non-rotating) is not significant for syntonization at the 10^{-18} accuracy level.

All effects that need to be taken into account for the calculation of the remaining terms are listed in tables Ia and Ib, together with orders of magnitude and present day uncertainties of the associated corrections.

Effect	Order of magnitude	Uncertainty
Earth's grav. potential	7×10^{-10}	10^{-17}
Centrifugal potential	1×10^{-12}	$< 10^{-18}$
Volcanic and coseismic (highly localised)	$< 10^{-16}$	
External masses (moon, sun)	10^{-17}	$< 10^{-18}$
Solid Earth tides	10^{-17}	$< 10^{-18}$
Ocean tides	10^{-17}	$< 10^{-18}$

Table Ia: Effects on syntonization with respect to TCG of clocks on the Earth's surface; Orders of magnitude and uncertainties of the corrections.

Syntonization with respect to TCG of Earth-bound clocks is limited at the 10^{-17} accuracy level by uncertainties in the determination of the potential of the Earth at the location of the clock. Hence only effects whose influence on (4) is larger than this limit are considered in table Ia.

Effect	Order of magnitude	Uncertainty
Earth's grav. potential	$< 6\times10^{-10}$	few 10^{-18}
		$< 10^{-18}$ (a>15000 km)
2nd order Doppler	$< 3\times10^{-10}$	$< 10^{-18}$ (a>15000 km)
External masses: Moon	4×10^{-13}	$<10^{-18}$
(w = 300000 km) Sun	4×10^{-14}	$<10^{-18}$
Venus	6×10^{-18}	$<10^{-18}$
Solid Earth tides	10^{-18}	$<10^{-18}$
Ocean tides	(at low altitudes)	
Polar motion		
Atmospheric pressure		

Table Ib: Effects on syntonization with respect to TCG of clocks on board terrestial satellites; Orders of magnitude and uncertainties of the corrections.

The gravitational potential of the Earth, $U_E(\mathbf{w})$ can be expressed as a series expansion in spherical harmonics. However, owing to mass irregularities such a series must be considered divergent at the surface of the Earth (Moritz 1961). Nonetheless, due to the predominantly ellipsoidal shape of the Earth, one can use the first two terms of this series expansion as a first approximation (Allan & Ashby 1986, CCIR 1990, Klioner 1992). Thus,

$$U_E(\mathbf{w}) = GM_E/w + GM_E a_1^2 J_2 (1-3\cos^2\theta)/2w^3 + \ldots\ldots \quad (5)$$

where G is the Newtonian gravitational constant, M_E is the mass of the Earth, a_1 and J_2 are respectively the equatorial radius and the quadrupole moment coefficient of the ellispoidal model of the Earth and θ is the geocentric colatitude of the point of interest. Substituting (5) into the second term of (4) gives terms of the order of 7×10^{-10} and 8×10^{-13} respectively for points on the surface of the Earth.

The surface obtained when setting $U_E(\mathbf{w}) = W_0 - (\omega w\sin\theta)^2/2$ in (5), with ω being the rotation rate of the Earth and W_0 the potential (gravitational + centrifugal) on the geoid, differs from the ellipsoid of the Earth model by less than 10 m. Hence an estimate of the accuracy of (5) can be obtained by considering the maximal difference between the geoid and the reference ellipsoid which can amount to ≈100 m (Vanicek & Krakiwsky 1986). Therefore expression (5) for the Earth's gravitational potential should not be used if accuracies better than one part in 10^{14} are required. Several procedures can provide a better accuracy, in the range 10^{-14} down to a few parts in 10^{17}, as the one presented below.

On the coast the mean sea level can be determined using a tidal gauge. This level differs from the geoid by what is known as Sea Surface Topology (SST) which can amount to ±0.7 m (Torge 1989). The SST can be determined with an accuracy of ≈0.1 m (Torge 1989) using oceanographic methods and satellite altimetry which induces an uncertainty of $\approx1\times10^{-17}$ in the second term of (4). The uncertainty in the knowledge of the potential on the geoid W_0, which is of the order of ± 1 m^2/s^2 (Bursa 1992, 1993), contributes another part in 10^{17}. The potential difference between mean sea level and an arbitrary point far from the coast can be obtained by geometrical levelling with simultaneous gravimetric measurements. The accumulated uncertainty when using modern levelling techniques and gravimetry is below 0.5 mm/√km (Kasser 1989) and does therefore not exceed a few centimeters even over large distances. In many countries levelling networks have been established at accuracies of 1-2 mm/√km for primary points, the use of which would again induce errors at the centimetric level.

Therefore the constant part of the gravitational + centrifugal potential at any point on the Earth's surface can be determined with an accuracy of order 2.5 m^2/s^2 using a tidal gauge and good geometrical levelling. The main contributions to this uncertainty are due to inaccuracies in the determination of W_0 and the SST. This limits the calculation of (4) at the level of $(2\text{-}3)\times10^{-17}$ which sets the limit for syntonization of clocks with respect to coordinate time (TCG or TT) on the surface of the Earth.

For clocks on board satellites, uncertainties in the model of the potential, such as GEM-T3 (Lerch et al. 1992), and in the determination of the satellite orbit (5 cm seems a realistic value) limit the accuracy of syntonization at a few parts in 10^{18} for low altitudes (semimajor axis < 15000 km). The uncertainties decrease at higher altitudes and, above 15000 km, all terms necessary for the syntonization with respect to TCG of clocks on board satellites can be calculated to an accuracy better than one part in 10^{18}.

3. Transformation to TT

TCG is related to TT by a relativistic transformations, hence any clock that is syntonized with respect to TCG can also be syntonized with respect to TT. In this case the accuracy of syntonization may be limited by the uncertainty in the determination of the parameters participating in the transformation.

The IAU defined TT as a geocentric coordinate time scale differing from TCG by a constant rate, the scale unit of TT being chosen so that it agrees with the SI second on the geoid (IAU 1991). TT is an ideal form of the International Atomic Time TAI, apart from a constant offset.

It can be obtained from TCG via the transformation

$$dTT/dTCG = 1 - L_g \qquad (6)$$

with $L_g = W_0/c^2 = 6{,}9692903\times10^{-10} \pm 1\times10^{-17}$.

It follows that at present the accuracy of syntonization with respect to TT is limited at the 10^{-17} level due to uncertainties in the determination of the potential on the geoid W_0, even for clocks on board terrestial satellites.

This limit is inherent to the definition of TT and can therefore only be improved by a reduction of the uncertainty in the determination of W_0. If highly stable clocks on board terrestial satellites are to be used for the realization of TT at accuracies exceeding this limit it might proove necessary to change the definition. One possibility would be to turn L_g into a defining constant with a fixed value, which would at the same time provide a relativistic definition of the geoid (Bjerhammar 1985, Soffel et al. 1988).

4. Time varying effects

For several applications of highly stable clocks, one is interested in the stability of the relative rate between two clocks, and therefore only time varying effects need to be considered, which can be calculated at the 10^{-18} accuracy level even for clocks on the surface of the Earth. Table II gives a summary of all such effects estimated to exceed the 10^{-18} limit.

Effect	Order of magnitude	Uncertainty
Volcanic and coseismic (highly localised)	$< 10^{-16}$	
Geodynamic and man-made (localised, long-term)	$< 10^{-16}$	
External masses (moon, sun)	10^{-17}	$< 10^{-18}$
Solid Earth tides	10^{-17}	$< 10^{-18}$
Ocean tides	10^{-17}	$< 10^{-18}$
Polar motion (long-term ≈430 days)	10^{-18}	$< 10^{-18}$
Atmospheric pressure	10^{-18}	$< 10^{-18}$

Table II: Time varying effects on the Earth's surface for the determination of the relative rate of two clocks; Orders of magnitude and uncertainties of the corrections.

Volcanic, coseismic, geodynamic and man-made (e.g. exploitation of oil, gas, geothermal fields) effects are highly localised and only need to be taken into account at some particular locations.

Polar motion and tidal effects are of periodic nature with essentially diurnal and semi-diurnal tidal periods, and the Chandler period (430 days) for the movement of the pole. If the clocks in question are syntonized using repeated time transfers (see (3)) at picosecond accuracy, tidal terms can be neglected as their short periods prevent their amplitudes in the time domain from reaching one picosecond (Klioner 1992).

For atmospheric pressure variations of ±10 mbar on a global scale (corresponding to seasonal changes), the effect on the rate of a clock on the Earth's surface can reach ±2 parts in 10^{18} with local pressure changes ((anti)cyclones with pressure variations of up to ±60 mbar) giving rise to a correction of up to $\pm2.7\times10^{-18}$ (Rabbel and Zschau 1985).

5. Conclusion

We have presented a theory for the syntonization of clocks with respect to Geocentric Coordinate Time (TCG) including all terms greater than 10^{-18} for clocks on board satellites at altitudes exceeding 15000 km. For this purpose terms of order c^{-3} and c^{-4} in the metric can be neglected, which implies that the specification of coordinate conditions and the state of rotation of the reference system is not necessary.

Syntonization with respect to Terrestial Time (TT), an ideal form of TAI, is limited at the 10^{-17} accuracy level due to the uncertainty in the determination of the potential on the geoid W_0 inherent to its definition.

For clocks on the Earth's surface syntonization with respect to TCG or TT is limited at an accuracy of $(2\text{-}3)\times10^{-17}$ by uncertainties in the determination of the geopotential at the location of the clock.

We briefly discussed time varying effects that may influence the stability of the relative rate of two clocks. These can be calculated at the 10^{-18} accuracy level even for clocks on the Earth's surface.

At present atomic clocks are approaching stabilities of the order 10^{-16} (Maleki 1993) with further improvements expected in the near future. For comparisons of these highly stable clocks over large distances, and their application in experimental relativity, geodesy, geophysics etc.. a sufficiently accurate relativistic theory for their syntonization, like the one presented in this paper, seems indispensable.

Together with a previous paper (Petit & Wolf 1994) the results obtained here amount to a complete relativistic theory for the realization of a geocentric coordinate time scale at a synchronization and syntonization accuracy of one picosecond and 10^{-18} respectively.

References

Allan D.W. & Ashby N., 1986, in: Kovalevsky J., Brumberg V.A. (eds.), *Relativity in Celestial Mechanics and Astronomy*. Proceedings of the IAU Symposium No. 114 Leningrad 1985, Reidel, Dordrecht.

Bjerhammar A., 1985, Bull. Géod. **59**, 207.

Brumberg V. A. & Kopejkin, 1990, Celestial Mechanics **48**, 23-44.

Bursa M. et. al., 1992,, Studia Geoph. et Geod., **36**, 101-114.

Bursa M., 1993, I.A.G. Special Commission SCR - Fundamental Constants (SCFC), Circular No 5.

CCIR, 1990, International Radio Consultative Committee, Reports, Annex to Volume VII, 150.

IAU, 1991, IAU transactions Vol. XXIB, 1991, Proc. 21st Gen. Assembly Buenos Aires, Kluwer Acad. Publ., Dordrecht, Boston, London.

Kasser M., 1989, C. R. Acad. Sci. Paris, **309**, Série II, 695.

Klioner S.A., 1992, Celestial Mechanics and Dynamical Astronomy **53**, 81.

Kopejkin S.M., 1988, Celestial Mechanics **44**, 87-115.

Lerch F.J. et al., 1992, NASA Technical Memorandum 104555.

Maleki L., 1993, Proceedings 25th PTTI Meeting 1993, Nasa Conference Publication 3267, 549-560.

Moritz H., 1961, Österreichische Zeitschrift für Vermessungswesen **49**, 11-15.

Petit G. & Wolf P., 1994a, Astronomy and Astrophysics **286**, 971-977.

Petit G. & Wolf P., 1994b, Proceedings 8th EFTF, 407.

Rabbel W. & Zschau J., 1985, J. Geophysics **56**, 81.

Soffel M. et al., 1988, Manuscripta Geodetica **13**, 143.

Torge W., 1989, *Gravimetry*, de Gruyter, Berlin.

Vanicek P. & Krakiwsky E., 1986, *Geodesy the Concepts*, North Holland.

SESSION 15

SAW II

Chairman : J.M. HODE

RECENT ADVANCE IN ZnO PIEZOELECTRIC TRANSDUCER TECHNOLOGY (FILM DEPOSITION, CHARACTERISTICS, APPLICATIONS)

S.V. Kulakov, V.G. Nefedov, O.G. Gabaraev, O.B. Gusev, V.A. Zhukov

St. Petersburg State Academy of Aerospace Instrumentation
(67 B.Morskaia St., St.Petersburg, 190000, Russia)

Abstract: Research on ZnO thin - film piezoelectric transducers manufacturing still attracts attention. That is connected primarely with potential advantages of such transducers. This paper presents some results in ZnO thin - film transducers manufacturing obtained in St. Peterburg State Academy of Aerospace Instrumentation.

1. INTRODUCTION

From '60 to '80, there was a great interest in manufacturing piezoelectric transducers based on ZnO films. A lot of works were published both in the West and in the former Soviet Union. Even in 1980, Dr. Fred S. Hickernell pointed out in IEEE Ultrasonic Symposium that "The deposition of acoustic quality zinc oxide (ZnO) for application to bulk acoustic wave (BAW) and surface acoustic wave (SAW) devices has been in progress for approximately fifteen years with some very notable achievements taking place along the way" [1].

From the beginning of '80, the number of publications on this subject slightly went down, though the interest in this problem remains stable. Apparently, this decrease in publications can be explained by two reasons. The first one is disappointment among a number of researchers who were unable to achieve the promising results they had anticipated. The second one is the fact that a number of researchers did achieve some results suitable for commercial applications, and they do not want to make them public.

Nevertheless, some new interesting publications have recently appeared, in particular, the article by Japanese authors [2]. This means that the work on manufacturing high - quality ZnO thin - film piezoelectric transducers continues.

The research on ZnO thin - film piezoelectric transducers in St. Petersburg State Academy of Aerospace Instrumentation were started in the beginning of '80, primarily for manufacturing high performance acousto - optic devices. We are working on ZnO piezoelectric film deposition technology for industrial applications.

2. FILM PIEZOELECTRIC TRANSDUCER DEPOSITION

Electroacoustic transducers that convert electric signals into acoustic waves can solve the problem of information input into processors based on the principles of acoustoelectronics and acousto - optics. The tasks are frequency and optimum signal selection, spectral and correlation processing, multidimentional signal processing (i.g., from phased antenna arrays), information processing for multidimentional matrix problems, etc.

In general case, a piezoelectric transducer consists of metallic electrodes, a piezoactive element and bonding layers providing appropriate acoustic contact between the transducer and an acoustic substrate. Depending on the technology for piezoactive element manufacturing, piezoelectric transducers can be divided into plate and film ones. In the first case, a prepared piezoelectric is used as the active layer, and the necessary acoustic contact is obtained by thermal compression bonding, cold bonding, etc. In the second case, the piezoelectric layer is deposited by film technology, and its properties and thus transducer's properties are determined primarily by the parameters of the technological process.

Film transducers can be used to achieve the following results:

- To increase the carrier frequencies of generated acoustic waves to several tens gigahertz.
- To manufacture transducers with their area of up to 1.000 mm^2.
- To deposit transducers on curved surfaces:
- To combine all deposition processes in a single technological process.
- To use slightly modified conventional equipment for thin film deposition.

Unfortunately, ZnO film transducers have several disadvantages:

- Lower piezoelectric properties of the films in comparison with monocrystals;
- Lower value of electromechanical coupling coefficient of ZnO in comparison with $LiNbO_3$ widely used in plate transducers.

A film piezoelectric transducer is a layered structure consisting of metallic electrodes, a piezoelectric layer, and, if necessary, bonding, insulating, etc., layers.

The manufacturing of a workable transducer is a complex problem, and it includes:

- Deposition of film components with required properties;
- Providing a good acoustic contact between the film elements and acoustic substrate.

Many works are devoted to the problems of ZnO patterned film deposition, metallic layers depositions,

the study of film growth laws, and their characterization [1-7]. That is why we do not pay much attention to such questions. We only present some results obtained in our laboratory.

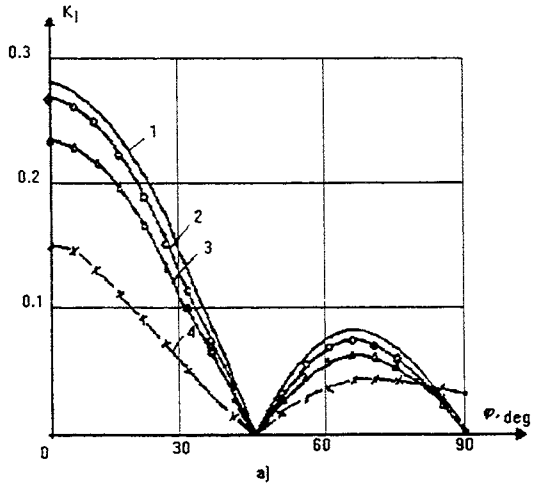
The integral characteristic primarily determining the performance of both the piezoelectrics and the transducer in whole is electromechanical coupling coefficient K. For polycrystal films, this coefficient is determined by the perfectness of the pattern, characterized by the following parameters:

- Pattern level R(hkl);
- Crystallite disorientation angle with respect to the pattern axis $\pm\sigma$;
- Pattern tilt angle with respect to the required position ϕ.

Fig. 1 shows the calculated values of electromechanical coupling coefficient for longitudinal Kl and shear Ks acoustic waves [8]. By controlling the pattern parameters, it is possible to obtain highly efficient conversion of electrical oscillations into longitudinal or shear acoustic waves. The presented data make us possible to formulate the requirements to the perfectness of the crystal pattern. For example, if 10% drop (in comparison with a monocrystal) in Kl is considered appropriate, pattern axis tilt angle can reach 15-20°, when crystallite disorientation angle is +6°. This means that the wide - spread opinion that the crystal structure should be close to monocrystal one is not always correct. The real value of the electromechanical coupling coefficient is more often determined by reciprocal orientation of the polar axes in adjacent crystallites that cannot be detected by X-ray structure analysis, but can be found by decrease in piezoelectric response with a piezoelectric tester.

The authors chose a DC ion - plasma technique in Ar-O_2 mixture for ZnO patterned film deposition. The target was a Zn disc with the diameter of 120 mm. Ar-O_2 ratio was 1:3. The deposition velocity was 0.1-2.0 μ m/sec.

As the experiments showed, the main factors determining the perfectness of the crystal pattern are: deposition velocity, the composition and pressure of the gas mixture, substrate temperature, deposition time, the geometry of substrate and target position.



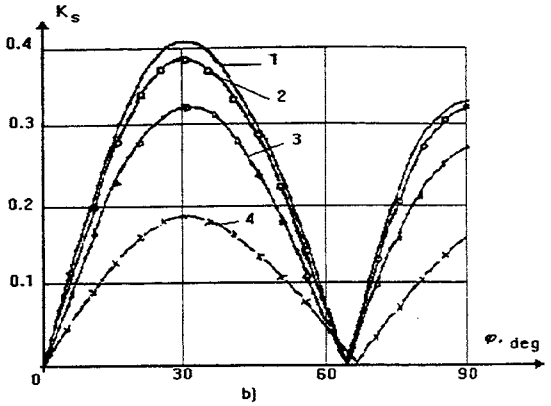
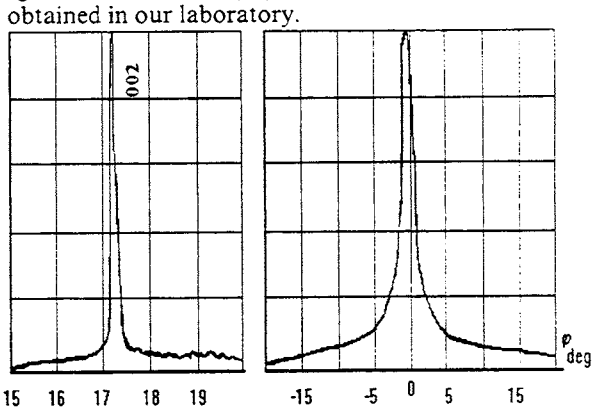


Fig. 1. The electromechanical coupling coefficient for longitudinal (a) and shear (b) acoustic waves: 1 - σ=0; 2 - σ=±6°; 3 - σ=±12°; 4 - σ=±18°.

Fig. 2 shows the X-ray grams of ZnO patterned films obtained in our laboratory.



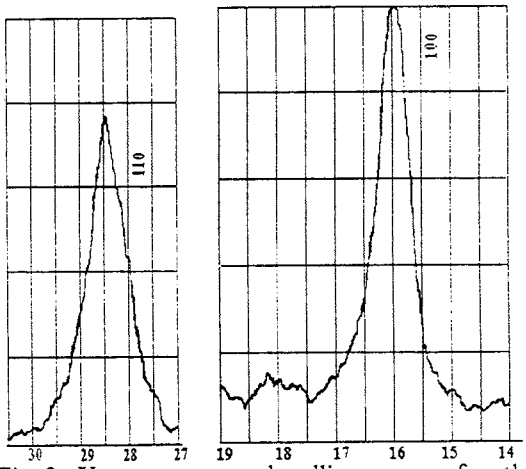


Fig. 2. X-ray grams and rolling curves for the films with pattern axis normal to the film plane (a) and lying in the film plane (b).

The pattern axis orientation was controlled by choosing an appropriate angle between substrate plane and the flux of atoms bombarding it. Decrease in disorientation angle for the films with the "lying" pattern was made by diaphragming in the flux of atoms bombarding the substrate and by forming a rectangular sputtering zone on the target.

The well - known advantage of the magnetron sputtering method is that it can be used for deposition metallic, semiconductor and dielectric films even applying direct current. We used these properties in order to utilize completely all possibilities of this method for transducer manufacturing.

A three magnetron sputtering system maintained in a single vacuum vessel was developed. Positioning a substrate in turns before the corresponding magnetron and varying the technological parameters (gas mixture composition, substrate temperature, electric discharge parameters, etc.), transducer's film structure was completely formed during a single technological process and in one technological space (Fig. 3).

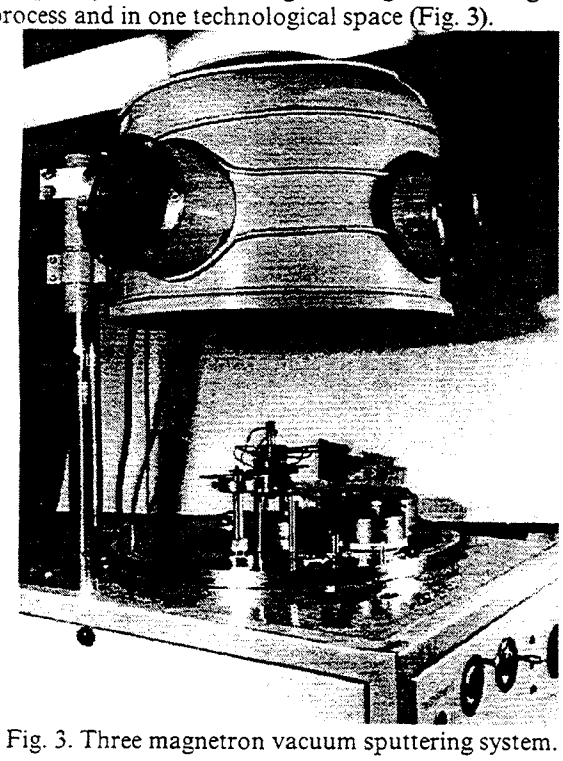


Fig. 3. Three magnetron vacuum sputtering system.

Al was used as a material for the metallic electrode. The quality of the Al film greatly depended on the deposition velocity and substrate temperature. The higher velocity, the higher grain size, more perfect crystal structure, the lower specific resistance. Nevertheless, large - grain structure did not allow us to obtain fine topological layout due to uneven etching. The target made of Cu-Si-Al alloy proved to be more suitable for quality film deposition. Cu reduces electric migration of Al and decreases surface roughness. Si improves Al adhesion. Under the substrate temperature of 160-200°C, Al has (111) orientation, the pattern axis being perpendicular to the substrate. The average crystallite size is 0.5 μm. In order to manufacture transducers, the films with the thickness of 0.3..5 μm were deposited under the deposition velocity of 2 nm/sec.

The main causes of faults in the acoustic contact are:

- Separation of transducer's film elements from each other or from the acoustic substrate due to weak adhesion;
- Cracking in the transducer or in the subsurface layer in the acoustic substrate because internal stresses exceed the ultimate stress in the used materials.

The magnetron technique used for deposition of all transducer's film elements considerably increases the adhesion due to higher energy of atoms bombarding the substrate (10 times higher than for thermal deposition) and can be done without special adhesion sublayers in most cases.

Film element deposition is usually done in unequilibrium conditions, which results in structural and thermal stresses in them. The structural stresses are connected with defects in crystal structure, etc. These stresses can be removed by annealing.

The thermal stresses are primarily determined by the difference in the coefficients of thermal linear expansion in contacted materials. The theoretical estimations of such thermal stresses appeared under the temperature difference of 250..300°C show that they are around 10..10 N/m. The authors measured stretching stresses of $(0.7..1.0)*10^9$ N/m in ZnO films, that is in good agreement with theoretical calculations [4]. These values are of the same order as the ultimate stresses for most materials used in transducer and acoustic substrate manufacturing (ZnO, Al_2O_3, $LiNbO_3$, SiO_2, TeO_2, $PbMoO_4$, GaAs, InP). Moreover, thermal stresses for TeO_2 and $PbMoO_4$ exceed their ultimate stresses. We believe that this fact was the reason for the authors of [9] to say that it was impossible to deposit ZnO film on these materials.

The research conducted in our Academy enabled us to manufacture a film piezoelectric transducer on such materials. We solved the problem by depositing a metallic layer separating the stresses in ZnO film and the acoustic substrate, and by using a two-stage process of film deposition. During the first stage of the process, a thick layer (0.3 μm) with good pattern is formed. During the second stage, the required thickness is achieved by depositing ZnO on this layer under room temperature. The initial layer is an orienting substrate. Thus there is almost no causes for thermal stresses. Using such technology, piezoelectric transducers made of ZnO film of 5..6 μm were deposited on TeO_2 crystals, which makes it possible to generate acoustic oscillations with frequencies of 0.5..0.7 GHz.

3. EXAMPLES OF PIEZOELECTRIC ZnO FILM USED IN THE SYSTEMS OF FREQUENCY SELECTION AND AS WIDEBAND PIEZOELECTRIC TRANSDUCERS FOR ACOUSTO - OPTICS AND ACOUSTOELECTRONICS

A resonator in a hybrid chip

A composite resonator for a chip containing a VHF generator was made on a sapphire substrate. The resonator was an etched substrate area with piezoelectric ZnO film of 4 μm thickness for longitudinal oscillation. The total thickness was determined by the condition of phase synchronism

and was approximately half-wave for the resonant frequency of 210 MHz. The measured Q of the resonator was 10,000.

Acoustic lens element for an acoustic microscope

An acoustic lens element was made as an acoustic module and electric module feeding the piezoelectric transducer of the lens element. The acoustic module was a cylindrical sapphire acoustic substrate with a spherical acoustic lens with a generating aperture of 150 μm. Using a DC magnetron sputtering, a ZnO flat transducer with Al film electrodes was made at the face opposite to the lens. The electric module was made as two micro coaxial lines with a shorted inductive stub between them. This module electrically matched the transducer and a standard RF feeder in a wide frequency band. The working aperture of the lens element was round with 0.8 mm diameter. The central frequency was 1.4 GHz. The bandwidth was 0.4 GHz, VSWR=2.5.

Acousto - optic multichannel cell

A multichannel acousto - optic cell was intended for RF signal input into the optical processor of acousto - optoelectronic systems of a radar with a phased antenna array. The base for the multielement piezoelectric transducer was ZnO film with Al electrodes deposited on a GaP acoustic substrate by using the magnetron method. The film had 15x35 mm area. Its thickness corresponded to the resonant frequency of 1.1 GHz. The bandwidth was 0.3 GHz, VSWR=3. The diffraction efficiency was 2%W. The number of channels was 10. The time aperture was 1 μ sec. This modulator is shown in Fig. 4.

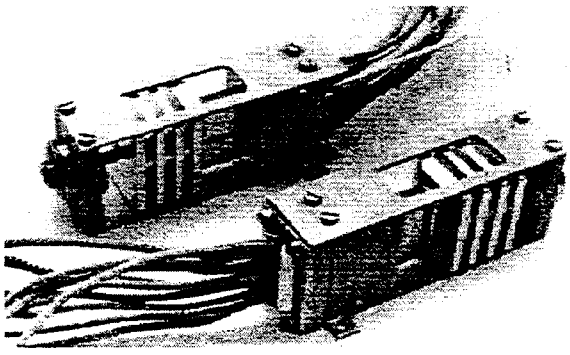


Fig. 4. Acousto - optic multichannel cell

A $LiNbO_3$ Bragg cell with anisotropic diffraction

The acoustic layout of this cell is presented in Fig. 5

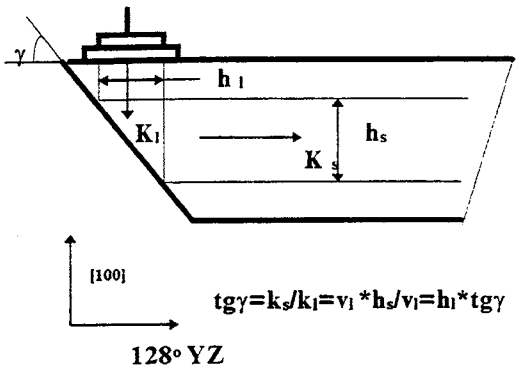


Fig.5

When this type of diffraction is implemented, the problem of efficient generation of the slow shear wave in the direction β (120°YZ) in $LiNbO_3$ in the frequency band of 2..3 GHz becomes topical. Plate shear piezoelectric transducers made of $LiNbO_3$ require the specific technology of ion etching, high quality of surface, etc. Nonresonance method of hypersonic generation does not show high efficiency. for those reasons, there was chose a ZnO film transducer and mode transformation from the longitudinal mode into the shear on. Table 1 presents the parameters of this cell. Experimental results were obtained by Prof. Victor Kludzin from our Academy.

Table 1

Cell material	$LiNbO_3$ crystal
Piezoelectric transducer	ZnO
Shear wave traveling direction	128°YZ
$V_{L\,[100]}$, cm/sec	$6.57*10^5$
V_S, cm/sec	$3.6*10^5$
Transformation angle, φ, °	61.4
Acousto-optic figure of merit for the longitudinal mode, M_{2L}, sec^3/g	$7*10^{-18}$
Acousto-optic figure of merit for the shear mode, M_{2S}, sec^3/g	$20*10^{-18}$
Longitudinal acoustic mode attenuation at frequency f=2GHz, α_L, dB/cm	~3.5
Shear acoustic mode attenuation at frequency f=2.5GHz, α_S, dB/cm	~20
Shear mode acoustic column sizes, Lxh, mm	2x1.5
3dB frequency band, GHz	2.0..2.5
Diffraction efficiency, P_{in}=1W, I_d/I_i	0.03

4. CONCLUSIONS

1. It is shown that the requirements to the pattern perfectness of ZnO film for piezoelectric transducers allow us to work with a standard equipment of chip manufacturing.

2. The problem of ZnO film deposition on TeO_2 and $PbMoO_4$, earlier considered to be unresolvable, is solved.
3. Acousto - optic devices manufactured with film transducers proved the fact that it is possible to form a transducer with large area and on curved surfaces.
4. It is proved that if it is necessary to manufacture a cell with bandwidth not wider than 0.3, ZnO films are not worse than other stronger piezoelectric materials.

5. REFERENCES

1. F.H. Hickernell, "ZnO Processing For Bulk- And Surface-Wave Devices". *Proc. of 1980, Ultrasonics Symposium.* v.1, 1980, Boston, pp 785-794.
2. M. Kadota, T. Kasanami, M. Minakata "Deposition and Piezoelectric Characteristics of ZnO Films by Using an ECR Sputtering System". *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control,* vol. 41, N4, 1994, pp 479-483.
3. N.F. Foster "Crystallographic Orientation of Zinc Oxide Films Deposited by Triode-Sputtering". *J. Voc. Sci. Tich.*, vol. 6, pp. 111-114, January, 1969.
4. T.D. Shermergor, N.N. Streltsova, *Film Piezoelectrics*, Moscow, "Radio i Sviaz", 1986. (in Russian)
5. A.I. Morozov, V.V. Proklov, B.A. Stankovskiy, *Piezoelectric Transducers for Radio Electronic Devices, Moscow*, "Radio i Sviaz", 1981. (in Russian)
6. B.V. Danilin, V.K. Syrchin, *Magnetron sputtering systems*, Moscow, "Radio i Sviaz", 1982. (in Russian)
7. *Design and Fabrication of Acousto-optic Devices.* Edited by A. Goutzoulis, D. Pape and S. Kulakov. Chapter 6 by V. Nefedov and D. Pape. Marcel Dekker, Inc. New Yourk,1994.
8. N.A. Kenigsberg, "Electromechanical Coupling Coefficients of ZnO Patterned Films", *Acousticheskiy Zhurnal*, Vol. 35, N. 2, 1989, pp. 368-370. (in Russian)
9. E.K. Kirchen and D.R. Jaeger "Today's capabilities of microwave (0,2 GHz to 18 GHz) acousto-optic devices" *SPIE Acousto-Optics.* San.Diego, Calif, 1976. Vol. 90.

SURFACE TRANSVERSE WAVES PROPAGATING ON PIEZOELECTRIC PLATES WITH THIN METAL STRIP GRATINGS ON TOP AND BOTTOM SIDE SIMULTANEOUSLY

S. Ballandras, E. Gavignet, E. Bigler

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

Abstract

A theoretical model taking piezoelectricity into account has been developed for the prediction of properties of Surface Transverse Waves (STW) excitation on piezoelectric plates of limited thickness supporting thin metal strip gratings deposited on one or both sides of the plate. A very general approach has been adopted for the model, in order to take into account all the parameters of the device (obstacles' shape, plate and obstacle thickness, shift of the pattern, etc.). Calculations have been performed for quartz plates in the vicinity of AT cut and Z cut, with plate thickness varying from 100 to 140 µm. The gratings are supposed to be symmetric. Periodicity of the gratings has been fixed to 20 µm. The dispersion curve obtained from the boundary condition equations emphasized the existence of more than only one stop-band as in classical STW devices. Higher frequency stop-bands are predicted, which would allow to design high frequency resonators or sensors using classical technology tools. Comparisons are proposed between behaviours of plates with single and double side gratings. First experimental results obtained on Z cut plate with a single side grating are reported. These quantitative results will be used for the design of high frequency STW devices.

Introduction

The propagation of Surface Transverse Waves (STW) under shallow groove or thin metal strip gratings on singly rotated quartz cuts appears to be one of the most promizing principle for the realization of high frequency devices like wide and narrow band filters [1] or ultra-stable resonators [2]. Indeed, it is now well-known that a STW device built on an AT-cut of quartz takes advantage of a high propagation velocity, close to the one of a shear bulk wave propagating in a BT-cut [3], together with a first order temperature compensation which has been experimentally emphasized [4]. Also, a strong effort has been developed to improve the first theoretical models proposed by Auld [5] and Plessky [6], particularly by including piezoelectric properties of the substrate [7, 8]. However, most of these models deal with STW propagating on semi-infinite substrates which corresponds to the practical case of a substrate very much thicker than the wavelength. Auld *et al.* have considered the case of an isotropic plate of finite thickness corrugated on the top side with shallow grooves [9]. Also, Danicki has studied the case of an isotropic plate of thickness close to the wavelength corrugated by shallow grooves of sinusoidal shape on both sides [10]. However, these different studies did not provide sufficient numerical results for clearly emphasizing the interest of such structures for the design of STW devices.

In this paper, a theoretical study of STW properties propagating on piezoelectric plates of finite thickness is proposed. Both cases of single side and double side corrugation are investigated. The corrugation considered is composed of thin metal strip gratings, which are widely used in the fabrication of STW devices. The mechanical and electrical boundary conditions are described and the domain of validity of the model is defined.

Calculations have been applied for AT-cut and Z-cut plates, wave propagation being along the rotated Z' axis, with aluminum strip gratings and results are reported under the form of dispersion curves relating the angular frequency ω to the wavenumber β_0. It is shown that the devices studied present more than one stopband as found in the usual analyses on semi-infinite substrates. Also, first experimental results obtained on a Z-cut plate with a single side corrugation are exposed and compared to the theoretical prediction. As a conclusion, explanations of the obtained experimental results will be proposed, which allow to propose further investigations for this new class of thin plate resonant devices.

Theoretical analysis

The paper deals with the propagation of shear transverse polarized waves on a plate of finite thickness. Then, according to Bloch-Floquet theorem in the case of periodic structures, the mechanical displacement u_1 and the electrical potential ϕ will be written as follows

$$u_1 = \sum_{n=-\infty}^{+\infty} A_n \, e^{-\alpha_n a_2} \, e^{-j\beta_n a_3} \, e^{j\omega t}$$

$$\phi = \sum_{n=-\infty}^{+\infty} B_n \, e^{-\alpha_n a_2} \, e^{-j\beta_n a_3} \, e^{j\omega t} \qquad (1)$$

where

$$\beta_n = \beta_0 + \frac{2\pi n}{\Lambda} \quad , \quad n = 0, \pm 1, \pm 2, \ldots$$

In these expressions, A_n and B_n are respectively the mechanical and electrical amplitudes of the wave, α_n is relative to the behaviour of the wave along a_2 (penetration depth), β_0 is the wavenumber along a_3 (propagation direction), Λ is the grating's period and ω the angular frequency. Since the grating is supposed to be infinite along a_3, an infinity of harmonic modes are assumed to be guided by the structure. Also, since the strips of the grating are supposed to be infinite along a_1, u_1 and ϕ will not depend on this variable.

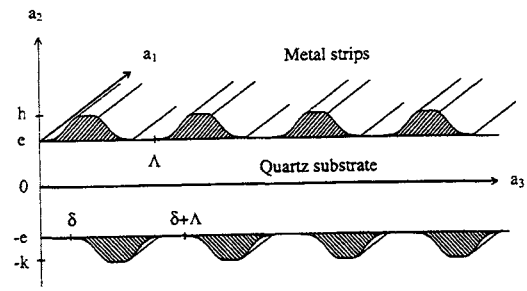


Fig. 1 : *Profile of a thin crystal plate with single or double side obstacle gratings*

The propagation of STW on a piezoelectric plate can be modelled by using piezoelectric propagation equation together with Poisson's condition for nonconductive media [8] written as follows

$$\rho_o \frac{d^2 u_1}{dt^2} = C_{55} \frac{\partial^2 u_1}{\partial a_3^2} + 2 C_{56} \frac{\partial^2 u_1}{\partial a_2 \partial a_3} + C_{66} \frac{\partial^2 u_1}{\partial a_2^2} + e_{35} \frac{\partial^2 \phi}{\partial a_3^2} + \left(e_{36} + e_{25} \right) \frac{\partial^2 \phi}{\partial a_2 \partial a_3} + e_{26} \frac{\partial^2 \phi}{\partial a_2^2} \quad (2)$$

$$e_{35} \frac{\partial^2 u_1}{\partial a_3^2} + \left(e_{36} + e_{25} \right) \frac{\partial^2 u_1}{\partial a_2 \partial a_3} + e_{26} \frac{\partial^2 u_1}{\partial a_2^2} - \varepsilon_{33} \frac{\partial^2 \phi}{\partial a_3^2} - 2 \varepsilon_{32} \frac{\partial^2 \phi}{\partial a_2 \partial a_3} - \varepsilon_{22} \frac{\partial^2 \phi}{\partial a_2^2} = 0 \quad (3)$$

where C_{IJ} and e_{iJ} are respectively the 2nd order elastic and piezoelectric constants in contracted indices, ε_{ij} are the dielectric coefficients and ρ_0 the mass density of the material.

Considering expressions (1), a general equation can be expressed relating ω, β_n and α_n. This equation will be written as a 4th degree polynomial in α_n, for a given set of values for β_n and ω

$$\alpha_n^4 (C_{66}\varepsilon_{22} + e_{26}^2) + \alpha_n^3 . 2j\beta_n \left(C_{66}\varepsilon_{23} + C_{56}\varepsilon_{22} + e_{26}(e_{36}+e_{25}) \right) - \alpha_n^2 \beta_n^2 \left(\varepsilon_{22}C_{55} + \varepsilon_{33}C_{66} + 4\varepsilon_{23}C_{56} - \varepsilon_{22}(\rho_o \omega^2 / \beta_n^2) + (e_{36}+e_{25})^2 + 2e_{26}e_{35} \right) - \alpha_n . 2j\beta_n^3 \left(C_{56}\varepsilon_{33} + C_{55}\varepsilon_{23} - \varepsilon_{23}(\rho_o \omega^2 / \beta_n^2) + e_{35}(e_{36}+e_{25}) \right) + \beta_n^4 \left(C_{55}\varepsilon_{33} + e_{35}^2 - \varepsilon_{33}(\rho_o \omega^2 / \beta_n^2) \right) = 0 \quad (4)$$

Finding the roots of this polynomial ensures the compatibility of equations (2) and (3). Then, for a given set of (β_n, ω), four values of α_n can be considered. In the case of semi-infinite substrates, the mechanical displacement and the electrical potential must be zero when $a_2 = -\infty$, implying that only two values of α_n will be physically representative (energy trapping condition along the surface). In the present case, there is no physical selective criterion for α_n and all the four values will be taken into account. This corresponds physically to waves bouncing up and down by total reflection along both sides of the plate. It implies that the mechanical displacement and the electrical potential will be composed of four partial waves for each mode n, yielding the following expression

$$u_1 = \sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} A_{nr} e^{-\alpha_{nr} a_2} \right) e^{-j\beta_n a_3} e^{j\omega t}$$

$$\phi = \sum_{n=-\infty}^{+\infty} \left(\sum_{r=1}^{4} \mathbb{C}_{nr} A_{nr} e^{-\alpha_{nr} a_2} \right) e^{-j\beta_n a_3} e^{j\omega t} \quad (5)$$

where the mechanical displacement and the electrical potential are related together across Poisson's condition (3), providing the following expression of the coefficient $\mathbb{C}_{nr}$

$$\mathbb{C}_{nr} = \frac{e_{35}\beta_n^2 - j(e_{36}+e_{25})\alpha_{nr}\beta_n - e_{26}\alpha_{nr}^2}{\varepsilon_{33}\beta_n^2 - 2j\varepsilon_{23}\alpha_{nr}\beta_n - \varepsilon_{22}\alpha_{nr}^2} \quad (6)$$

Equations (4), (5) and (6) need to be completed now by the boundary conditions in order to ensure the unicity of the solution of the problem.

Mechanical boundary conditions

First, the case of a double-side grating (Fig. 1) will be carefully exposed in a very general approach. The single side grating's case will appear as a simplification of this more complicated development.

The mechanical boundary condition description requires to define the dynamical stress conditions at the surface of the strip and also at the interface between the strip and the substrate. The top surface of the strip is supposed to be stress-free. This condition will be written according to the geometrical definitions of Fig. 2 as follows, with e equal to the half thickness of the plate

$$T'_{ij} . \nu_j'^{+} = 0 \quad at \quad a_2 = e + hf(a_3)$$
$$T'_{ij} . \nu_j'^{-} = 0 \quad at \quad a_2 = -e - kg(a_3) \quad (7)$$

In equation (7), the primed terms are related to the strip. $\nu_j'^{+}$ and $\nu_j'^{-}$ denote respectively the normal outside the strip surface on the top and the bottom side of the plate. h and k are respectively the thickness of the upper and lower strips and $f(a_3)$ and $g(a_3)$ describe their respective profile.

The expression of $\nu_j'^{+}$ and $\nu_j'^{-}$ are given below, assuming that h and k are both very smaller than the grating period Λ.

$$\nu_1'^{+} = 0 \qquad \nu_1'^{-} = 0$$
$$\nu_2'^{+} = 1 \qquad \nu_2'^{-} = -1$$
$$\nu_3'^{+} = -h \frac{\partial f(a_3)}{\partial a_3} \qquad \nu_3'^{-} = -k \frac{\partial g(a_3)}{\partial a_3} \quad (8)$$

It is then possible to develop equations (7) by using Taylor-McLaurin series, limited to the first order with respect to the strip thickness because h and k are both assumed to be very lower than Λ. As in the case of previous models developed for semi-infinite substrates [8], the propagation equation in the metal will be used to simplify the boundary condition equations. These equations will be written for both sides of the plate as follows where ρ' represents the mass density of the metal of the strip

$$T'_{12}(e, a_3) = h \left(f(a_3) [\rho' \omega^2 u'_1 (e, a_3) + \frac{\partial T'_{13}(e, a_3)}{\partial a_3}] + \frac{\partial f(a_3)}{\partial a_3} T'_{13}(e, a_3) \right)$$

$$T'_{12}(-e, a_3) = -k \left(g(a_3) [\rho' \omega^2 u'_1 (-e, a_3) + \frac{\partial T'_{13}(-e, a_3)}{\partial a_3}] + \frac{\partial g(a_3)}{\partial a_3} T'_{13}(-e, a_3) \right) \quad (9)$$

Now, the continuity equations at the interfaces between the metal and the plate will be defined. Indeed, they express the fact that the displacement and the nonzero stress component normal to the plate are equal from each side of the interface. These conditions yield the following expressions

$$u'_1(-e, a_3) = u_1(-e, a_3) \qquad u'_1(e, a_3) = u_1(e, a_3)$$
$$T'_{12}(-e, a_3) = T_{12}(-e, a_3) \qquad T'_{12}(e, a_3) = T_{12}(e, a_3) \quad (10)$$

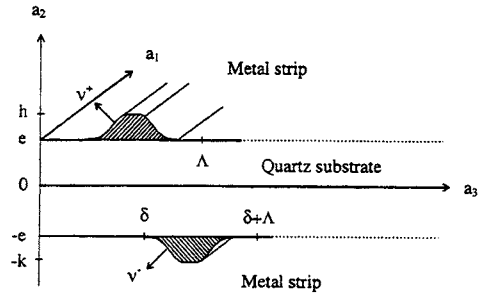


Fig. 2 : *Definition of the normal vectors outward the top and bottom metal strips*

Equations (5), (6), (9) and (10) will be used now to provide two compact expressions representative of the mechanical boundary conditions. These expressions are reported below, where C'_{44} is the thickness shear elastic constant of the isotropic metal

$$\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[C_{66}\,\alpha_{nr}+jC_{56}\beta_n+\mathbb{C}_{nr}(e_{26}\,\alpha_{nr}+je_{36}\beta_n)\right]\times e^{-\alpha_{nr}a_2}\,e^{-j\beta_n a_3}Anr\;\Big|_{a_2=e}$$

$$=-h\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[(\rho'\omega^2-C'_{44}\beta_n^2)f(a_3)-jC'_{44}\beta_n\frac{\partial f(a_3)}{\partial a_3}\right]\times e^{-\alpha_{nr}a_2}\,e^{-j\beta_n a_3}Anr\;\Big|_{a_2=e}$$

$$\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[C_{66}\,\alpha_{nr}+jC_{56}\beta_n+\mathbb{C}_{nr}(e_{26}\,\alpha_{nr}+je_{36}\beta_n)\right]\times e^{-\alpha_{nr}a_2}\,e^{-j\beta_n a_3}Anr\;\Big|_{a_2=-e}$$

$$=+k\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[(\rho'\omega^2-C'_{44}\beta_n^2)g(a_3)-jC'_{44}\beta_n\frac{\partial g(a_3)}{\partial a_3}\right]\times e^{-\alpha_{nr}a_2}\,e^{-j\beta_n a_3}Anr\;\Big|_{a_2=-e}\qquad(11)$$

It is possible to eliminate the dependance versus a_3 in equations (11) by using orthogonality properties of the different modes, as demonstrated by Auld [5]. Multiplying each term of equations (11) by $exp\ (j\beta_q a_3)$ and after integration over one grating's period along a_3, the final form of the mechanical boundary conditions are obtained

$$\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[Kqnr+\frac{\Lambda}{h}\delta_{nq}(F_{nr}+\mathbb{C}_{nr}Z_{nr})\right]e^{-\alpha_{nr}e}Anr=0\qquad(12)$$

$$\sum_{n=-\infty}^{+\infty}\sum_{r=1}^{4}\left[Lqnr-\frac{\Lambda}{k}\delta_{nq}(F_{nr}+\mathbb{C}_{nr}Z_{nr})\right]e^{+\alpha_{nr}e}Anr=0$$

with

$$Kqnr=\left(\rho'\omega^2-C'_{44}\beta_n\beta_q\right)I_{nq}$$

$$Lqnr=\left(\rho'\omega^2-C'_{44}\beta_n\beta_q\right)Jnq$$

$$Fnr=C_{66}\,\alpha_{nr}+j\,C_{56}\beta_n$$

$$Znr=e_{26}\,\alpha_{nr}+j\,e_{36}\beta_n$$

and

$$Inq=\int_o^{\Lambda}f(a_3)\,e^{-j(\beta_n-\beta_q)a_3}\,da_3$$

$$Jnq=\int_o^{\Lambda}g(a_3)\,e^{-j(\beta_n-\beta_q)a_3}\,da_3$$

A straightforward derivation shows that in the case of a single side corrugated plate, the mechanical boundary conditions will be obtained by setting L_{qnr} equal to zero in equations (12).

Electrical boundary conditions

A very precise description of the electrical boundary conditions arising at the surface of a STW device requires the development of complicated models [11]. However, a previous work [8] has demonstrated that simple approximations can provide satisfying results for weakly coupled piezoelectric materials (like quartz) by considering a condition of zero potential at the surface. Indeed, the academic case physically corresponding to a substrate covered by a thin metal layer (of negligible mass) electrically shorted offers significant simplification for operational modeling of STW propagation taking piezoelectricity into account. For both cases of single and double side grating plates, the electrical boundary conditions will be written as follows

$$\phi(+e,a_3)=0\ \ ;\ \ \phi(-e,a_3)=0\qquad(13)$$

which yields after development and simplification using orthogonality properties of the modes (as previously for mechanical boundary conditions) the final form of electrical boundary conditions

$$\sum_{r=1}^{4}\mathbb{C}qr\,Aqr\,e^{\alpha_{nr}e}=0$$
$$\sum_{r=1}^{4}\mathbb{C}qr\,Aqr\,e^{-\alpha_{nr}e}=0\qquad q=0,\pm1,\pm2\ldots\qquad(14)$$

It is shown that boundary conditions provide a set of 4 equations with 4 unknowns. The unicity of the solution will be ensured by the nullity of the determinant of this boundary condition system. It is then possible to solve completely the mathematical problem using numerical procedures.

For a given value of ω, different values of β_n and α_{nr} will be found starting from equations (4), (12) and (14). It can be pointed out that infinite Fourier series used in these equations must be troncated because of obvious practical reasons. Since only thin obstacle gratings are considered, two coupled modes ($n = -1$ and $n = 0$) will be sufficient to describe satisfyingly the properties of STW propagation [5-8]. This approach has been applied for the present work and calculations have been performed for particular configurations of gratings deposited on quartz plates.

Rectangular metal strip on quartz plates

Thin metal strip gratings can be easily deposited on crystal plates from a technological point of view, but the profile of the strips remains very difficult to control. However, most of the studies devoted to STW [5-8] have shown that the rectangular profile approximation is very easy to implement and appears to be well-suited for theoretical predictions of STW properties. Assuming that the profile of the strips on top and bottom side of the crystal plate are identical ($f(a_3) = g(a_3)$, $h = k$) and perfectly aligned, integrals Inq and Jnq will be equal, which will considerably simplify the mechanical boundary conditions in equation (12). The function $f(a_3)$ describing the profile of the strips will be expressed for one period Λ of the grating by using Heaviside functions as follows

$$f(a_3) = H(a_3) - H\left(a_3 - \frac{\Lambda}{2}\right) \qquad (15)$$

In the case of a misalignment of the 2 patterns for a double side grating device (cf. Fig. 2), integrals I_{nq} and J_{nq} will be related as follows

$$I_{nq} = J_{nq}\, e^{-j(\beta_n - \beta_q)\delta}$$

where δ represents the amount of misalignment in the a_3 direction.

Equation (15) will be used to calculate the integral I_{nq} versus the different values of n and q. After integration, the following results are obtained

$$I_{nq} = \begin{cases} \Lambda/2 & \text{for } n = q \\ 0 & \text{for } n-q \text{ even} \\ \dfrac{\Lambda}{j\pi(n-q)} & \text{for } n-q \text{ odd} \end{cases}$$

For a practical numerical resolution of the problem, different calculation algorithms are used in order to determine values of β_n and α_{nr} for a given value of angular frequency ω. Equation (4) is solved numerically and a modified Newton method has been adapted to the present problem for the calculation of β_n [8].

As indicated in introduction of the present paper, calculations have been performed for singly rotated quartz plates, in the vicinity of AT-cut, for which STW properties are particularly attractive (high propagation velocity, thermal compensation effects). Also, calculations have been done for Z-cut for two major reasons. First, the idea was to check if properties highlighted for STW on thin AT-cut plate can be obtained for other singly rotated quartz cuts. The second reason was the possibility to realize experiments using Z-cut plates in very short delays. The thickness of these plates was 140 µm, then different parameters of the design have been fixed, starting from this value. Particularly, the acoustic wavelength has been chosen equal to 40 µm, which provides a thickness to wavelength ratio of 3.5. Since classical STW decrease very slowly in respect with the depth of the substrate, such a value of ratio is suitable for the purpose of the present study. This implies that the grating period Λ will be equal to 20 µm, with a metal strip width set to 10 µm (cf. Fig. 2). A metal strip thickness of 150 nm has been used, which corresponds to a commonly used value, easy to perform from a technological point of view.

Because of the high number of parameters of the model, most of them have been fixed for theoretical predictions. However, for AT-cut plates, the plate thickness has been varied from 120 µm to 140 µm (by step of 10 µm) to emphasize STW property sensitivity to this parameter.

Theoretical results are presented in Fig. 3 under the form of dispersion curves relating the angular frequency ω to the imaginary part of the wavenumber β_0. These curves emphasize the existence of multiple frequency stopbands in the case of double side grating devices. Moreover, the position of the stopbands in regard with ω is clearly influenced by the thickness of the plate. The curves also show that the value of the imaginary part of β_0, which is only nonzero within the stopband is not constant versus the frequency and varies strongly within the frequency domain of analysis. A general tendency of variation indicates that this value increases together with the frequency but it can be remarked that for certain plate thicknesses, the distribution of β_0's imaginary part does not follow a regular behaviour.

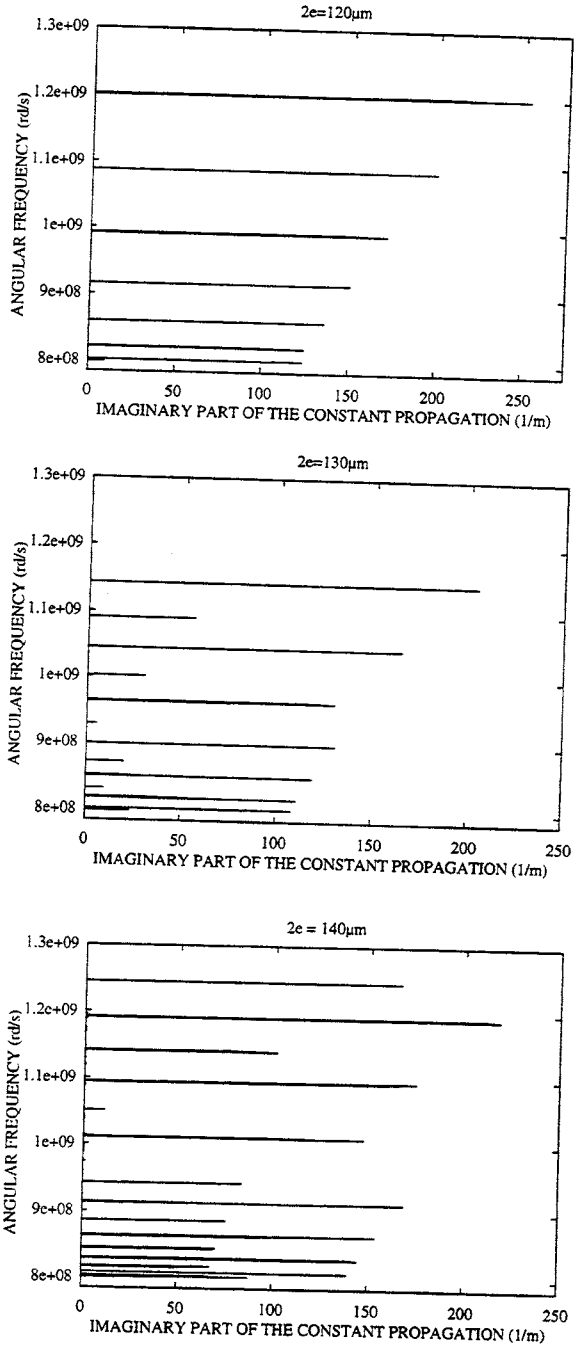


Fig. 3 : *Dispersion curves relating ω to βo for different thicknesses of AT-cut plage of quartz (double side metal strip grating)*

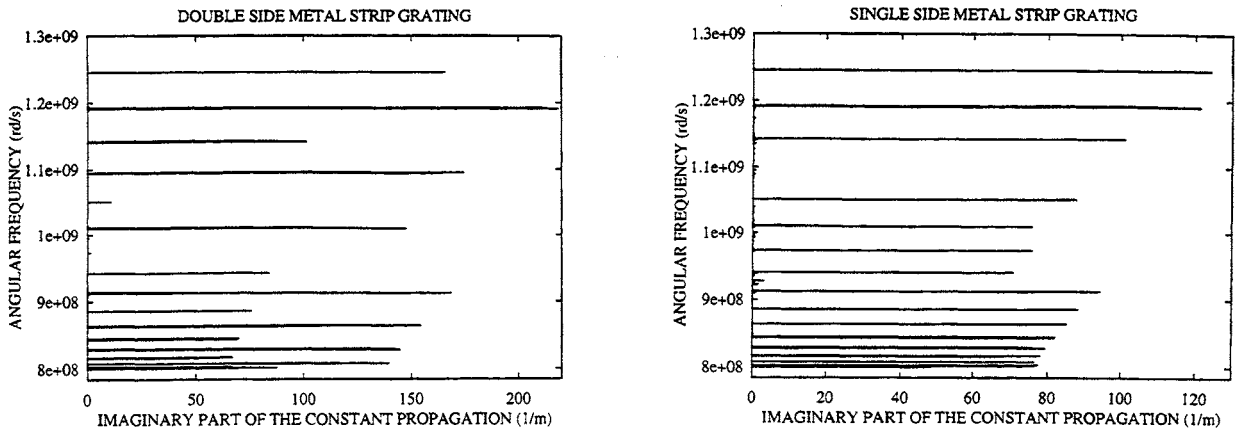


Fig. 4 : *Comparison between single and double-side metal strip grating devices built on an AT-cut plate of quartz of 140 µm thick*

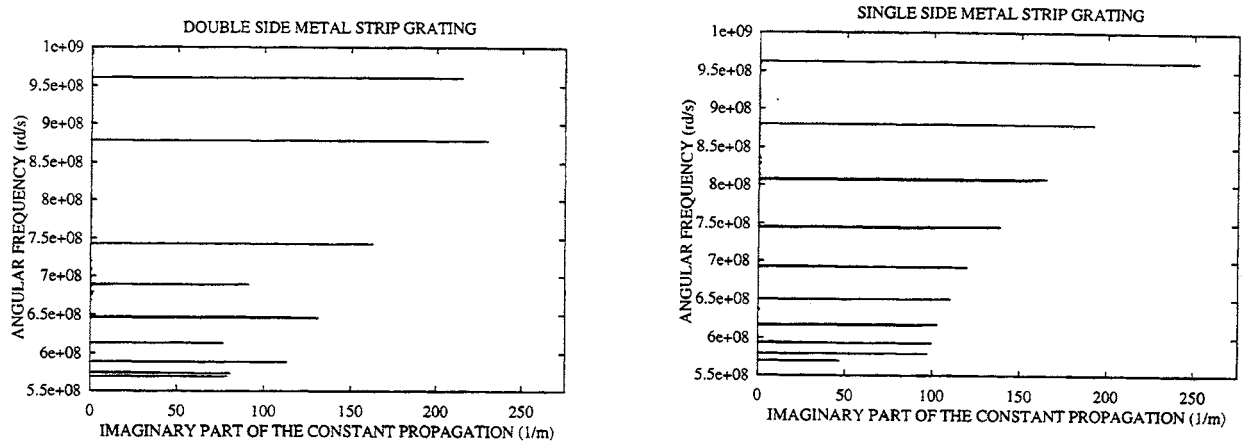


Fig. 5 : *Comparison between single and double-side metal strip grating devices built on an Z-cut plate of quartz of 140 µm thick*

The differences of STW properties propagating on plates with single or double side grating are emphasized on Fig. 4 for an AT-cut plate of 140 µm thick. It can be remarked that in this case, the single side grating device presents some more stopbands than the double side one. Most of the stopbands are common for both devices, but the value of imaginary part of β_0 is quite two times higher for the double side grating device than for the single one. This would yield a better energy trapping for the double side grating device because of a higher dispersion in the stopbands along propagation direction a_3. Calculations have been also performed for the same device built on a Z-cut plate of quartz. The previous conclusions about β_0 are no longer valid in this case when analyzing the dispersion curves in Fig. 5, but it can be remarked that also in this case, most of the stopbands are common to both devices. Elastic, piezoelectric and dielectric constants for quartz are those given by Slobodnik in Ref. [11].

First experimental results

Synchronous delay lines have been built on Z-cut plates of quartz in order to check the validity of the theoretical results obtained. The experimental device is composed of several interdigital transducers separated by gratings of different dimensions in order to test the largest number of configurations. For the first experiments only the single side grating configuration will be tested. The noncorrugated side has been locally covered with a thin metallic layer in order to emphasize the sensitivity of STW properties in thin plate to electrical boundary conditions.

For the present study, only one configuration has been tested which is shown in Fig. 6. Frequency responses have been measured for the device without grounding.

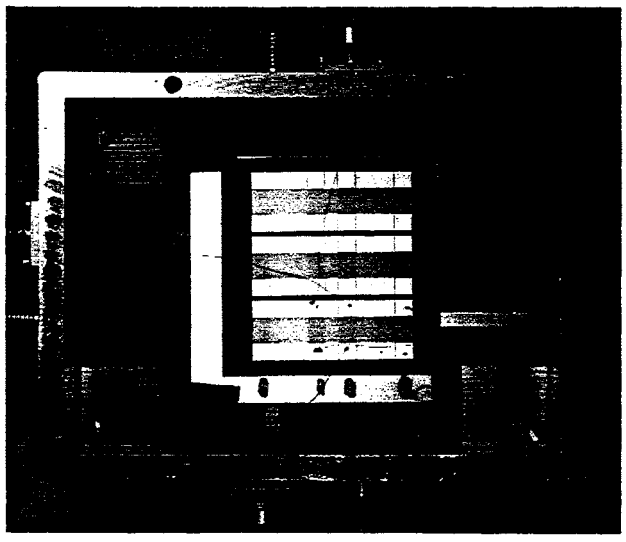


Fig. 6 : *Design of experimental STW synchronous devices built on Z-cut plate of quartz of 140 µm thick. The device is marked on the which has been tested*

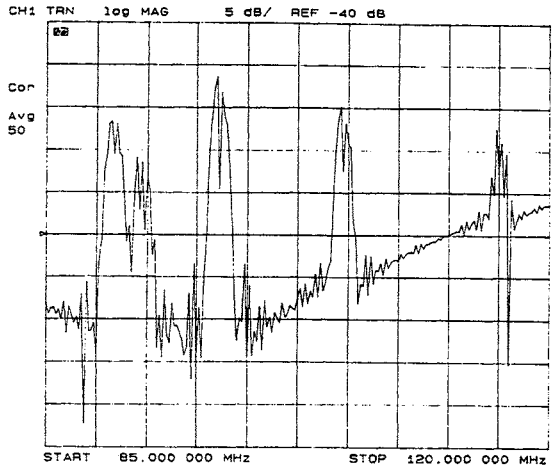


Fig. 7 : *Frequency responses obtained for the device presented in Fig. 6*

Results are plotted in Fig. 7. It is shown that for this configuration several modes are excited, which can be related to theoretical predictions as follows. Since the experimental devices are synchronous delay lines, the end of each frequency response groups would correspond to the beginning of a frequency stopband. It is then possible to compare these values together in order to check the relation between them. Table I allows this comparison, emphasizing some correlations between theoretical and experimental set of data. It can be remarked that some theoretical frequency stopbands are not found experimentally. This can be explained by the fact that excitation of all the modes which can propagate in the device requires more complicated transductor configuration. Futher experiments will allow to check this hypothesis.

Conclusion

An analysis of STW properties for propagation on thin piezoelectric plates with single or double side thin metal strip gratings has been proposed. A general model has been developed in this approach, which has been detailed in the paper. This model includes possibilities to vary a lot of design parameters like cut angle, thickness of the plate, periodicity of the grating, height of the obstacle, etc. For weakly coupled piezoelectric materials, the mass loading effect of the grating acts principally in the trapping of STW. Then, simplified electrical boundary conditions have been used in the present work. This approach provides satisfying results in regard with first experimental results obtained in the case of single side grating devices built on Z-cut plate of quartz. The multiple frequency stopbands theoretically emphasized can be experimentally found, as shown in the paper. This significates that the model presented here is reliable enough for the design of new STW devices for sensor or resonator applications. This study will be continued in measuring STW properties with single and double side grating on AT-cut plate of quartz. Explanation of the multiple frequency stopband phenomena can be proposed in relating these stopbands to plate modes which can be excited by an interdigital transducer on singly rotated quartz plates. The presence of grating would allow to trap these plate modes. Comparison between the frequency positions of the stopbands and the plate mode frequencies will allow to conclude about the relevance of the proposed explanations.

Table I
Comparison between experimentalfrequency responses and theoretical frequency stopband position

Order of frequency stopband	Theoretical values (MHz)	Experiments (MHz)
1st	90.6	90.4
2nd	92.0	92.2
3rd	94.3	--
4th	98.1	97.8
5th	110.2	106.4
6th	118.6	117.2

Acknowledgements

The authors would like to thank C. Bonjour and W. Daniau for their help in the preparation of experimental devices.

References

[1] T. Thorvaldsson, V.P. Plessky, S. Muckenhirm, M. Joray, "GHz range STW resonators and narrow band filters", Proc. 1994 Ultrasonics Symp., Cannes, France, nov.. 1994.

[2] T.L. Bagwell and R.C. Bray, "Novel surface transverse wave resonators with low loss and high Q", Proc. 1987 Ultrasonics Symp., IEEE cat. n° 87CH2492-7, pp. 319-324.

[3] S. Ballandras, E. Gavignet, E. Bigler, "Influence des effets de température quasi-statiques du premier ordre sur les dispositifs à ondes de surface transverses sur sillons gravés", J. Phys. III France, vol. 5, pp. 161-174 (1995).

[4] I.D. Avramov, "1 GHz voltage controlled oscillator using a SSBW/BGW combined mode resonator filter", Proc. 45th Ann. Freq. Cont. Symp., Los Angeles, june 1991.

[5] B.A. Auld, J.J. Gagnepain, M. Tan, "Horizontal shear waves on corrugated surfaces", Elect. Lett., vol. 12, n° 24, pp. 650-651 (1976).

[6] Yu.V. Gulyaev and V.P. Plesskii, Sov. Tech. Phys. Lett., vol. 3, n° 3, pp. 87-88, (march 1977).

[7] V.M. Bright and W.D. Hunt, "Analysis of Bleustein-Gulyaev wave propagation under thin periodic metal electrodes", J. Appl. Phys., vol. 70, n° 2, pp. 594-602 (1991).

[8] E. Gavignet, S. Ballandras, E. Bigler, "Theoretical analysis of surface transverse waves propagating on a piezoelectric substrate under shallow-groove or thin metal strip grating", to be published in J. Appl. Phys.

[9] B.A. Auld, A. Renard, J. Henaff, "STW resonances on corrugated plates of finite thickness", Electronics Letters, vol. 18, n° 4, pp. 183-184 (1982).

[10] E. Danicki, D. Bogucki, "Wave propagation and scattering in elastic plate with periodically grooved surface", Archives of Acoustics, vol. 18, n° 1, pp. 113-130 (1993).

[11] A.J. Slobodnik, E.D. Conway, R.T. Delmonico, "Microwave Acoustics Handbook, vol. 1A : Surface Wave Velocities", Air Force Cambridge Research Labs, 1973.

GREEN'S FUNCTION ANALYSIS OF THE MULTISTRIP COUPLERS FOR NARROW BAND LOW LOSS FILTERS

VICTOR P. PLESSKY AND SVETLANA N. KRASNIKOVA

Advanced SAW products SA, CH-2022 Bevaix, Switzerland
Dalniaia Sviaz, St. Petersburg, Russia

ABSTRACT

Green's function theory of multistrip couplers (MSC) is developed and compared with some known theoretical and experimental data. The approach is similar to one used in Pollock's and co-authors paper[1]. The difference between previous work and the present approach is that we directly use numerical data for effective dielectric permittivity not reducing it to simple Ingebrigsten approximate formulas. In this way the case of leaky waves can be considered with all main effects, like leaky losses, generation of the bulk waves and others included. We also implicitly formulate the electro-neutrality conditions and impose only the differences between the potentials of corresponding connected electrodes in two tracks to be equal.

A few examples are analyzed:

- *a simple shorted electrode reflector structure, as a reference point*
- *reflecting MSC with period of the electrode structure equal to half of SAW wavelength (2 electrodes per wavelength)*
- *reversing MSC with 3 electrodes per wavelength*

Our own measurements on reversing MSC are briefly described.

1. INTRODUCTION

In a new generation of low loss SAW filters for mobile communications, some well-known ideas can find a new implementation. The multistrip coupler (MSC), proposed many years ago by Marshall and Paige [2] is a good example of this kind. Recently number of authors have demonstrated very low loss filters based on the MSC and reversing multistrip coupler (RMSC) [1,3,4].

In coupled resonator filters (CRFs) very often two tracks are used , each one being a two port SAW resonator, and the resonators are cascaded in series using the interdigital transducers (IDTs). Normally the IDTs are designed so that their impedances are matched in the passband of the filter. The drawback of this scheme is that the transducers talk to one another in a wide frequency band , which results in a relatively high level sidelobe to the right hand side of passband. For many application this is a real difficulty. Using reflecting or reversing MSC this problem can be addressed adequately and the results [3,4] look very encouraging. The difference is that in the case of the MSC the signal is directed from the input transducers to the output only in the passband of the filter , while the IDTs, being bi-directional, are "talking" at all frequencies. The bulk waves have also lower chances to pass to the output transducers , when the MSCs are used.

There are a few papers dealing with the theory and experiments on MSC and reflecting (reversing) MSC. The first theory of MSC was proposed by Maerfeld and Tournois [5], and then generalized [6,7] for the reflecting MSC working at the Bragg frequency (2 electrodes per wavelength). Reflecting (reversing) MSC were proposed by Feldmann and Henaff [8] using different periods in the two tracks to fulfill synchronism conditions. The reversing multistrip couplers with the same period in two tracks were first proposed by Pollock and others [1] and then re-invented by Danicki [9]. In Pollock's paper [1] the Green's function approach was already used, but the choice of materials was limited to 128 LiNbO3, while here we are more interested in the "leaky-wave" cuts. The approximate description based on Ingebrigsten's formula, used in this paper is also too restrictive for the strong coupling materials and cannot give accurate results for the leaky waves. Finally, the increase of speed of the computers now makes even very intensive numerical calculations more acceptable than some years ago.

2. THEORETICAL BACKGROUND

The Green's function (see, e.g. [10]) relates the charges and the potentials on the surface of the piezoelectric media :

$$\varphi(q) = \frac{1}{|q|\varepsilon_{eff}(q/\omega)}\sigma(q) = G(q,\omega)\sigma(q) \qquad (1)$$

where

$$\sigma(q) = \int_{-\infty}^{+\infty}\sigma(x)e^{-iqx}dx$$

and

$$\varphi(q) = \int_{-\infty}^{+\infty}\varphi(x)e^{-iqx}dx$$

are the Fourier components of the charge distribution $\sigma(x)$ and the potential on the surface $\varphi(x)$. Time dependent factor is taken in the form $e^{-i\omega t}$ and is omitted throughout

this paper. The effective dielectric permittivity $\varepsilon_{eff}(q/\omega)$ depends only on "slowness" of the wave $1/V = q/\omega$ and can be calculated for a given cut of a crystal in advance only ones and then used as a known function. We used a software developed in [11] and material constants from [12] for calculations. Fig. 1 gives an example of the effective permittivity for 64-LiNbO3 - "leaky wave" cut. The Green's function $G(q,\omega)$ (or effective permittivity) contains complete information on all types of waves which are excited on the free piezoelectric surface by charges placed on it. So, our consideration is limited to the case of thin electrodes, although for the reversing MSC (see below) with 3 electrodes per period at the fundamental frequency mechanical reflections are absent.

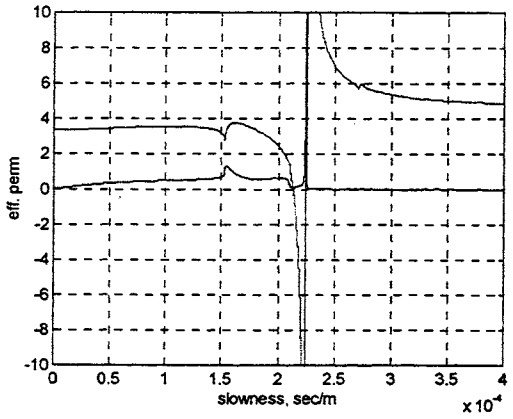


Fig. 1 Real and imaginary parts of effective dielectric permittivity for 64-LiNbO3

The zeros, poles, quazipoles and other peculiarities of the effective permittivity function directly correspond to the waves which can be excited in the substrate [10] . In particular, the velocity of the leaky wave on a free surface of 64-LiNbO3 can be found from the effective permittivity curve and is equal to $V_f = 4692\ m/\sec$ with complex zero for slowness $S_0 = 2.1313*(1+i9.38*10^{-4})*10^{-4}\ sec/m$ which corresponds to the attenuation of $5.9*10^{-3} neper/\lambda$ or $5.1*10^{-2}\ dB/\lambda$. These numbers are close to ones published earlier [13]. It means that if the wave passes an additional distance of 20λ, as in numerical examples discussed below, its amplitude decays at a factor of 0.89 or the propagation loss is *-1.02* dB. That must be taken into account while analyzing the numerical results below.

2.1. GEOMETRY AND CHARGE DISTRIBUTION

We consider here the simplest geometry of finite number *N* of thin metal (perfectly conducting) electrodes of width *l* placed periodically with distance *p* in one acoustic channel (for simple reflector) or two channels for the MSC. The electrodes in the second channel are supposed to have the same x-coordinates (wave propagation direction) and the same apertures as in the first channel. The transition area between the two tracks is ignored, being supposed smaller than the aperture of both channels.

The electrodes are connected either directly one to another in pairs for standard MSC or , for the reversing MSC, in each group of three electrodes where the first one of the first track is connected to the second electrode from the second track, and vise-versa ,while the thirds electrodes are interconnected directly (Fig. 2).

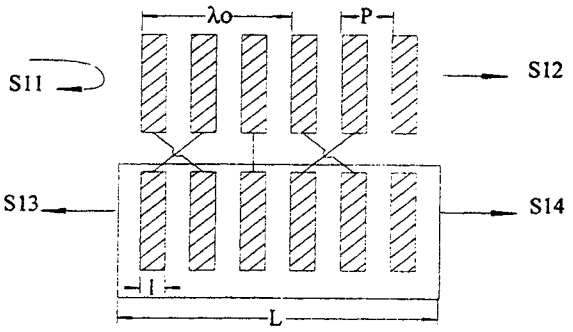


Fig 2

The charge distribution on single "n-th" electrode can be expressed in the following form

$$\sigma_n(x) = \frac{2}{\pi\, l} \frac{\sum_{k=0}^{K} A_{k,n} Ch_k\left(\frac{x-x_n}{l/2}\right)}{\sqrt{1-\left(\frac{x-x_n}{l/2}\right)^2}} \qquad (2)$$

where $Ch_k(y)$ are Chebyshev polynomials, x_n is the coordinate of the center of the electrode number *n*, and $A_{k,n}$ corresponding amplitude factors. The use of formula (2) gives three clear advantages:

1) the charge peculiarities at the edges of the electrodes are adequately described from the very beginning,

2) the Fourier components of charge can be obtained in an analytical closed form due to the formula:

$$\int_{-1}^{+1} \frac{Ch_k(y)}{\sqrt{1-y^2}} e^{iay}\, dy = (i)^k \pi\ J_k(|a|) \qquad (3)$$

for $a > 0$.

Here $J_k(z)$ are the Bessel functions of the first kind of order *k*. (For $a < 0$ a complex conjugated value should be taken.)

3) Using these formulas we can express a total charge on "n-th" electrode

$$Q_n = \int_{x_n - l/2}^{x_n + l/2} \sigma(x)\, dx = A_{0,n}. \qquad (4)$$

That is, the charge is equal to the amplitude factor of the zero order Chebyshev polynomial, which is very convenient for formulation of electroneutrality conditions.

2.2. ELECTRONEUTRALITY CONDITIONS

Now using formula (3) we can calculate spectral components of charge distribution , then using Green's function (a known function for in spectral q,ω-domain) we can find spectral components of potential and finally to determine potential at any point using the inverse Fourier transform. But before

realizing this procedure in the next section, we discuss electro-neutrality conditions. The problem is that for line charge the potential is divergent logarithmically at large distances. Actually this difficulty is artificial because in all real systems the charges are compensated and this divergence doesn't appear. To avoid this difficulty for every charge Q_n we place an opposite charge $-Q_n$ uniformly distributed over whole area *[-L/2, L/2]* (Fig. 2) occupied by the electrode system. At the same time the neutrality condition is imposed implicitly for electrodes in every acoustic channel:

$$\sum_{n=1}^{N} Q_n = 0. \tag{5}$$

Actually it means that the additional charge distributed over whole area *[-L/2, L/2]* is zero.

3. THE MAIN EQUATIONS

Using the procedures described above one can get, for potential created by the charges on "n-th" electrode, at an arbitrary point x the following formula:

$$\varphi_n(x) = I_{1,n} + I_{2,n} \tag{6}$$

$$I_{1,n} = \frac{Q_n}{\pi}\int_0^{\infty} G(q,\omega)[\cos(q(x-x_n))J_0(ql/2) - \frac{\sin(\frac{qL}{2})}{\frac{qL}{2}}\cos(qx)]dq \tag{6a}$$

$$I_{2,n} = \frac{1}{\pi}\sum_{k=1}^{k=\infty} A_{k,n}\int_0^{\infty} G(q,\omega)[\cos(q(x-x_n) - \frac{\pi}{2}k)J_k(ql/2)]dq \tag{6b}$$

The second term in the first integral is due to the opposite sign charge distributed over whole area and it makes this integral convergent at $q \to 0$. The term $\exp(iq_0x)$ must be added in the right hand side of equation (6) for the incident wave (if any). Equation (6) are basic for this method. They contain $N*K$ unknown amplitude factors, if K Chebyshev polynomials are used for charge distribution description on every one of N electrodes. The potential must be constant for all points belonging to the same electrode. We can write this condition for K equidistantly separated points on each of the electrodes. In this way we add N unknown potential values, but we get $N*K$ equations. (For two tracks of MSC all these numbers must be doubled.) The missing N ($2N$ for MSC) equations for potentials and charges depend on concrete connections of electrodes.

- For shorted electrodes reflector (one track) we have all potentials equal (*N-1* equations) and electroneutrality condition.
- For open electrodes reflector (one track) we have N neutrality conditions for every single electrode.
- For standard MSC we get *N-1* conditions that the difference of potentials is the same between connected electrodes in both tracks , that every pair of connected electrodes is neutral (N equations) and one more condition of the neutrality of a whole system of electrodes in one track.
- For the reversing MSC analogous conditions are formulated for corresponding electrodes.

Some comment should be made on the equal difference of potentials condition. Since our approach is essentially one-dimensional, the potentials in each track are supposed to be determined only by charges on the electrodes situated in that channel. We cannot calculate the influence on the potentials produced by the charges of another track electrodes. As far as only *differences* of the potentials are considered this approximation must work as well as formulas for connected capacitors in standard electric networks.

4. NUMERICAL PROCEDURES

We have used formulas (6) directly , not introducing any approximation for Green's function (effective permittivity), which was calculated in wide range at 10000 points with higher resolution near zero and quazipole regions (weak Rayleigh wave pole was ignored , Fig. 1). Then the linear system of 2*N*K+2*N+1 equations (for MSC) was solved numerically and all charges (amplitude factors) could be determined. Formula (6) was then used again to find the potential of the scattered wave far away from the electrode grating. It was found that with the increase of the number K of Chebyshev polynomials used the results converge extremely rapidly. The results for K=1, when only one Chebyshev polynomial $Ch_0(y) = 1$ is used are not sufficiently accurate , but for 2 polynomials the results are practically the same as for 3 or 4 . More than 4 polynomials are not necessary to use. This technique [14] has a clear advantage over standard approaches where each electrode is subdivided into narrow strips with uniform charge distribution. In this case a great number 40 to 100 of these substrips is necessary for adequate description of the charge distribution, which makes the time of calculation unacceptably long for any structure of practical interest. In our case MATLAB software with no special measures taken to accelerate calculations needs about 20 minutes for one frequency point calculation for MSC with 39 electrodes and K=3 Chebyshev polynomials used.

5. RESULTS AND DISCUSSION

Figure 3 shows the reflection and transmission coefficient of a leaky wave for a shorted electrodes, N=39, grating on the surface of 64-LiNbO3.

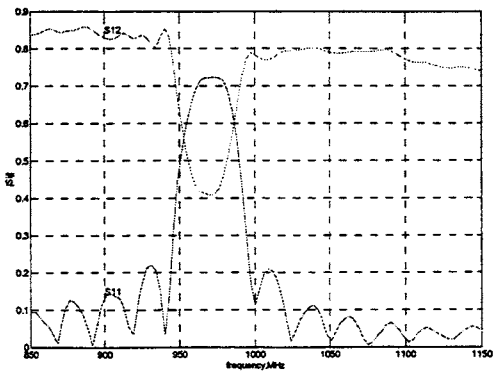


Fig. 3 Shorted electrodes reflector (N=39, K=3, m/p=0.5) characteristics

The observation points were 20λ in front or behind the grating. The amplitude of the incident wave in this and all other numerical examples was taken to be 1 in the center of the grating, which gives additional losses factor of 0.84. So, to get real magnitudes all values in plots of Fig. 3 to Fig. 7

must be divided by this number. Relative stopband width can be roughly estimated as $\frac{\Delta\omega}{\omega} \approx 2.1\%$, the reflectivity per wavelength being about 6.7% , and the velocity at the center of the stopband is 4520 m/sec which is consistent with what is predicted by perturbation theory [10].

Figure 4 illustrates the scattering matrix coefficients for a MSC with 2 electrodes per wavelength near its fundamental frequency. As for Fig. 2, S11 means the reflection backward in the same channel, S12 is the transmission thorough the grating, S13 - is the reflection in the neighboring channel (track-changing reflection) and S14 is cross-transmission. Qualitatively, the behavior of the curves is similar to measured many years ago for Rayleigh waves [6] and to predicted by COM model [7]. Again the additional propagation loss is present in all scattering parameters. The experimental data known to us for MSC characteristics at Bragg frequency are practically limited to mentioned paper and are hardly available for RMSC.

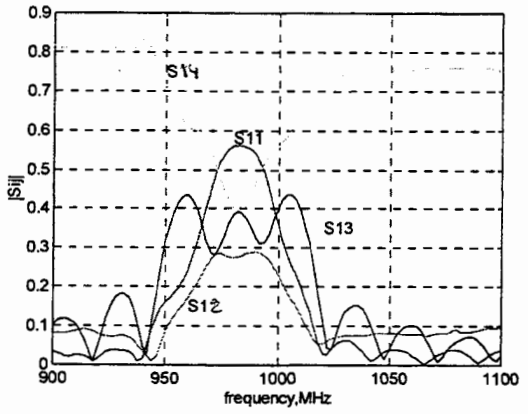


Fig. 4 Transmission and reflection coefficients of a MSC with N=39 , m/p=0.5, K=3 on 64-LiNbO3

Next Figure 5 shows a performance of a reversing multistrip coupler (RMSC) in a wide frequency range.

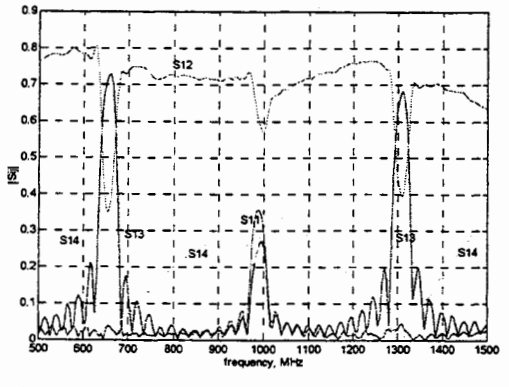


Fig. 5 RMSC performances, N=39, m/p=0.5, K=3

It is seen, that RMSC has a very useful characteristics at a fundamental frequency (see for details Fig. 6): low loss for track-changing reflection and a strong suppression of the close-in sidelobes. Moreover, there is practically no backward reflection in the same track and cross-transmission S14 is small in the passband but is increasing at other frequencies.

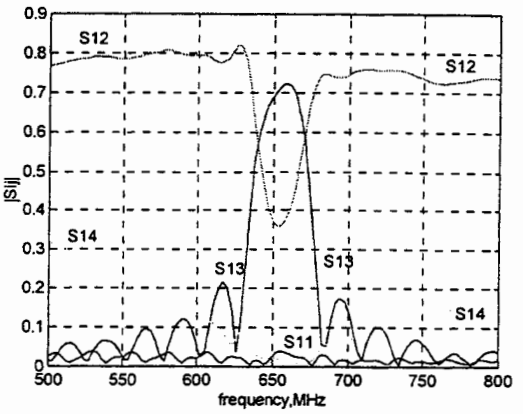


Fig. 6 RMSC performances (as on Fig. 5) at fundamental frequency range

Such performances are evidently useful for ring filter designs [1,3,4]. At low frequencies RMSC works to some extend as "normal" MSC transferring part of wave energy in parallel track in forward direction. It is also interesting to note that there is quite strong interaction on Bragg frequency when the period of the electrode structure is close to half of the incident wavelength. At a second overtone , Fig. 7, that is twice the fundamental frequency, the RMSC works almost as well as at the fundamental frequency, which is important for high frequency devices design.

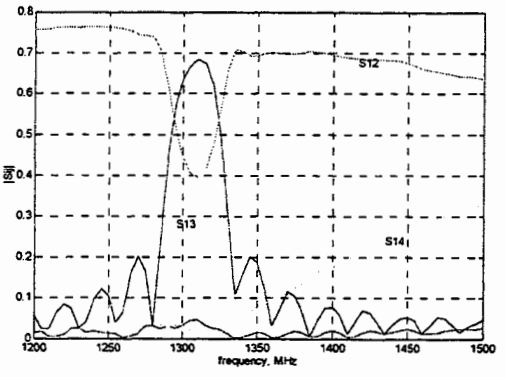


Fig. 7 RMSC (N=39, m/p=0.5, K=3) performance at second overtone.

5.1 RMSC EXPERIMENT

We have got some experimental data (Fig. 8) on RMSC on 128-LiNbO3 [15]. The structure included *25*3* electrodes (*3p=143.44 μm*) of *600 μm* aperture in each track. No comparison was made with the present theory , but the theory, presented by Danicki [9] gives comparable results (as far as the width of the passband and the levels of sidelobes are concerned). The minimal insertion losses were normalized to the same value. Note, that Green's function theory (Fig.5 and Fig. 6) gives the close-in left hand sidelobe also of higher level than that on the right hand side, as in the experiment.

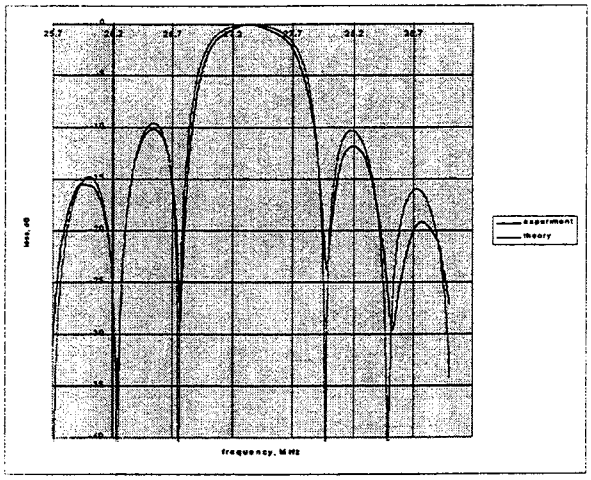


Fig. 8 RMSC on 128 LiNbO3

6. CONCLUSIONS

- Multistrip couplers had a limited use in low loss filter designs up to now probably because their performances were not studied in detail neither theoretically nor experimentally.
- The characteristics of the reversing MSC have some attractive features: no reflection in the same channel, no scattering into the bulk and hence no energy losses and fast roll-off of characteristics skirts.

REFERENCES

[1] W. Pollock, J. Schofield, R. F. Milsom, R. J. Murray and I. Flinn, "Low loss SAW filter using single-phase IDTs and no external tuning", 1983 IEEE Ultrasonics Symposium Proc., pp. 87-92.

[2] F. G. Marshall, E.G. S. Paige, "Novel acoustic-surface-wave directional coupler with diverse applications", Electronics Lett., Vol. 7, pp. 460-462, 1971.

[3] S. Doberstein, V. Malyukov, K. Nikolaenko, V. Razgonyaev, "Using of self-resonance approach for high selectivity low loss SAW filters", 1992 IEEE Ultrasonics Symposium Proc., pp. 151-154.

[4] S. A. Doberstein and V. A. Malyukov, "SAW ring filters with insertion loss of 1 dB", Proc. of 48th Frequency Control Symposium, pp. 330 - 336.

[5] C. Maerfeld and P. Tournois, " Perturbation theory for the surface-wave multistrip coupler", Electronics Lett., Vol. 5, pp. 115-116, 1973.

[6] J. Bording, K. A. Ingebrigtsen, "Scattering parameters of surface-wave multistrip directional couplers: a field approach", Electronics Letters, Vol. 9, No. 3, pp. 63-64, 1973

[7] B. P. Abbott, and D. C. Malocha, "Closed form solutions for multistrip coupler operation including the effects of electrode resistivity", 1990 IEEE Ultrasonics Symposium Proc., pp. 25-30.

[8] M. Feldmann and J. Henaff, "Design of multistrip arrays", 1977 IEEE Ultrasonics Symposium Proc., pp. 686-690

[9] E. Danicki, "Theory and applications of RMSC", 1992 IEEE Ultrasonics Symposium Proc., pp. 195-198.

[10] D. P. Morgan, *Surface-wave devices for signal processing*, New York,: Elsevier, 1985.

[11] B. P. Abbott, ETH, Zurich, unpublished

[12] G. Kovacs, M. Anhorn, M. E. Engan, G. Visintini and C. C. W. Ruppel, "Improved material constants for LiNbO3 and LiTaO3", 1990 IEEE Ultrasonics Symposium Proc., pp. 435-438.

[13] L. Boyer, J. Desbois, Y. Zang, and J. M. Hode, "Theoretical determination of the pseudo surface acoustic wave characteristic parameters", 1991 IEEE Ultrasonics Symposium Proc., pp. 353-358.

[14] V. P. Plessky, T. Thorvaldsson, "Rayleigh waves and leaky SAWs in periodic systems of electrodes: periodic Green functions analysis, Electronics Lett., Vol. 28, pp. 1317-1319, 1992.

[15] S. N. Krasnikova, O. S. Losev, "Design of SAW filters based on reflecting MSCs", Private communication, 1994

9 TH EUROPEAN FREQUENCY AND TIME FORUM - BESANÇON - 8 - 9 - 10 MARCH 1995

Session SAW II Number 15

COMPARATIVE PERFORMANCES OF 460 MHz LOW-LOSS SAW RING AND DUAL-TRACK IMAGE IMPEDANCE CONNECTED FILTERS

S.A.Dobersntein and V.A.Malyukhov

Omsk Research Electronic Equipment Institute, Masiennikov str., 231, Omsk, 644063, Russia. Fax: (381-2) 57 90 38

At present several types of low-loss SAW filters were suggested for RF stages of the modern mobile communication systems.[1] In this paper we compare two constructions of the low-loss SAW filters which are developed for the center frequency of 460 MHz on 128°YX $LiNbO_3$. The first construction is a dual-track three-transducer filter with the reflectors at the ends i.e. so-called dual-track image impedance connected (IIC) filter.[1] The second construction is a ring filter using reflective multistrip couplers (RMSCs) with 3 electrodes per λ (λ is the SAW wavelength at the center frequency).[2] Both filter constructions are developed for self-matching condition when filter specified real input/output impedances are achieved by interdigital transducer (IDT) static capacitance compensation with radiation susceptance. The optimization of the developed SAW filter topology is provided for achieving low insertion losses. For the ring filter we used the results of the previouse work[3] and for the dual-track IIC filter we determined the optimal number of reflector electrodes to provide SAW minimum transmission losses between input and output, the distance between input and output IDT to provide the specified filter bandwidth, obtained filter specified real input/output impedance and also real impedance at the connection point of two acoustic tracks by self-matching effect, chose optimal Al-electrodes thickness to reduce the conduction losses. The optimization of filter topologies was realized by successive approximation using computer programs on the basis of the equivalent circuit model. At 50 Ohm the dual-track IIC filter had insertion loss of 2.5 dB, amplitude ripple of 0.1 dB, 1 dB bandwidth of 2.1 MHz and stopband attenuation about 60 dB at +/- 60 MHz offset from the center frequency. Chip size is 2x1.5x0.7 mm. At 50 Ohm the ring filter has shown insertion losses of 1.2-1.5 dB, amplitude ripple of 0.1 dB, 1 dB bandwidth of 6 MHz, stopband attenuation over 50 dB at +/- 60 MHz offset from the center frequency. Chip size is 2x1.5x0.7 mm. The ring filter with phase weighted IDTs has shown insertion losses about 2 dB, stopband attenuation over 50 dB at +/- 20 MHz offset from the center frequency. The phase weighted dual-track IIC filter provided insertion loss of 3.5 dB. The cascade of two unweighted ring filter has shown stopband attenuation of 70-80 dB and insertion losses below 3 dB. When the cascaded dual-track IIC filters are used such stopband attenuation is provided but with greater insertion losses of 5-6 dB. We have presented the investigation results of the influence of fabrication tolerance (metalization thickness h and finger width d) on the SAW filter insertion loss and bandwidth. We mounted our SAW filters in 7.5x8.5x2.9 mm planar packages compatible with SMT and used them in 440-470 MHz mobile communication equipment.

1.G.Fischerauer,I.Schropp. Proc. IEEE Ultrason. Symp.,1993.
2.W.Pollock et al. Proc. IEEE Ultrason. Symp.,1983,p.87-92.
3.S.Dobersntein,V.Malyuknov. Proc. IEEE Freq. Control Symp.,1994.

SESSION 16

LASER FREQUENCY STABILIZATION

Chairman : P. THOMANN

COMPARISON OF CALCULATED PHASE-SENSITIVE RECOVERABLE SIGNALS USING A FABRY-PEROT CAVITY IN REFLECTION AND IN TRANSMISSION

Elio Bava, Fabio Massari

Politecnico di Milano, Dipartimento di Elettronica e Informazione, P.zza L. da Vinci 32, Milano, Italy

Abstract

This work is concerned with the analysis of the signals which can be recovered in a phase-sensitive detector after either transmission through or reflection from a Fabry-Perot interferometer of a phase-modulated light. The frequency responses are obtained and compared, the effects of residual amplitude modulation which result in offsets and lineshape distortions are analyzed, the conditions to reduce the in-loop frequency noise are examined taking into account the electronic noise, when the source is locked to the cavity by means of a high gain control system.

1. Introduction

Fabry-Perot cavities are widely used as frequency discriminators for laser frequency fluctuations measurements, line narrowing and short-term stabilization. To this purpose optical heterodyning techniques with phase-sensitive detection are used with interferometers operating in the transmission or in the reflection mode. A few papers analyse the static frequency response of these devices [1,2], the transient response to a frequency step [3], the use of a reflection cavity as a discriminator of laser quantum noise [4].
This work is mainly concerned with the comparison of the two operation modes of a Fabry-Perot, with phase-sensitive detection techniques. The difference between the calculated beat signals recoverable in the two topologies (reflection and transmission mode) are investigated varying the modulation parameters, that is the modulation index and the modulation frequency to cavity halflinewidth ratio x_c, in the range of experimentally acceptable values. As far as possible the discussion does not consider the two operation modes separately, unless to point out specific different behaviors, even if the sensitivities may be rather different. In general numerical computation is needed to describe the two operations, however in a few cases, mainly when $x_c \rangle 1$, the analytical development leads to meaningful expressions.
Dealing with the residual amplitude modulation of the light frequency modulator, attention is paid to evaluate offsets and to put in evidence regions where lineshape distortions are most evident. Moreover the conditions are examined to reduce the noise level at the detector input when a high gain electronics is used in the servo loop. According to the different amounts of the shot noise and of the intrinsic noise of the electronics transferred to the detector input, the optimum modulation index is evaluated.

2. Lineshape and sensitivity of the synchronously detected signal

Phase-sensitive detections of a phase-modulated light transmitted through or reflected by a Fabry-Perot cavity may be expressed in the same formal way. Detailed derivations of the detected signals are found in [1] and [2] respectively. Indeed, if ω_l and ω_m are the laser and the modulation angular frequencies and m the phase modulation index, the field with amplitude $\mathcal{A}$ incident into the cavity can be written

$$E_i(t) = \mathrm{Re}\left\{ \mathcal{A}\left[1 + a(t)\right] \sum_{n=-\infty}^{\infty} J_n(m) \exp\left[j\left((\omega_l + n\omega_m)t + \phi(t) \right) \right] \right\} \quad (1)$$

where $a(t)$ and $\phi(t)$ are deterministic or random processes accounting for amplitude and phase modulation or fluctuations.

The reflected or transmitted fields produce an output photocurrent I_p at the frequency $p\omega_m$ $(p \geq 1)$

$$I_p(t) = \frac{\eta e}{h\nu_l} c\varepsilon_0 \mathcal{A}^2 S Y_p(t) \quad (2)$$

where S is the detector area, $\nu_l = \omega_l / 2\pi$, η is the detector quantum efficiency, e, c, ε_0 and h are the electron charge, the velocity of light, the vacuum permittivity and Planck's constant respectively and $Y_p(t)$ is a normalised current which includes also the cavity response. Its expression is

$$Y_p(t) = \mathrm{Re}\left\{ \sum_{n=0}^{\infty} J_n(m) J_{n+p}(m) \left[F_n^* F_{n+p} + (-)^p F_{-n} F_{-(n+p)}^* \right] \exp\left(jp\omega_m t \right) \right\} +$$

$$+ \mathrm{Re}\left\{ \sum_{n=1}^{p-1}{}' (-)^{n+p} J_n(m) J_{p-n}(m) F_n F_{-(p-n)}^* \exp\left(jp\omega_m t \right) \right\} \quad (3)$$

where $\sum'$ holds for $p \geq 2$, with

$$F_k = \frac{1}{2\pi} \int_{-\infty}^{\infty} G(\omega) H\left[j(\omega + \omega_l + k\omega_m) \right] \exp(j\omega t) d\omega \quad (4)$$

and

$$G(\omega) = \int_{-\infty}^{\infty} \left[1 + a(t) \right] e^{j\varphi(t)} e^{-j\omega t} dt \quad (5)$$

and $H(j\omega)$ is the cavity response. When fluctuations are neglected F_k reduces either to the transmission or to the reflection coefficient at frequency $\omega_l + k\omega_m$.
Suitable expressions for H in the two operation modes are

$$H(j\omega) \equiv T(j\omega) = -T \frac{\exp(-j\omega\tau/2)}{1 - R\exp(-j\omega\tau)} \quad \text{(in transmission)} \quad (6a)$$

$$H(j\omega) \equiv \Gamma(j\omega) = \sqrt{R} \frac{1 - (1-A)\exp(-j\omega\tau)}{1 - R\exp(-j\omega\tau)} \quad \text{(in reflection)} \quad (6b)$$

where R, T and A, respectively the reflectivity, transmissivity and absorption of either mirror, are related by $R+T+A=1$ and τ is the round-trip time in the cavity.
The specific expressions of $I_p(t)$ in the two cases are readily obtained by developing the formulas given in [1] and [2] respectively.
Expression (3) shows that independently of the operation mode, for $\omega_m \rangle\rangle \Delta\omega_c$ ($\Delta\omega_c$ is the cavity FWHM) and $\omega_l - \omega_0 = \Delta\omega \langle\langle \omega_m$ (ω_0 is the cavity resonance), only terms with n=0 have an efficient discrimination power on $\Delta\omega$, whereas the remaining terms nearly cancel each other, leading to an asymptotic expression [1,2]

$$Y_p = 2J_0(m) J_p(m) H''(j\Delta\omega) \left[H''(jp\omega_m) \cos p\omega_m t + H'(jp\omega_m) \sin p\omega_m t \right] \quad (7)$$

where $H = H' + jH''$.
According to the previous expression, the lineshape is described by $H''(j\Delta\omega)$, and exhibits a maximum at $\Delta\omega \approx \Delta\omega_c/2$, whereas the signal amplitude depends on m through the product of $J_0(m)$ and $J_p(m)$, which peaks at m=1.08, 3.95 and 6.84, for p=1, 3, 5 respectively.
A few relationships are of interest to compare the reflection to the transmission mode of operation in this nonstatic regime. From (6a), (6b), we obtain $(AR \approx A)$:

$$\frac{\Gamma''(jp\omega_m)}{\Gamma'(jp\omega_m)} \approx \frac{(1+R-A)}{(1-R-A)} \tan\left(\frac{\tau p\omega_m}{2} \right) \quad (8a)$$

$$\frac{T''(jp\omega_m)}{T'(jp\omega_m)} = -\frac{(1+R)}{(1-R)} \tan\left(\frac{\tau p\omega_m}{2} \right) \quad (8b)$$

$$\left|\frac{\Gamma'(jp\omega_m)}{T''(jp\omega_m)}\right| \approx \frac{2\sqrt{R}}{(1+R)}\frac{(1+R-A)}{(1-R-A)}\,sin\left(\frac{\tau\, p\omega_m}{2}\right). \qquad (8c)$$

The sensitivity D_A of the system is obtained by performing $\left.\frac{dI_p}{d\Delta\omega}\right|_{\Delta\omega=0}$, the derivative of (2) with respect to $\Delta\omega$ around the line centre $\Delta\omega = 0$. General expressions have been reported in [1,2]

The dependence of the modulation index m_{opt}, which maximises D_A, on the cavity parameter $x_c = 2\,\omega_m/\Delta\omega_c$ in the reflection mode, shown in figure 1, should therefore be compared with figure 5 in [1]; numerical computations on lossless symmetric cavities yield the same dependence of m_{opt} on x_c in both cases.

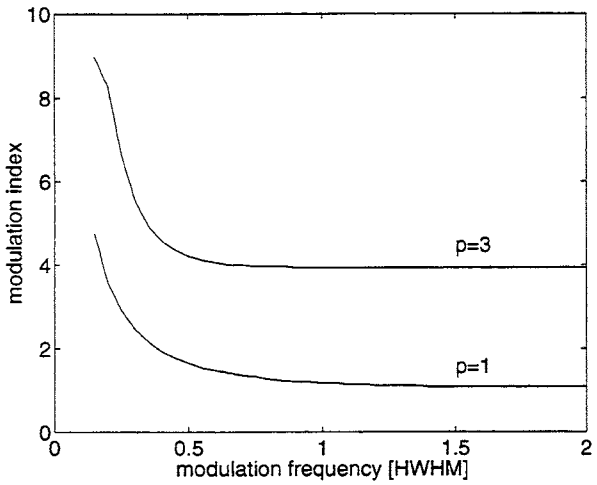


Figure 1: optimum modulation index as a function of the cavity parameter x_c

Figure 2 shows the maximum achievable sensitivity at p=1, assuming, as in [1], $\Delta\omega_c/(2*2\pi)$=35 kHz and Finesse=7000.

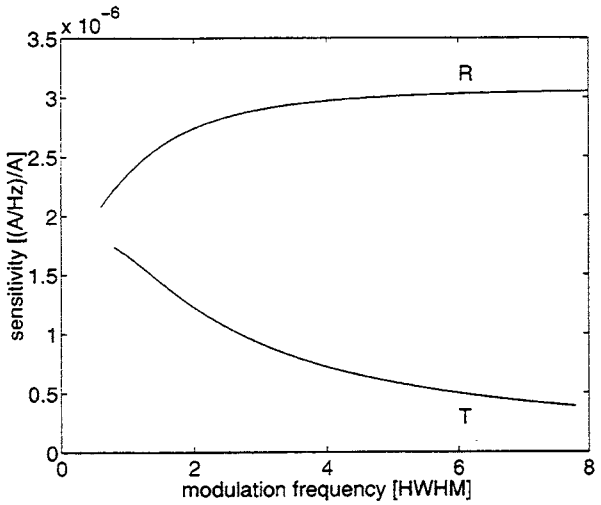


Figure 2: normalised sensitivity $D_A \Big/ \frac{\eta e}{h\nu_l} c\varepsilon_0 \mathcal{A}^2 S$ at p=1 in reflection (R) and transmission (T)

When $\omega_m \rangle\rangle \Delta\omega_c$, the sensitivity, from (7), is proportional to $\left.\frac{dH''}{d\Delta\omega}\right|_{\Delta\omega=0} \approx \tau\frac{(1-R-A)}{(1-R)^2}$ both in reflection and in transmission.

3. Offsets and distortions of residual amplitude modulation at ω_m

Let us suppose the laser frequency modulator to produce a residual amplitude modulation $a(t) = a_1 sin(\omega_m t + \phi_1)$, $\phi(t) = 0$ in (1). After substitution in (3), the normalised current becomes

$$Y_p = \mathrm{Re}\left\{\sum_{n=-\infty}^{\infty} H(j(\Delta\omega + n\omega_m))H^*(j(\Delta\omega + (n+p)\omega_m))L_n L^*_{n+p}\exp(-jp\omega_m t)\right\} \qquad (9)$$

where

$$L_n = J_n(m) + \frac{a_1}{2j}\left[J_{n-1}(m)\exp(j\varphi_1) - J_{n+1}(m)\exp(-j\varphi_1)\right] \qquad (10)$$

The effect of $a(t)$ is to introduce offsets and distortions in the detected current, which is otherwise an odd or even function with respect to $\Delta\omega = 0$ for p odd or even respectively. In general numerical computations are needed to evaluate the amount of these effects, however the analysis can offer means to recognise the most important conclusions.

For $\Delta\omega = 0$ a rather compact expression is found for the two offset components of $Y_{p,0}(t)$. By using (9) and (10) and taking into account that $H(j\Delta\omega) = H^*(-j\Delta\omega)$, it follows

$$Y_{p,s,0} = a_1 \sum_{n\geq -(p-1)/2}\left(sin\phi_1 A_{1n} I_n + \cos\phi_1 A_{2n} R_n\right) \qquad (11a)$$

$$Y_{p,c,0} = -a_1 \sum_{n\geq -(p-1)/2}\left(-sin\phi_1 A_{1n} R_n + \cos\phi_1 A_{2n} I_n\right) \qquad (11b)$$

where

$$R_n + jI_n = H(jn\omega_m)H^*\left[j(n+p)\omega_m\right] \qquad (12)$$

$$A_{1n} = J_{n+p}\left(J_{n+1} + J_{n-1}\right) + J_n\left(J_{n+p+1} + J_{n+p-1}\right) \qquad (13a)$$

$$A_{2n} = J_{n+p}\left(J_{n+1} - J_{n-1}\right) - J_n\left(J_{n+p+1} - J_{n+p-1}\right) \qquad (13b)$$

Note that in the case of a matched and lossless Fabry-Perot a relationship holds for the two expressions obtained from (12), for $H \equiv \Gamma$ or $H \equiv T$:

$$\Gamma(jn\omega_m)\Gamma^*\left(j(n+p)\omega_m\right) = T(jn\omega_m)T^*\left(j(n+p)\omega_m\right) \bullet$$
$$\bullet\frac{4R}{T^2}\,sin\frac{n\omega_m\tau}{2}\,sin\frac{(n+p)\omega_m\tau}{2} \qquad (14)$$

which shows that the two possible expressions of (12) are strictly related in amplitude and phase, except for $n = 0$.

In general when $\omega_m \rangle\rangle \Delta\omega_c$, in (12) $I_n \langle\langle R_n$ for $H(j\omega) \equiv \Gamma(j\omega)$, and the offsets are proportional to $\cos\phi_1$ (11a) and $sin\phi_1$ (11b) respectively, as was already pointed out in [2]. In particular strong contributions to the offsets arise from the term n=0, if the cavity is not well matched [5].

The same result does not apply in general for $H(j\omega) \equiv T(j\omega)$, because, when n=0, (12) is approximately jI_p and $R_p \approx 0$.

Let us evaluate the distortions where they are expected, that is for $\Delta\omega = \delta\omega$ and $\Delta\omega = \delta\omega \pm \omega_m$, $\delta\omega \langle\langle \omega_m$. It is useful to separate in Y_p terms which exhibit a dispersion with $\delta\omega$ from those nearly independent.

In the former case, after some algebra, (9) can be written

$$Y_p = \mathrm{Re}\left[\exp(-jp\omega_m t)\right.$$
$$\left\{H(j\delta\omega)H^*\left[j(\delta\omega + p\omega_m)\right]L_0 L^*_{-p} + H\left[j(\delta\omega - p\omega_m)\right]H^*(j\delta\omega)L_{-p}L^*_0 + \right.$$
$$\left.\left.\sum_{n=-\infty, n\neq 0, \neq -p}^{\infty} H\left[j(\delta\omega + n\omega_m)\right]H^*\left[j(\delta\omega + (p+n)\omega_m)\right]L_n L^*_{n+p}\right\}\right] \qquad (15)$$

Under the same hypothesis $\omega_m \rangle\rangle \Delta\omega_c$, in the most important mode of operation in this region, that is $H \equiv \Gamma$, the following expression is found (p odd) for the distortions $\Delta Y_{p,s,\delta\omega}$ and $\Delta Y_{p,c,\delta\omega}$ of the normalised photocurrent given by expression (7)

$$\Delta Y_{p,s,\delta\omega} \approx a_1 \cos\phi_1 \Gamma'(jp\omega_m)\left\{\Gamma'(j\delta\omega)\left[2J_1 J_p - J_0\left(J_{p+1} - J_{p-1}\right)\right] + \right.$$
$$\left. + \Gamma'(jp\omega_m)\sum_{n=-(p-1)/2, n\neq 0}^{\infty} A_{2n}\right\} \qquad (16)$$

$$\Delta Y_{p,c,\delta\omega} \approx a_1 sin\phi_1 \Gamma'(jp\omega_m)\left\{\Gamma'(j\delta\omega)J_0\left(J_{p+1} + J_{p-1}\right) + \right.$$
$$\left. + \Gamma'(jp\omega_m)\sum_{n=-(p-1)/2, n\neq 0}^{\infty} A_{1n}\right\} \qquad (17)$$

Expressions (16) and (17) give the patterns of figures 3 and 4 respectively, when a unit residual amplitude modulation depth and a dephasing for maximum distortion are assumed.

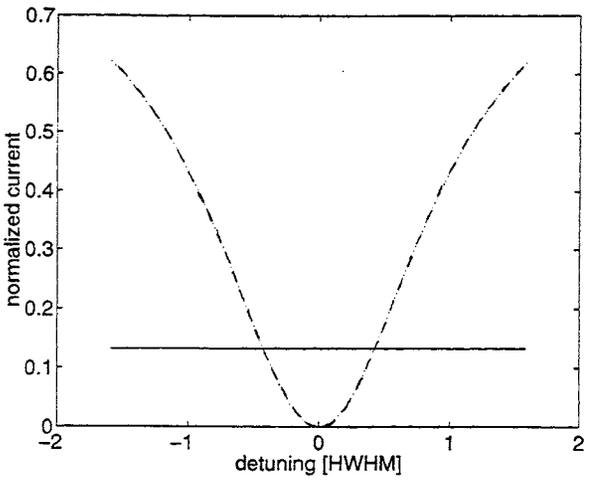


Figure 3: distortion of the lineshape (in-phase component $\Delta Y_{p,s,\delta\omega}$) around the cavity centre frequency. $\Gamma(0)=0$, $m=1.08$, $x_c = 40$, $a_1 = 1$, $\varphi_1 = 0$. The contribution proportional to $\Gamma'(j\delta\omega)$ (dash-dot line) has been separated from the constant offset (solid line).

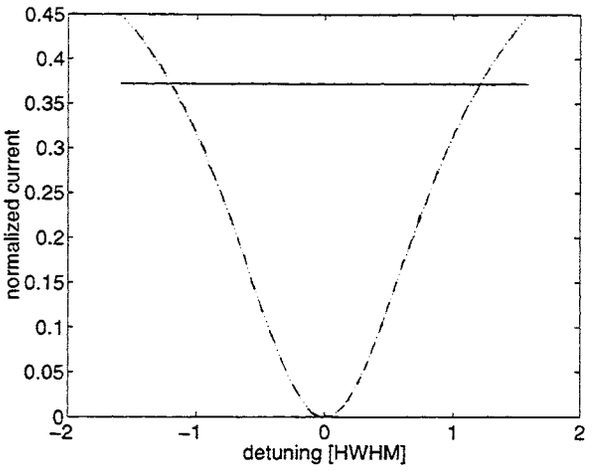


Figure 4: distortion of the lineshape (in-quadrature component $\Delta Y_{p,c,\delta\omega}$). $\Gamma(0)=0$, $m=1.08$, $x_c = 40$, $a_1 = 1$, $\varphi_1 = 90^0$.

Following the same procedure in the latter case, evidence of offsets and distortions is obtained after some algebra:

$$Y_{p,\delta\omega\pm\omega_m} = \mathrm{Re}\exp(jp\omega_m t)\left\{\sum_n H\left[j(\delta\omega+n\omega_m)\right]H^*\left[j(\delta\omega+(n-p)\omega_m)\right]L_{n\mp1}L^*_{n-(p\pm1)}\right\} \quad (18)$$

For sake of simplicity here only the case $\delta\omega = 0$ is evaluated. The presence of a(t) makes $\neq 0$ the sum of the normalised currents at opposite $\Delta\omega$ values:

$$Y_{p,s,\omega_m} + Y_{p,s,-\omega_m} \approx -a_1\cos\phi_1\left\{\ \Gamma(0)\Gamma'(jp\omega_m)\left(A_{2,-1}+A_{2,p-1}\right)+ \right.$$
$$\left. +\left|\Gamma(jp\omega_m)\right|^2\sum_{n\neq0,\neq p}A_{2,n-1}\ \right\} \quad (19a)$$

$$Y_{p,c,\omega_m} + Y_{p,c,-\omega_m} \approx -a_1\sin\phi_1\left\{\ \Gamma(0)\Gamma'(jp\omega_m)\left(A_{1,-1}+A_{1,p-1}\right)+ \right.$$
$$\left. +\left|\Gamma(jp\omega_m)\right|^2\sum_{n\neq0,\neq p}A_{1,n-1}\ \right\} \quad (19b)$$

where $A_{1,n}$ and $A_{2,n}$ have been defined in (13a) and (13b).

4. Frequency response

The simultaneous presence of amplitude and phase noise in the light sidebands leads, in general, to offsets at the output of the phase-sensitive detector, unless the two processes are statistically independent. This was shown in a rather general way, although limited to the case $\omega_m \gg \Delta\omega_c$, in [6] for $H \equiv \Gamma$, but, following the same procedure it can be proved also for $H \equiv \mathrm{T}$. In the simple case of fluctuations at frequency q, an exact evaluation of the offset is obtained by substituting in (1) $\omega_i = \omega_0$, $a(t) = a\sin(qt+\alpha)$, and $\phi(t) = b\sin qt$. The relaxation oscillatons of a solid-state laser are adequately described by the previous single freqency amplitude noise, provided α is an appropriately chosen mean angle. The following continuous components result

$$\overline{Y_{ps}} = 2\sum_{k\geq1,n\geq-(p-1)/2} J_n(m)J_{n+p}(m)B_k\,\mathrm{Im}\left\{H^*\left[(n+p)\omega_m+kq\right]H\left[n\omega_m+kq\right]+\right.$$
$$\left. -H^*\left[(n+p)\omega_m-kq\right]H\left[n\omega_m-kq\right]\right\} \quad (20a)$$

$$\overline{Y_{pc}} = 2\sum_{k\geq1,n\geq-(p-1)/2} J_n(m)J_{n+p}(m)B_k\,\mathrm{Re}\left\{H^*\left[(n+p)\omega_m+kq\right]H\left[n\omega_m+kq\right]+\right.$$
$$\left. -H^*\left[(n+p)\omega_m-kq\right]H\left[n\omega_m-kq\right]\right\} \quad (20b)$$

with

$$B_k = a\sin\alpha\ J_k(b)\left[J_{k+1}(b)+J_{k-1}(b)\right] \quad (21)$$

The previous equations show that bias is present in both quadratures and is $\propto a\sin\alpha$ provided $b \neq 0$. In the limit $b \ll 1$, $\sum_{k\geq1}$ reduces just to the term $B_1 = \frac{ab}{2}\sin\alpha$.

To obtain the linearized-system frequency response, at first we evaluate the output at frequency q in the limit $b \to 0 (a = 0)$. Under the hypothesis $\omega_m \gg \Delta\omega_c$, it follows

$$Y_{ps}(t) \approx 2J_0(m)J_p(m)b\,\mathrm{Re}\left\{e^{-jp\omega_m t}jH^*(p\omega_m)\left[-H(0)\sin qt+\left|H(q)\right|\sin(qt+\theta)\right]\right\} \quad (22)$$

where θ is the phase of H. From (22), if $H \equiv \Gamma$, assuming $\Gamma(p\omega_m) \approx \Gamma'(p\omega_m)$ the following frequency response is obtained for the sine component

$$H_\Gamma(jq) \approx 2J_0(m)J_p(m)\Gamma'(jp\omega_m)\frac{\Gamma(jq)-\Gamma(0)}{jq} \quad (23)$$

After substitution of the reflection coefficient in (23), the dependence on q of H_Γ turns out, if $R \to 1$, $\propto\left(1+j\frac{2\tau R}{1-R}q\right)^{-1}$, which is close to $\left(1+j\frac{q}{\Delta\omega_c/2}\right)^{-1}$.

The frequency noise to amplitude conversion behaviour, which is approximated by the asymptotic expression (23), is shown in figure 5 for x_c=40

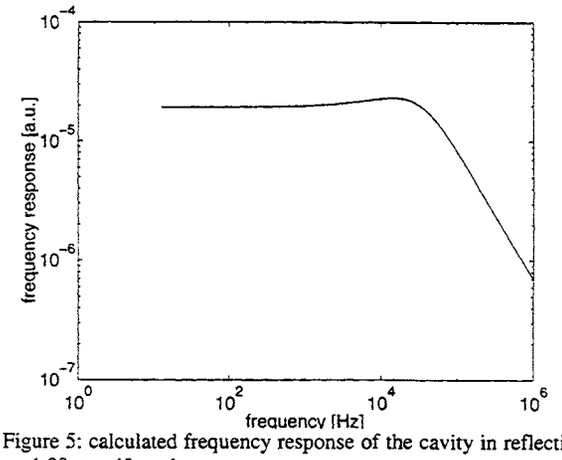


Figure 5: calculated frequency response of the cavity in reflection; m=1.08, x_c=40, p=1

Expression (22) agrees, for p=1, with the diagrams reported in [7], with expressions (6,7) in [4] and with the approximation presented in [3]. Moreover, the dependence of H_Γ on q is the same as in [10], where the discrimination responses were calculated for unmodulated light, provided the working point in eq. (10) of ref.[10] is $\Delta\omega = 2\pi\delta\nu_0 = 0$.

5. Signal to noise ratio and closed loop spectral density of frequency noise

Performances of servo-loops for frequency stabilisation and closed loop frequency noise are routinely evaluated by measuring the feedback error signal [8]. When the laser is tightly locked to the reference cavity, source fluctuations are strongly reduced, and the achievable reduction of the spectral density of frequency noise is limited by shot noise at the photodetector and by the electronic noise transferred to the detector input. Indeed, if $D_A\,[A/Hz]$ is the sensitivity of the frequency discriminator and $S_A\,[A^2/Hz]$ is the spectral density of the equivalent current noise at the detector input, the low frequency part of the closed loop spectral density of frequency noise is given by

$$S_{\Delta\nu,cl} = S_A / |D_A^2| \tag{24}$$

where

$$S_A = 2e\frac{e\eta}{h\nu_l} P_i \left\{ (1-Q)\left[J_0^2(m) H^2(0) + 2\sum_{k=1}^{\infty} J_k^2(m) \left| H(jk\omega_m) \right|^2 \right] + Q' \right\} + N \tag{25}$$

$N\,[A^2/Hz]$ is the equivalent input noise current density, P_iQ is the input power coupled to spatial modes of the cavity other than the fundamental one, which we assume to impinge on the photodetector after off-line reflection ($Q' = R\frac{1+(1-A)^2}{1+R^2}Q$ in reflection, whereas Q'=0 in transmission).

Thus, for a detector with given noise spectral density and maximum input power, an optimum modulation index m_{opt} is to be chosen, depending on H. This problem, as far as the reflection topology with $\Delta\omega_c \langle\langle \omega_m$ is concerned, has been addressed in [2], whereas the discussion in [9] was restricted to the shot noise limited frequency noise, disregarding electronic or background noise and mode-matching problems.

In general, m_{opt} ranges between a limiting value set by the electronic noise, which regardless of $H(0)$ maximises the sensitivity D_A according to figure 1), and the optimum modulation which exists when the shot noise dominates over the electronics noise; the latter may strongly depend on $\Gamma(0)$ in reflection, and differs from the former in particular when p=1 and high modulation to cavity halfwidth ratios are considered. At higher harmonics, as a matter of fact, the shot noise is rather insensitive to the carrier contribution ($\propto J_0^2 H^2(0)$), and the signal to noise ratio is only slightly impaired if m is set to maximise D_A.

We define a suitable adimensional noise parameter M as

$$M = \frac{\frac{Nh\nu_l}{2e^2\eta P_i} + Q'}{1-Q} \tag{26}$$

When $Q = 0$ (the incident power is coupled to a single spatial mode of the cavity), M is the ratio of the electronic to the shot noise spectral density.

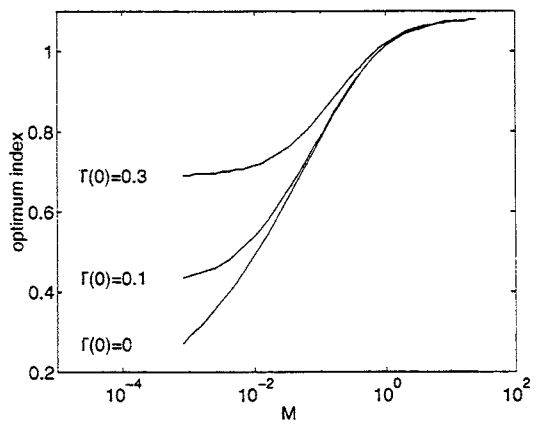


Figure 6: optimum modulation index dependence on the noise parameter M in reflection mode, varying the in resonance reflection coefficient $\Gamma(0)$; x_c=20, HWHM=35kHz.

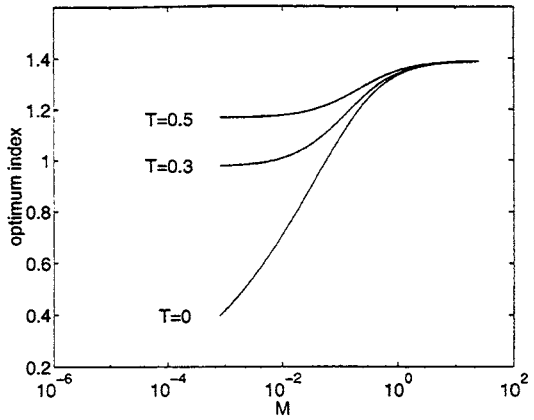


Figure 7: optimum modulation index dependence on the noise parameter M in reflection mode; x_c=0.67, HWHM=35kHz.

Figures 6, 7 and 8 show the dependence of m_{opt} on M in reflection (6,7) and transmission (8), with p=1 and x_c=20 (6) and x_c=0.67 (7,8) respectively. In the latter case, the shift of the curve in figure 8 appears as a straightforward consequence of assuming a symmetric cavity, according to (6b).

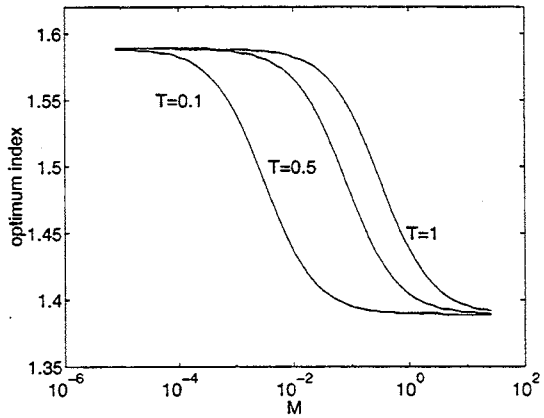


Figure 8: optimum modulation index dependence on the noise parameter M in transmission mode, varying the in resonance transmission coefficient $T(0)$; x_c=0.67, HWHM=35kHz.

6. References

[1] D. Hils and J. L. Hall, Rev. Sci. Instrum. **58**, 1406 (1987)

[2] E. Bava, M. Bonaiti and F. Massari in *Proceedings of the VIIIth European Frequency and Time Forum*, (München, D, 1994) 1039.

[3] M. Houssin, M. Jardino and M. Desaintfuscien, Rev. Sci. Instrum. **61**, 3348 (1987).

[4] Y. Cheng, P. L. Mussche and A. E. Siegman, IEEE J. Quantum Electron. **QE-30**, 1498 (1994).

[5] E. Bava and F. Massari, *Proceedings of the VIIth LEOS annnual meeting*, (Boston, USA, 1994) 157.

[6] E. Bava and F. Massari, Politecnico di Milano, Technical Report 94-106, submitted to Rev. Sci. Instrum.

[7] M. Zhu and J. L. Hall, J. Opt. Soc. Am. B **10** , 802 (1993).

[8] T. Day, E. K. Gustafson and R. L. Byer, IEEE J. Quantum Electron. **QE-28**, 1106 (1992).

[9] K. Nakagawa, A. S. Shelkovnikov, T. Katsuda and M. Ohtsu, Opt. Commun. **109**, 446 (1994).

[10] P. Tremblay and R. Ouellet, IEEE Trans. **IM-40**, 204 (1991).

Acknowledgments

This work has been partially supported by the Italian Space Agency (ASI) and by the Italian National Council for Research (CNR).

Ultra-stable cavity-stabilized laser diode at 729 nm

B. Fermigier, M. Desaintfuscien, J. Dupont, M. Houssin and F. Plumelle

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

Abstract

The experimental setup to cool a 750 nm commercial laser diode to obtain 729 nm is described. This wavelength is the clock transition of an optical wavelength/frequency standard using a confined Ca^+ ion. Preliminary laser linewidth in a self-aligned external cavity is also described and characterized. A Pound-Drever servo configuration with an ultra-stable cavity is in progress. Minimum attainable linewidth due to fundamental noises is estimated.

I. Introduction

Laser cooled ions confined in electromagnetic traps form the basis of an optical frequency standard with very high stability and accuracy, superior to that of the Cs beam standard.Ca^+ is a particularly attractive candidate because all the useful transitions involved in the experiment can be reached with laser diodes [1] [2] [3].

The clock transition, at 729 nm, will be probed by a cooled laser diode. This laser, whose linewidth must be very narrow to take advantage of the line Q, will be stabilized on an ULE cavity using the optical Pound-Drever method [4] [5].

We present a diode cooling system in order to obtain a lasing wavelength of 729 nm with commercial laser diodes near 750 nm. This diode is placed in an external cavity to obtain a preliminary linewidth narrowing. The cooled external cavity laser is characterized by showing its frequency noise spectrum.

Then the minimum attainable linewidth for the servo-controlled laser is calculated.

II. Realization of a cooled external-cavity laser diode

Unfortunately there is no 729 nm commercial laser diode but the 729 nm wavelength can be obtained by cooling a commercial laser diode (SHARP LT030/031 near 750 nm). We measured that the emitted wavelength of this diode is nearly a linear function of the temperature with a 0.2 nm/K slope (Fig. 1). As a consequence, our working temperature is around -80°C. The laser diode must be put under vacuum to avoid condensation and reduce the heat losses. Moreover, since the diode may be damaged by repeated temperature variations, it will be permanently kept at this low temperature.

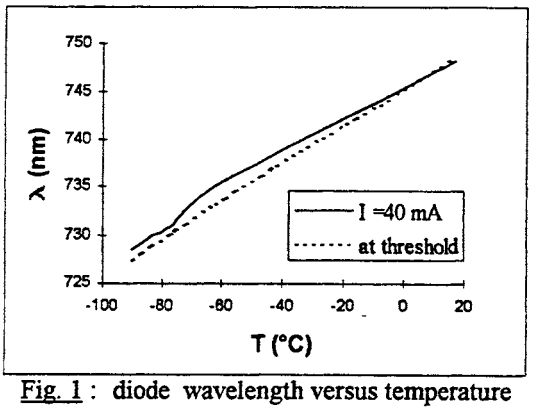


Fig. 1 : diode wavelength versus temperature

Firstly we tried to use Peltier thermoelectric coolers that are easy to control, well designed for compact setup and can work round the clock. However

the achievement of so low temperatures requires multi-stage thermoelectric coolers and we faced heat evacuation problems.

Then we turned to a liquid nitrogen cooling system with a small Peltier element for the temperature regulation (Fig. 2). A Dewar tank full of nitrogen is used ; a copper strap provides heat exchange between the diode and this liquid nitrogen tank. We have chosen a supple strap that permits to absorb nitrogen vibrations. The length and the section are chosen so that the diode temperature approaches -80°C, Peltier being off. So the Peltier works at low current to obtain the required temperature.

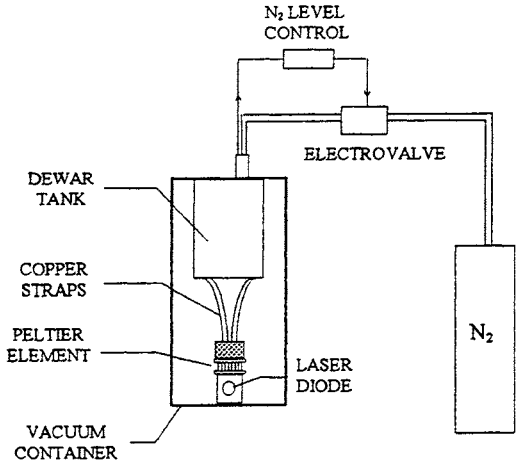


Fig.2 : nitrogen cooling system

Moreover the laser diode is put in a self-aligned external cavity to obtain a preliminary linewidth reduction and an enhanced wavelength tuning due to the grating feedback. This setup includes the classical external cavity components and a cylindrical lens as shown in fig. 3 [6]: the output facet of the diode is antireflection coated; the laser beam is collimated by the first lens and the grating is in Littrow configuration.The cylindrical lens and the grating separated by a focal length distance form a system that gives a reflected beam parallel to the incident one in the vertical plane while keeping the grating wavelength selectivity in the horizontal plane. This cavity is less sensitive to mechanical perturbations and misalignment. It permits the suppression of positioning components to align the cavity. The cylindrical lens is parallel to the grating to obtain a collimated output beam.

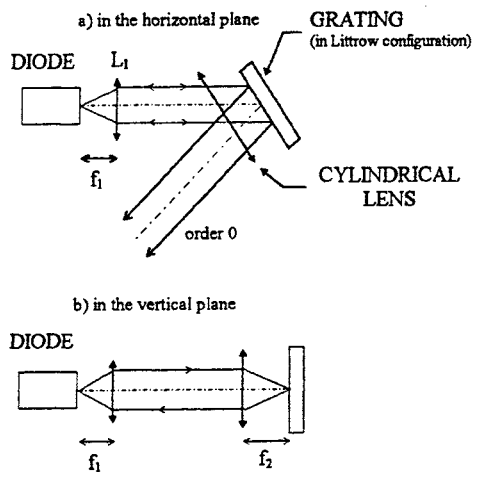


Fig. 3 : self-aligned external cavity.

We have measured the wavelength tunability of our cavity according to the diode temperature (Fig. 4). This curve shows that the tunability range decreases with the temperature. We observe easier competition between the cavity and the solitary diode modes. We think that this phenomenon is due to the coating of the diode which is not so effective at 729 nm because it has been realized (by maximizing the LD threshold) at ambient temperature and thus at 750 nm.

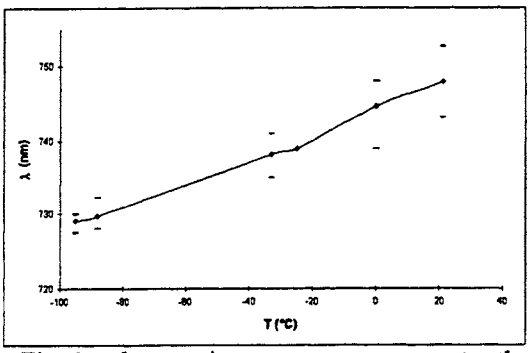


Fig. 4 : the continuous curve represents the external cavity wavelength at the threshold I_{th} and the plots show the wavelength tunability at 1.1 I_{th} .

Concerning the preliminary linewidth reduction, the curve (fig. 5) shows the frequency noise spectrum up to 40 kHz. It is achieved by a BF servo-control on a Fabry-Perot fringe side. We can note an important noise level at low frequencies because of mechanical and acoustic disturbances. Above several kHz the noise is white. Such a white noise level (around 200 Hz/$\sqrt{\text{Hz}}$) would correspond to a linewidth of 120 kHz.

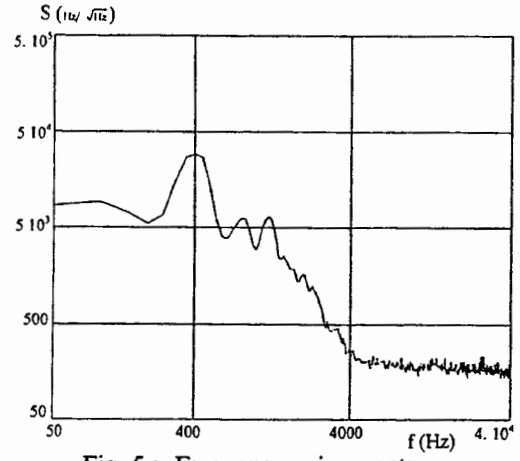


Fig. 5 : Frequency noise spectrum

II. Servo-controlled laser linewidth

The method we have chosen is the Pound-Drever technique (Fig. 6).

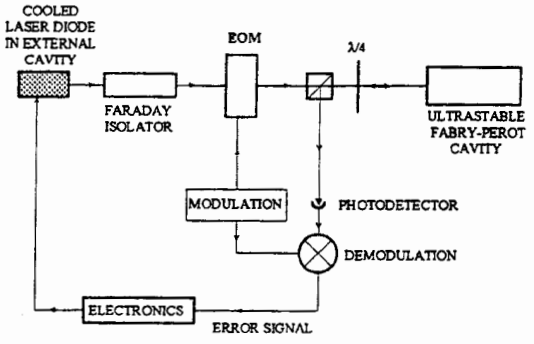


Fig. 6 : Pound-Drever servo-loop technique

The laser diode optical field E_0 is modulated at the frequency ω_m with an electro-optic modulator. The signal E_i is injected in a Fabry-Perot cavity which gives a reflected beam E_r. The incident field can be written :

$$E_i = E_0 J_0 e^{i\omega t} + E_0 J_1 e^{i(\omega+\omega_m)t} + E_0 J_{-1} e^{i(\omega-\omega_m)t}$$

The reflected intensity is then :

$$|E_r|^2 \approx P_{DC} + P_{\omega_m}$$

with

$$P_{DC} = 2E_0^2 J_1^2 r^2 + \frac{E_0^2 J_0^2 r^2 \left(\frac{\nu}{\Delta\nu_c/2}\right)^2}{1+\left(\frac{\nu}{\Delta\nu_c/2}\right)^2} \quad (1)$$

and

$$P_{\omega_m} = 2rE_0^2 J_0 J_1 \frac{\left(\frac{\nu}{\Delta\nu_c/2}\right)}{1+\left(\frac{\nu}{\Delta\nu_c/2}\right)^2} . 2\sin\omega_m t \quad (2)$$

(ν : laser detuning from the cavity resonance frequency ;
F : FP finesse ; L : FP length ;
$\Delta\nu_c = c/2LF$: FP linewidth)

Then the photocurrent is :

$$I_{tot} = \frac{\rho q}{h\nu} |E_r|^2$$

and the useful signal after demodulation is:

$$I_{useful} = \frac{\rho q}{h\nu} 2rE_o^2 J_0 J_1 \frac{\left(\frac{\nu}{\Delta\nu_c/2}\right)}{1+\left(\frac{\nu}{\Delta\nu_c/2}\right)^2} \quad (3)$$

$$\approx \gamma . \nu$$

where the error signal slope γ is :

$$\gamma = \frac{\rho q}{h\nu} \frac{4rE_o^2 J_0 J_1}{\Delta\nu_c} \quad (4)$$

γ is proportional to the length L and the finesse F of the Fabry-Perot. With a 10 cm long FP, a finesse of 15.000 and a laser power of 1 mW, we calculate:

$$\gamma \approx 10^{-9} \text{ A / Hz}$$

Unfortunately in the servo loop this signal is degraded by different noise sources. If the loop gain is sufficiently high, the FM noise level can be ultimately reduced to a value limited by these noises. We consider different noise sources: shot noise of the photodiode, laser intensity noise and electronic noise.

Shot noise

The frequency noise spectral density of the shot noise of the fast photodiode receiving the reflected intensity is :

$$S_{shotnoise} = \sqrt{2qI_{tot}} \approx \sqrt{2q}\sqrt{\frac{\rho q}{h\nu}}\,2E_oJ_1r$$
$$\approx 10^{-12}\ A/\sqrt{Hz}$$

Laser intensity noise

It appears at two levels in the photocurrent expression:

- in the useful part (2) and then (3), the error signal slope γ is only slightly modified.
- in the continuous part (1), the variations of E_0^2 lead to a frequency noise density of:

$$S_{I.N.} = I_{DC}.\sqrt{RIN}$$
$$\approx 10^{-11}\ A/\sqrt{Hz}$$

where the typical relative intensity noise of the laser diode (RIN) is around 10^{-12} /Hz.

Electronic noise

The NEP of the photodiode and the first amplification stage is around $10^{-12}\ A/\sqrt{Hz}$.

From these results, we note that the limitation due to the laser intensity noise is preponderant. This photocurrent noise spectral density corresponds to a laser frequency noise density of :

$$S_{laser} = S_{I.N.} / \gamma \approx 0.01\ Hz\ /\sqrt{Hz}$$

which gives a 0.3mHz linewidth if we suppose a lorentzian profile.

IV. Conclusion

For the Calcium experiment, we need a 1 Hz linewidth. Our estimation shows that these three noises will not be our limitation. We still have to realize a very low noise electronic loop with high BF gain and 1 MHz bandwidth [7]. Moreover strong optical feedback reducing is necessary.

References

1. F. Plumelle, M. Desaintfuscien and M. Houssin, IEEE Trans. Instr. Meas. **42**, n°2 (1993).
2. S. Urabe, K. Hayasaka, M. Watanabe, H. Imajo, R. Ohmukai and R. Hayashi, Appl. Phys. **B57**, pp 367-371 (1993).
3. M. Knoop, M. Vedel and F. Vedel, J. Phys.,II France **4**, pp 1639-1650 (1994).
4. D. Hils and J. L. Hall, Frequency Standards and Metrology, Proc. of the 4th Symposium, A. de Marchi, Ed. Belin: Springer-Verlag (1989).
5. K. Nakagawa, M. Kourogi and M. Ohtsu, Opt. Lett. **17**, n°13, pp 934-936, (1992).
6. N. Dimarcq, M. Houssin, E. Aucouturier and M. de Labachelerie, OSA Annual Meeting/ILS-X, Dallas (Oct. 1994).
7. M. Ohtsu and S. Kotajima, IEEE J. Quant. Elec. **QE-21**, pp 1905-1912 (1985).

An Ultrastable High Power Nd:YAG Laser for the Detection of Gravitational Waves

F.Bondu, A.Brillet, P.Fritschel*, F.Perrone, F.Cleva, L.Latrach and C.N.Man

CNRS, VIRGO-LAL, bat 208, Campus, 91 405 Orsay Cedex, France
* Permanent Address: MIT , Cambridge, Massachusetts 02139, USA

Abstract:

The detection of gravitational waves is certainly one of the most exciting scientific and technical challenges for the coming years, which will see the construction of the French-Italian VIRGO antenna near Pisa, and of the two US LIGO antennas in Washington State and in Louisiana.

The realisation of the laser source is one of the critical difficulties: to detect sub-nanoradian changes of the interferometric phase in these kilometric Michelson interferometers requires a very stable and high power laser.

We will explain the specifications of the VIRGO laser : a power of 10 Watts and a very high spectral purity of better than 10^{-2} Hz/Hz$^{1/2}$ between 10 Hz and a few kHz (a linewidth in the submillihertz range.). Future space interferometers, like the LISA project, require similar stabilities, but in a more difficult frequency range (0.1 to 10^{-3} Hz), where thermal effects, and even ageing of the materials become relevant.

We will describe this laser and the principle of its frequency stabilisation on a high finesse, very rigid, Fabry-Perot cavity made of ULE, with a high level of thermal, seismic, and acoustic isolation.

We will report results of the first stability measurements, obtained by locking the laser to such a cavity and measuring their fluctuations with a second, identical, cavity used as a frequency discriminant.

The capability of the system to stay locked for long periods of time (typically more than a week), and to relock automatically, allows for an evaluation of the mid-term fluctuations (from 1s to 1000s). We will provide the first results on this aspect, which are important as regards the feasibility of LISA .

Introduction:

The laser source for VIRGO is an ultra-stabilized high power CW laser. Above the Fourier frequency of 200 Hz, the sensitivity of VIRGO should be limited by the shot-noise : the effective laser power required to reach the first goals is about 1 kW. This power will be obtained by using a 20W laser, and the "recycling" of the light reflected by the interferometer [Vinet et al.]

In order to split the problem of stabilizing and obtaining high power output in the laser, we have chosen the solution of injection-locking a high power monomode laser with a low power master oscillator.

In the following experiments, the low power lasers (100mW and 300 mW) were MISER type monolithic diode-pumped lasers.

The high power laser is a home-made diode-pumped Nd:YAG slab laser operating in a monomode output and injection locked by the 100mW master oscillator. The high power laser has been designed with a ring traveling-wave geometry ideally suited for its inherent directional isolation against feedback of light into the master laser. Its nominal output power is 5 W monomode, and the current operation for the injection locking experiment is 2 W.

1 - Principles of frequency stabilization and frequency noise measurement:

We use the Pound-Drever-Hall Technique : the laser beam is phase modulated at 14 MHz and the light reflected by a Fabry-Perot cavity (FP1) has power fluctuations at the modulation frequency proportional to frequency shifts. The signal coming from a fast photodiode is demodulated and sent to an electronic servosystem (see figure 1). We use a 300 mW Lightwave laser, and the cavity is described below. The open loop transfer function (including frequency filtering effect of the cavity) is 1/f from about 30 kHz to 2 MHz, $1/f^4$ from 1.5 kHz from 30 kHz, and $1/f^5$ below 1.5 kHz. The open loop unity gain bandwidth is approximately 1 MHz. In order to optimize the bandwidth and the dynamic range of the frequency corrections, we use 3 transducers : a very slow thermal one, compensates the drifts of the laser frequency below 0.1 Hz ; a faster PZT acts up to 30kHz, and an external electro-optic phase modulator corrects for the remaining high frequency phase fluctuations.

The spectrum of the error signal when the loop is closed gives informations on the electronic noise and on the remaining differential fluctuations between the laser and the cavity. To evaluate the common mode fluctuations, we need an independent measurement, which is provided by a similar second cavity : the laser beam, locked to FP1 is sent to FP2, which is used as a frequency discriminant. The error signal from FP2 contains the frequency noise of FP2, and the remaining laser frequency noise, of which it provides an upper limit, if the 2 cavities are sufficiently uncorrelated.

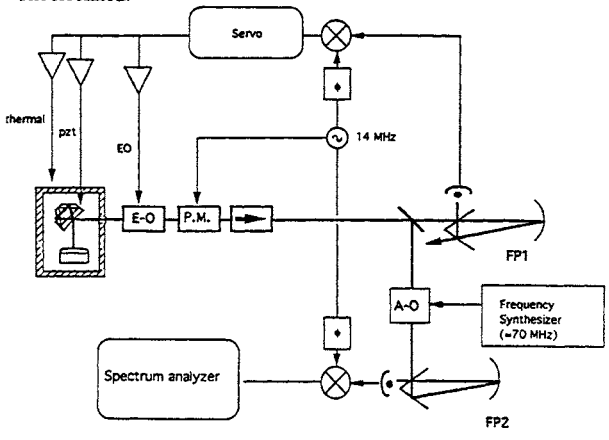


Figure 1 : Principles of frequency stabilization on a cavity and measurement of the absolute frequency noise. To have a high unity-gain bandwith (1 MHz), we use three loops.

2 - Optical and mechanical setups:

The cavities are made in ULE (Ultra Low Expansion). Their shape was designed in order to maximize the resonant frequencies (no resonance below 10 kHz, for a length of 30 cm). The material, besides its thermal properties, has a high mechanical quality factor (we measure about 6.10^4). So coupling of seismic noise with internal resonances is lowered as much as possible, and thermal noise, the final limiting noise in this experiment, is lowered too : high mass (13 kg), high resonant frequency; high quality factor. The light propagates through a cylindrical hole inside the cavity. High-quality mirrors are optically contacted to make a triangular ring cavity. The measured finesse for each cavity is about 32000, and the measured contrast is 0.6 (relative difference of reflected light at resonance and out of resonance). (see figure 2). Each cavity is in a separate vacuum chamber ; the vacuum level is about 5 10^{-8} mbar. Since the lengths of the two cavities are not strictly identical, we heat FP2 by a few degrees and to place an acousto-optic shifter in front of it in order to adjust the resonant frequency. We monitor the temperature of FP2 with a Pt resistor. The frequency shift of the acousto-optic is around 70 MHz. It is generated by a low noise frequency synthetiser. A drawback of this setup is that only FP1 can be operated close to the optimal temperature, as regards its thermal expansion (3.10^{-9} K^{-1}), while FP2 is about 5 times worse.

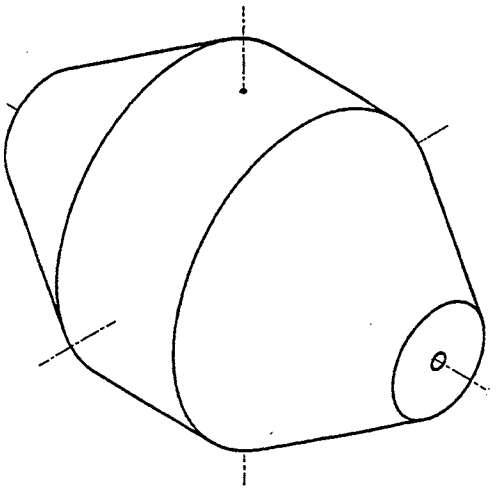


Figure 2 : ULE cavities for a Nd:Yag frequency stabilization. Shape was designed to lower thermal noise and avoid coupling of seismic noise with internal resonances of the cavity.

The error signal from the second cavity may be low-pass filtered (about 0.1 Hz), and sent to shift the frequency of the acousto-optic : we can then measure long-term drifts of the relative frequency between the two cavities, by monitoring the synthesizer frequency.
We were careful about seismic isolation. First, the tank with the experiment rests on air cushions. Second, we use a pendulum : it is a second order low-pass mechanical filter for the seismic noise in the horizontal direction above its resonant frequency (about 0.5 Hz). A gas-spring super attenuator, designed by the VIRGO Pisa group, is used to lower seismic noises in the vertical direction. Then, stems supporting the cavities act as springs and filter seismic noise above 25 Hz.
In order to alleviate temperature fluctuations and acoustic noise, the tank is covered with rock-wool.

3 - Measurements:

All precautions described above being taken, we were able to measure very low frequency noise in the range of interest for the VIRGO experiment : 2.10^{-3} Hz/√Hz at 2000 Hz (sum of the noises of the two cavities), about 1 Hz/√Hz at 8 Hz, 3.10^{-4} Hz/√Hz at 20 kHz. In the figure 3, the upper curve is the free running laser frequency noise (measured with a low-finesse cavity) ; the following curve is the error signal from FP2 ; the lower one is the error signal on FP1. By comparing the upper and lower curves, we see that the loop gain is higher than 10^4 at 10 kHz.

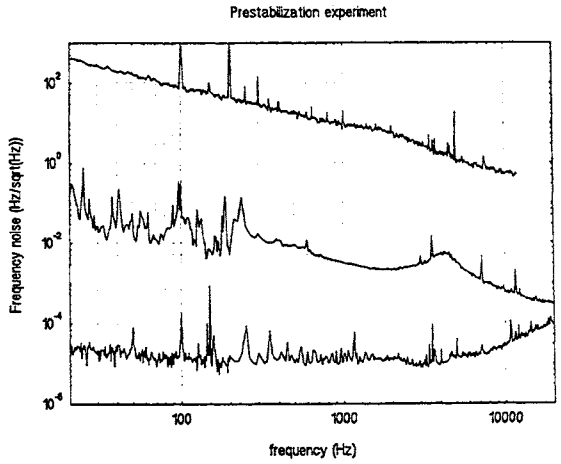


Figure 3 : Frequency noise measurement. The lowest curve is the error signal for cavity 1 (due to electronic noise), the second one is the sum of the noises of the two cavities, and the third one the free-running laser spectrum (measured on a low-finesse cavity).

The spectrum of the error signal from FP2 is partially understood : at high frequency (20 kHz), it reaches the expected shot noise level of 2.10^{-4} Hz.Hz$^{-1/2}$. Between 10 and 200 Hz, we estimate that the peaks (of seismic origine) are due mainly to the Doppler effect.resulting from relative motions of the two cavities, rather than to their length variations.
Notice that the noise amplitude at 20 kHz corresponds to relative frequency fluctuations of $\delta\nu / \nu \approx 5.10^{-19}/\sqrt{Hz}$!

For comparison with other stable oscillators, it is interesting to compute the Allan variance corresponding to this noise with the formula

$$\sigma^2(2,T,T) = \int_{f_{min}}^{f_{max}} S_y(f) 2\sin^2(\pi Tf)\left[\frac{\sin(\pi Tf)}{\pi Tf}\right]^2 df$$

where

$$S_y(f) = < y(t)y(t+T) >$$

y(t) being the ratio of the measured frequency fluctuation to the main frequency (3.10^{14} Hz). We can see then (figure 4) that we obtain 5.10^{-16} for σ at 0.02 ms and 2.10^{-15} at 0.1 s. The short term stability of this laser may be the best one ever.
The medium-term Allan variance is much worse, because of thermal variations in the tank. In the figure 5 we can see that these drifts may be of the order of 1 MHz (3.10^{-9}) on a day. They are clearly correlated to the thermal drifts that are corrected by the very-slow channel of the servo, but we don't understand yet why they are larger than expected.

Anyway, the measured frequency fluctuations fulfill the VIRGO specifications for the prestabilization stage. These results are reproduced routinely, in a rather noisy environment : people circulation, no thermal control, pulsed high power klystrons only tens of meters away, etc...

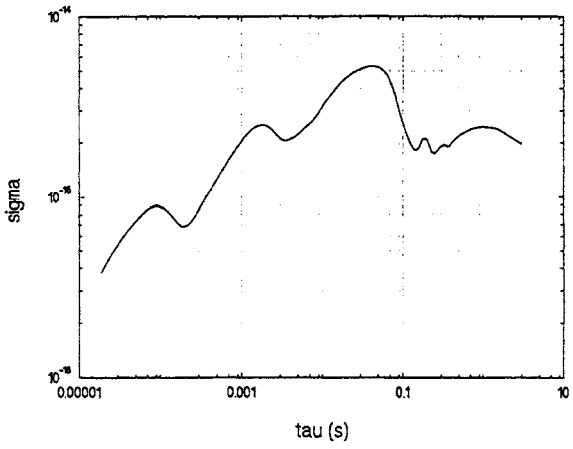


Figure 4 : Allan root variance corresponding to frequency noise.

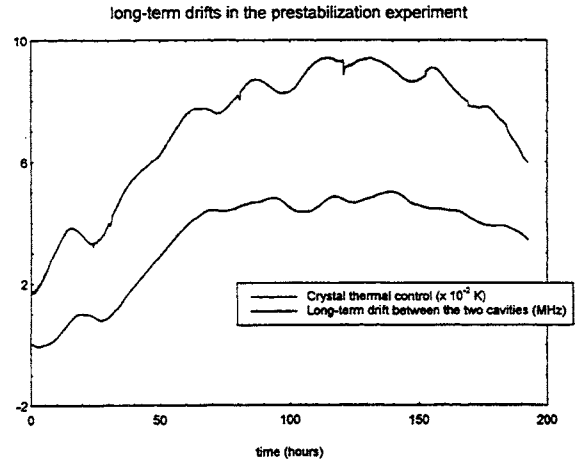


Figure 5 : Long term fluctuations of the frequency of the second cavity versus the first one. The correlation with the YAG crystal thermal control can be seen (upper curve). The laser keeps locking for more than one week.

The servo system includes an auto-locking system : if the laser unlocks (which happened only when we hit strongly the experimental tank, or when we generated a specific kind of electrical perturbation), the servoloop detects a decrease of the light transmitted by the cavity, and generates a frequency sweep of the laser, which stops when the transmission is recovered. If properly aligned, the system does not unlock more than once a week, and it relocks in a few seconds at most.

The first conclusion, important for the VIRGO experiment, is that VIRGO specifications for prestabilization stage are fulfilled. The second one is that it is a very stable oscillator indeed. The short term stability, better than 10^{-15}, is certainly one of the best ever achieved. It is still limited by seismic effects, and can easily be improved by more than one order of magnitude with one more stage of seismic isolation, and the use of vacuum for the acoustic insulation.

The system is reliable enough to study long term drifts, and the first measurements are promising. I particular, with an adequate thermal control and the operation of both cavities at optimal temperature, we know we could get close to the specifications of the LISA project.

4 - Injection-locking:

In a steady state operation of a free running laser, the circulating intensity I_0 in the cavity is such that it saturates the optical gain of the active medium to equal optical losses. An external wave tuned near resonance with such a cavity would hence give rise to an infinite circulating intensity inside the cavity. Actually this amplification is limited because of the saturation in the gain medium. The optical gain devoted to build-up the free running oscillation is then lowered enough that this oscillation is turned off. The amplification of the external wave begins to raise to the intensity I_0 , or slightly above and with a phase shifted from the external wave by an amount proportional to the frequency difference between the two waves $\Phi(\nu_1 - \nu_0)$. This condition implies also that the detuning is inside some range called the locking range expressed as:

$$\nu_1-\nu_0 \leq \Gamma = \frac{1}{2\pi} T\, FSR \sqrt{\frac{P_m}{P_s}}$$

where T is the output coupler transmission,
FSR the free spectral range,
P_S the slave power in the considered transverse mode, slave is relevant to the oscillation built in the injected cavity
P_m the power of the injected signal coupled to the slave mode, m refers to master for the injecting laser
Γ is the locking range.

The interesting properties are first the transfer of the single frequency operation from the master laser to the slave laser and we will see that the frequency fluctuations of each laser are converted into phase fluctuations at the output via $\Phi(\nu_1 - \nu_0)$ which can also be seen as frequency fluctuations since $f = \frac{d\Phi}{dt}$. From the laser equations [Siegman], we can derive the behaviour of the injected output signal $\delta\nu_2$, versus the fluctuations of the slave cavity $\delta\nu_0$ and the frequency fluctuations of the master laser $\delta\nu_1$

$$\delta\nu_2 = \frac{\Gamma}{\Gamma+f}\delta\nu_1 + \frac{f}{\Gamma+f}\delta\nu_0$$

This expression shows that the external perturbations encountered in the injection process a low-pass filtering of pole Γ (Γ being of the order of 1MHz in our experiments). Injection locking also filters the amplitude fluctuations as it acts as a second order filter with a damping ratio $\frac{\Gamma}{f_{sp}}$ and this transfer function can be written as: $\frac{1}{f_{sp}^2 - f^2 + 2jf\Gamma}$.

where f_{sp} is the frequency at which occurs the typical relaxation oscillation (100 kHz in our experiment) of solid-state laser. One can see that this relaxation oscillation of the slave laser is completely damped and needs no active servo loop. The Figure 6 shows the function $\frac{\delta Pout}{Pout} / \frac{\delta Ppump}{Ppump}$ versus frequency for the laser in the case of free running and injected (with a 1MHz locking range) operation.

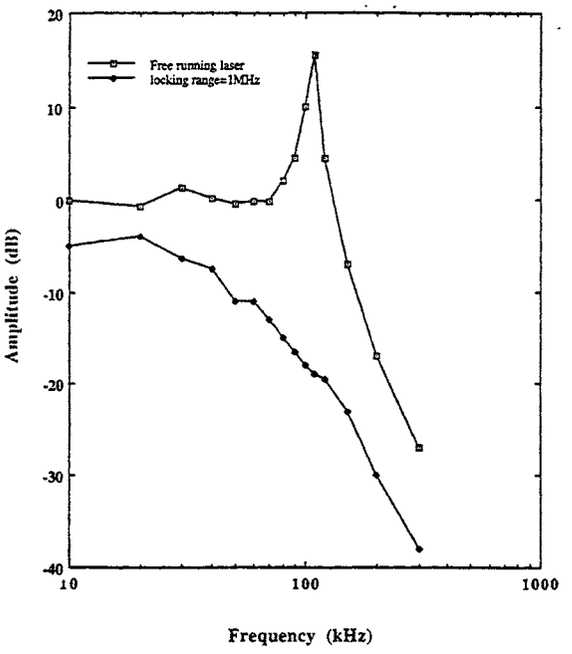


Figure 6: Relative output amplitude normalized to the relative pump amplitude fluctuations vs frequency.

A schematic of the experiment is given in Figure 7. In order to compensate for the large mechanical and acoustical perturbations, we use the Drever-Hall technique (also known as Pound-Drever) to maintain the slave laser in the injection range with a phase modulator impinging sidebands at 10 MHz and the error signal applied on two transducers carrying mirrors of the slave laser. The unity gain bandwidth of the whole servo-loop is about 100 kHz.

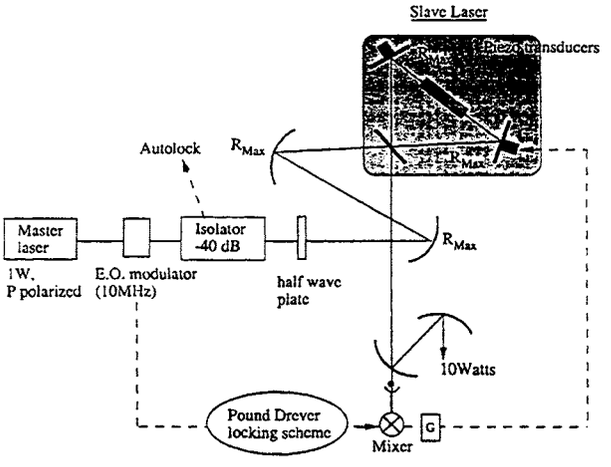


Figure 7: Optical scheme of the injection locking and the Drever-Hall locking technique.

The Figure 8 gives the residual frequency noise between the master laser and the slave injected laser vs frequency. The shot noise limitation is represented by the dashed curve and corresponds to 2 mW of light. The upper straight curve figures the noise of the free-running master laser, typically in $1/f$, to which one must compare the performance of the residual noise. We can remain that the free-running frequency noise of the slave laser is up to 300 kHz rms from 1 Hz to 1 kHz and up to 30 kHz rms from 1 kHz to 10 kHz. We can see that the performances we have obtained are largely beyond the request even with an initial noisy slave laser; the residual noise around a few kHz is due to the vibrations of the water-cooled slab, this noise will not be present in the next laser used in the Virgo experiment.

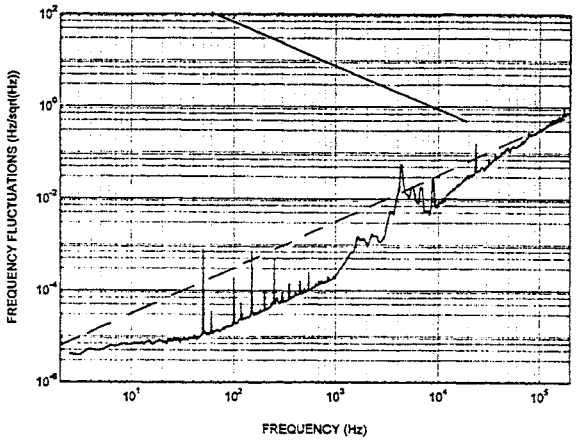


Figure 8: Residual frequency noise between the master and slave lasers. The dashed curve shows the shot noise limit calculated for 2 mW of detected light. The upper straight curve is the measured frequency noise of the master laser having a 1/f behaviour.

Conclusion:

We have described a laser system comprising a frequency stabilized laser to a high finesse optical cavity and we have obtained for Virgo the performances expected in the range of 10 Hz to 10 kHz. For other long term applications, these performances are already good compared to other prestabilized lasers and the limitations are uniquely due to the temperature drift of the cavity and they can be improved with a thermal control.

References:

A.E.Siegman, "Lasers", University Science Books, 1986
J.Y.Vinet, B.Meers, C.N.Man, and A.Brillet, Phys Rev D 38 , 2, 433 (1988)

Acknowledgement:

We wish to thank M.Dialinas for the shape design of the reference cavity, J.P.Coulon and J.Cachenaut for the electronics parts. We are grateful to REO Inc. for the optical fabrication of the cavity. Also many thanks to J.C.Lucenay and T.Redon for the mechanical helps.

POSTER SESSION

OSCILLATORS

A LOW COST OVENIZED CRYSTAL OSCILLATOR

P. Eskelinen*, J. Frilander*, J. Engelberg*, and M. Matola**

*Kotka Institute of Technology, **RF Consulting Co, Finland

ABSTRACT

A stable ovenized crystal oscillator for a normal laboratory environment can be constructed without specially manufactured quartz crystals, if minimal size is not necessary. Roughly one tenth the cost of commercial devices is achievable while maintaining a long term frequency drift below 0.01 ppm and uncertainty vs. temperature less than 0.02 ppm. The effects of thermal design, heating algorithm and mechanical structure can be matched with the final operating environment. Good co-operation between electrical engineering and mechanical design helps minimizing the cost-performance ratio.

GOALS OF THE DESIGN

This project was focused to produce an oven oscillator which could be operated with comparable uncertainty to commercial devices but restricted to a normal lab environment, where temperature is between, let´s say 15 and 30 degrees. Also the intention was to study the effects of low-cost quartz crystals and different mechanical solutions to the oscillator performance. The total cost of the unit was to be kept within a fraction of the industrial counterparts. The ultimate size of the unit was not a key parameter, as installation in a 19" rack is to be assumed.

BASICS OF OPERATION

A small crystal oscillator is encapsulated in a metal box. The temperature inside the box is measured with a thermistor. Outside the box several heating resistors are mounted. This network is connected to a regulator assembly, which performs a conversion between thermistor resistance and heating power applied. The whole oscillator module is surrounded by thermally insulating material and finally the unit together with the regulator circuit is installed in a sheet metal casing.

The idea of this construction is obviously to have a suitably large thermal mass inside the insulation and thus together with the insulating properties to produce a known thermal time constant. Besides, the double metal casing can provide good RFI protection.

The basic oscillator has a drastic effect on the final results obtainable through its own thermal behaviour and secondly by aging, as described in [1.]. To have a grisp background for the oven design several encapsulated oscillators were tested in a cyclic oven. A typical temperature drift of 0.285 ppm/ degree was obtained, which is quite reasonable for an off-the-self crystal. The measured aging was about 0.45 ppm within a three week test period as shown in Fig.1. Finally, the sensitivity of the varactor tuned frequency adjustment was measured to be 200 Hz/V.

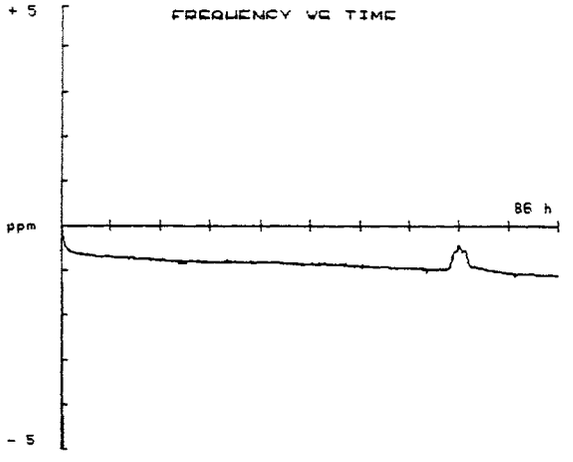


Fig.1. Crystal aging was measured for 86 hours during this test. The temperature was maintained within 0.2 degrees.

DESIGN OF THE OVEN

After having chosen the limiting dimensions a numerical thermal analysis was performed with respect to the required heating power and achievable warm-up time. An iterative model was created based on known thermophysics from e.g. [2.].

The heat transfer through the insulation can be approximated as in (1.)

$$\Phi = \lambda * A * \frac{\theta}{s} \qquad (1)$$

where A is the area of the oscillator box, s is the insulation thickness, θ is the difference in temperature and λ is the thermal conductivity of the insulation.

The temperature of the inner box can be calculated from (2)

$$T=T_s+(T_o-T_s)*e^{-\frac{A}{m*c}*t} \qquad (2)$$

where m and c are the thermal capacity and mass of the oscillator, t is the elapsed time and T_s and T_o are the inside and outside temperatures respectively.

Finally the thermal behaviour can be expressed in (3) as

$$T_s(t+\Delta t)=T_s(t)+\frac{Pke^{\frac{56.84}{T_s(t)}}-\lambda\frac{A}{s}(T_o(t)-T_s(t))}{mc}\Delta t \qquad (3)$$

where P is the maximum available heating power, k is the thermistor constant and t is the elapsed time. Results of a computer simulation based on the above assumptions suggested a final oven temperature of 46 degrees.

The placement of the heating resistors and the point where the oven temperature is measured was of major concern. The unit should preferably be operable irrespective of its attitude and the direct effect of heating to the thermistor resistance should be minimized.

The proper oven temperature is a compromise. Because no special purpose crystals were used and as the incidental oscillator power would anyhow produce a temperature rise roughly 10 degrees above ambient, the oven heating was aimed at about 50 degrees.

OVEN ELECTRONICS

In order to have a dynamic range, which is large enough a bridge type measuring chain was selected. The difference voltage is amplified and further fed to control the duty cycle of a square wave oscillator circuit. This in turn drives a power stage, which operates in switching mode and delivers power to the heating resistors. A block diagram of the oven electronics is shown in Fig. 2. A simple P-type control algorithm could also be used as this keeps the circuit design cheap and easily tuned.
The relative dynamic range of the regulator was measured to be between 14 % and 96 % of maximum available power.

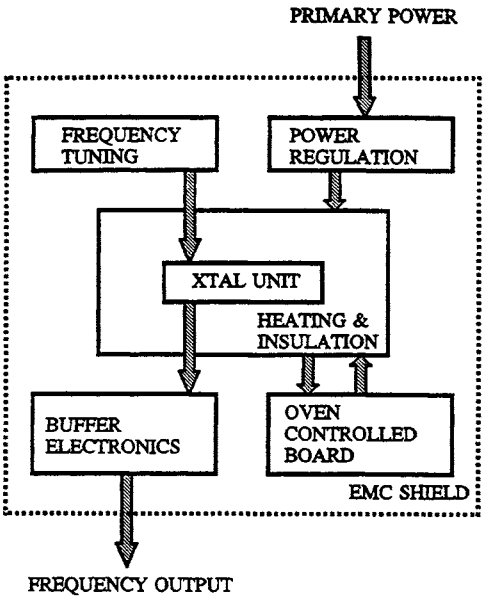


Fig.2. The oven electronics is kept simple. The block diagram is shown here with oven control and power supply.

THE CONSTRUCTIONS TESTED

Two lay-outs were used. The first prototype has a circuit board area of 45x35 mm whereas the second one is only 26x26 mm. The electronics was identical on both designs except the crystal, were an extremely low-cost room temperature unit of unknown origin (sold by Farnell Electronics) and a specially designed 46 degree oven quartz (of Quarzkeramik GmbH, a subsidiary of Rohde&Schwarz), were tried.

The prototype, which used a commercially available die cast aluminium housing is used as a reference. The oscillator box was buried in a styrofoam block. This design has obviously the benefit of easy adjustment and modifications.

The second oscillator has the smaller circuit board and was mounted inside a welded stainless steel box. The obvious profit comes from the extra thickness of insulating material which in this case was milled from a polyurethane block. Despite the smaller size, the actual thermal time constant of this package was almost six times greater (8 seconds) than that of the aluminium alloy one. Styrofoam was used here as insulating material. The module is shown in Fig. 3.

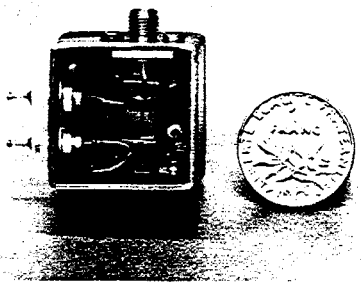


Fig.3. The final oscillator module is installed in a welded steel box.

TEST RESULTS

The separately tested parameters were the effects of

-circuit lay-out in the same thermal construction
-crystal type
-tuning varactor thermal behaviour
-supply voltage variations

Besides these, the whole oscillator with oven and surrounding insulation has been tested up to now both in normal room temperature environment and in a thermal test chamber. The test system described earlier in [3.] and improved in [4.] was used. Near-carrier measurements and analysis of phase noise was performed with a R&S modulation analyzer and an HP FFT analyzer.

One of the first results with the prototype is shown in Fig. 4. To be noted is the temperature rise of about .5 degrees at roughly 1.5 hours from start (the lowest curve of the plot) causing a rapid decrease of heating power (the slightly noisy line) from 49.2 % to 47.9 % of maximum. On the 1 ppm frequency scale no change of output frequency can be seen.

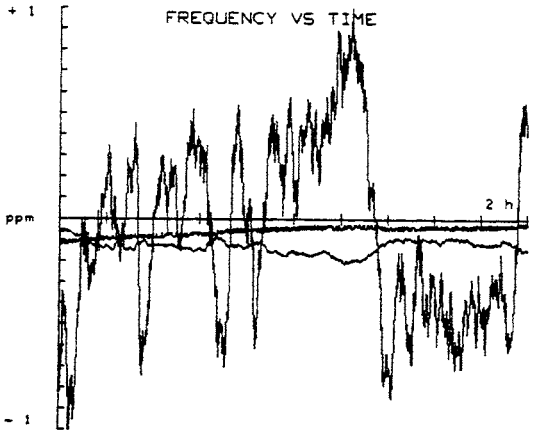


Fig.4. The very first test of the total oven configuration. Note the 0.5 degree ambient ramp at about 1.5 hours and the respective change in duty cycle.

This phenomena was simulated with the computer as well. The results are shown in Fig. 5. The same roughly linear temperature ramp of .5 degrees was assumed within 10 minutes and the initial oven temperature was set at 48.88 degrees. As can be seen, the range of heating power is from 47 to 48.3 % with a difference of 1.3 % which is just the same as the one actually measured.

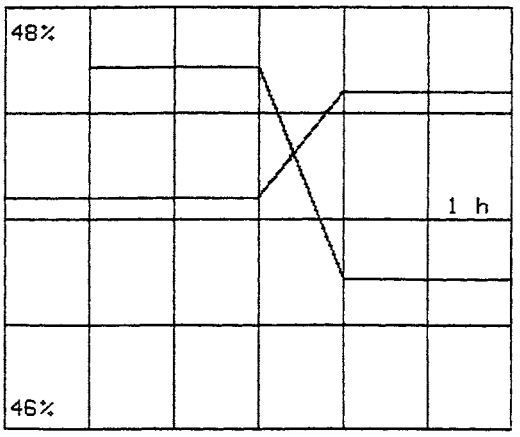


Fig.5. The phenomena of Fig. 4 simulated through a simple computer model.

The prototype oven temperature seemed to be stabilized at about 49 degrees with an average duty cycle of 0.47 and maximum available power of 2.1 W. The theoretically calculated power is less indeed, but it is obvious, since thermal leaks were unavoidable in the prototype design. Two major leaks were notified: the uneven seal between the halves of insulating material and the coaxial output cable, which was actually hot when touched with hand.

The measured oscillator performance of the two configurations is summarized in Table 1.

The achievable improvement in oscillator characteristics due to minimal board area is demonstrated in Fig. 6. As already mentioned, the thermal capacity of the smaller oscillator unit is about six times larger because of different material

A special system was built for supply voltage tests with some ideas copied from [5.]. Normally the oscillator supply voltage varies around the typical battery voltage, which in our case is nominally 14.5 V (lead-acid battery). Fig. 8 shows the results of a test where the supply voltage has been varied around its nominal. The results are summarized by linear regression in Table 2. with separate columns for oven ON and oven OFF.

Table 1. Critical parameters of the oscillator

Parameter/Value	Prototype with die-cast case	PU-foam, steel case
Size (osc. unit)	52x38x31 mm	26x31x31 mm
Weight	.073 kg	0.08 kg
Frequency	5 MHz	5 MHz
Agign (/day)	<3E-9	<4E-9
Agign (/1 s)	<4E-10	<.4E-10
Temperature drift	<1E-8 / deg	<2.8E-8 / deg
Power (warm-up)	2.85 W	3.2 W
Power (operating, 24 C)	1.35 W	1.35 W
Warm-up time	<1 h to 1E-10	<1h to 1E-10
Residual FM	<2.8 Hz (rms)	7.9 Hz (rms)
Frequency retrace	<1 E-10 (24 h OFF)	<.2E-6
Insulation	styrofoam	polyurethane
Crystal	low cost	Quarzkeramik

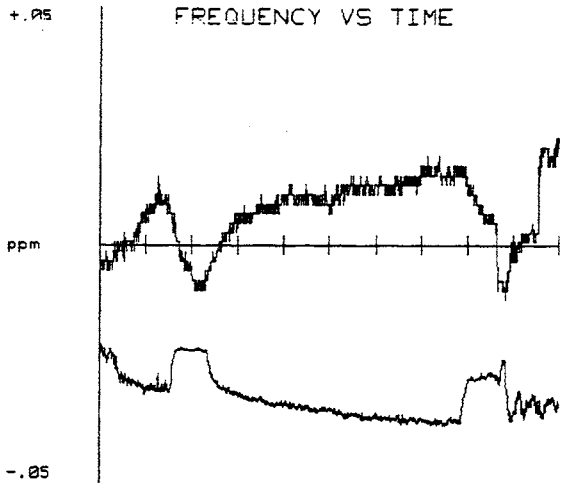


Fig.6. Frequency drift vs. temperature of the smaller (steel-case) oscillator design.

Table 2. The frequency error vs. supply voltage.

Voltage	a (bias) ppm	b (slope) ppm	e %
Dec, oven ON	0.861	-0.043	92
Inc, oven ON	0.998	-0.049	94
Dec, oven OFF	0.496	-0.035	96
Inc, oven OFF	1.009	-0.067	97

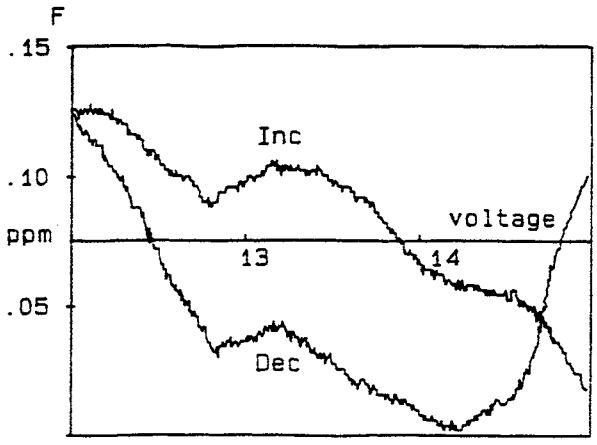


Fig.7. Frequency as a function of the supply voltage. The time interval between successive measurements was 10 seconds.

REFERENCES

1. Hewlett Packard, "Fundamentals of Quartz Oscillators", Application Note 5952-7507

2. D. Halliday, R. Resnick, J. Walker, "Fundamentals of Physics", 4th edition 1993

3. Eskelinen P. et al, "Practical arrangements of frequency and time keeping at the Kotka Institute of Technology", Sähkö-Tele 6/1993

4. Eskelinen P. et al, "Tarkkoja aika- ja taajuusmittauksia", Viestimies 3/1994

5. Leinonen S., final year project "A PC Controlled Voltage Calibrator", 1994.

Session POSTER Number

Time series analysis : an application to the stability of a crystal controlled oscillator

C. Eckert, J.A. Miehé, B. Cunin, *Groupe d'Optique Appliquée, Centre de Recherches Nucléaires, 67200 Strasbourg, France.*

The instantaneous time phase fluctuations of a crystal controlled oscillator are studied by applying false neighbours methods. The oscillator drives the acousto-optic mode-locker of a Nd:YAG laser synchronously pumping a dye laser. A synchroscan streak camera, swept with the oscillator signal, records the picosecond laser pulses. Denoting by T(t) the laser temporal jitter and by $\Theta(t)$ the time phase of the oscillator, the measured fluctuations are realizations of the random process $\Theta(t)$ - T(t). The variance of T(t) being about $(3ps)^2$, the magnitude (~50 ps) of the observed dispersions is mainly due to the oscillator.

Usual methods (correlation dimension, Lyapunov exponent) proposed for detecting low-dimensional chaos in complex time series are based on the embedding theorem. It states that a scalar time series $\{x(t_0), x(t_0 + \Delta t), x(t_0 + 2\Delta t)...\}$ of individual measurements with a sampling time Δt can be used to reconstruct the space underlying the dynamics of the system. The parameters needed to construct the vectors $\{X(t_0), X(t_0+\tau_0), X(t_0+2\tau_0)...\}$, where $X(t_i) = \{x(t_i), x(t_i+\tau)..., x(t_i + (M-1)\tau)\}$, are the embedding space dimension M, the time delay $\tau = L\,\Delta t$ and the time interval $\tau_0 = J\Delta t$ between the first components of successive vectors. Unfortunately these methods can lead to ambiguous results. It has indeed be shown that colored random noise can give rise to features characteristic of deterministic chaos, and conversely that a strange attractor can be hidden by an instrumental noise or by the superposition of a colored noise to the deterministic process.

Geometrical considerations have recently been put forward to detect a possible attractor by estimating the proper parameters M, L and J of the embedding space. It amounts to see if the nearest neighbours of a reference point are true or false neighbours, that is points which appear to be neighbours because of undervalued space dimension and/or lags. If m, l and j are the parameters of the space under test, then for a pure stochastic process the False Nearest Neighbours Percentage (FNNP) decreases by increasing m until dropping to zero when m = M. Likewise, the Averaged Wavering Product (AWP) is minimal at l = L and j = J. Concerning the crystal controlled oscillator fluctuations, the AWP is independent of l and j. This was expected since the parameters L and J are intended to rule out dynamically close points, that is points for which $x(t_i + \Delta t) \sim x(t_i)$ due to a small sampling time, and in our case Δt = 40 ms. The determination of the proper embedding dimension M is more tricky the FNNP(m) stabilizes around 5 %. However this can be understood by considering that real data always have a residual noise and can contain both a stochastic and a chaotic component. Thus the time phase fluctuations of the crystal controlled oscillator are consistent with a strange attractor embedded in a space with M lying between 3 and 6. Applying a method of noise reduction, to separate the instrumental noise and the jitter T(t) from the interesting process Θ (t), should enable to specify the proper dimension.

ANALOG ASIC DESIGN FOR MINIATURE TEMPERATURE COMPENSATED CRYSTAL OSCILLATORS

Y. RICHARD - JM TAROT - SOREP SA (EUROWAVE Division) 35220 CHATEAUBOURG (FRANCE)

ABSTRACT

An analog bipolar full custom integrated circuit has been designed in order to develop a competitive industrial Temperature Compensated Crystal Oscillator housed in a small surface mount package.

This ASIC integrates all the functions of the TCXO. Only a few resistors and a quartz crystal are setted around it. A significant feature is that standard and cheap AT cut quartz crystals can be used.

Four blocks are integrated into the ASIC : an oscillator (with a varactor diode, capacitances and an output buffer), a temperature sensor, a polynomial third order generator and a summing amplifier. The compensation technics consists to generate a third order temperature dependent voltage and to apply it to the oscillator command input. This compensation voltage fits the temperature cubic variation of the voltage to be applied to the oscillator to keep output frequency constant in all temperature range.

An automatic test bench has been developped which allows a cost effective production to be achieved.

Industrial TCXO products have been developped. Their principal performances are : 5 V power supply, 1.4 mA current consumption, ± 1 ppm frequency stability in the [-20°C, +70°C] temperature range, -135 dBc/Hz phase noise at 10 KHz from the carrier, up to 30 MHz operating frequency and a package volume of 11.5 x 12.5 x 3.8 mm^3.

INTRODUCTION

An analog bipolar full custom integrated circuit has been designed in order to create a competitive industrial Temperature Compensated Crystal Oscillator (TCXO).

The TCXO has to be miniature and cheap. To achieve these goals, it was decided to develop an ASIC which integrates all the TCXO functions.

Only a few fitting resistors and a quartz crystal are used to bring the TCXO into operation. A significant feature is that the compensation technique allows standard and cheap AT cut quartz crystals to be used.

An automatic trimming and test bench has been developped in order to minimize production costs and to allow an easy industrialization of the TCXO.

TEMPERATURE COMPENSATION

The resonant frequency drift of an AT cut quartz crystal follows a temperature cubic law :

$$\Delta f/f = A1.(T - T0) + A2.(T - T0)^2 + A3.(T - T0)^3$$

with	$\Delta f/f$	:	relative frequency drift
	T	:	temperature
	T0	:	reference ambiant temperature
	A1, A2, A3	:	quartz dependent constants.

A1, A2 and A3 constants have values which depend on the angle cut of the AT quartz crystal. So it is necessary to measure these parameters for each quartz crystal used in order to build and accurate temperature compensation.

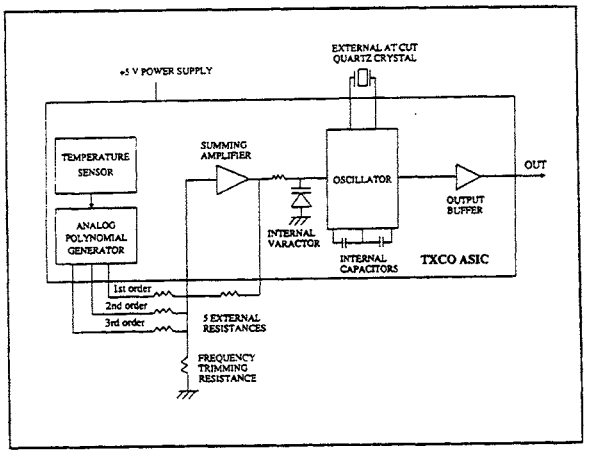
The frequency of the oscillator integrated into the ASIC is voltage controlled. This control voltage has to fit the cubic variation of the voltage to be applied in order to keep output frequency constant in all the temperature range

The analog compensation voltage is built from a balanced sum of three temperature dependent terms : the first at 1st order, the second at 2nd order and the last at 3rd order :

$$VC = P1.(V - V0) + P2.(V - V0)^2 + P3.(V - V0)^3$$

with	VC	:	control voltage of the oscillator
	V	:	voltage proportional to the temperature T
	V0	:	voltage proportional to the ambiant temperature T0
	P1, P2, P3	:	quartz oscillator dependent coefficients.

BLOCK DIAGRAM OF TCXO ASIC



INTEGRATED CIRCUIT

The integrated circuit is organized into four parts :

- a controlled voltage oscillator (with a varactor diode, capacitances and an output buffer)
- a temperature sensor
- a third order polynomial generator
- a summing amplifier.

The oscillator, a negative impedance type one, don't need external capacitances : two capacitances are integrated into the ASIC. The oscillator frequency is controlled by a command voltage applied on the integrated varactor diode via a resistance. The external quartz crystal sets the frequency in the range from a few MHz to several tens MHz.

A Band Gap generator delivers two kinds of voltage : a constant one with temperature and a proportional one to temperature. These two voltages are supplied to analog multipliers which generate the terms of the third order polynomial.

Then, each TCXO has to be caracterized individually : outputs of the analog polynomial generator (1st, 2nd and 3rd order) and output frequency are measured in function of temperature in all the temperature range. The fitting resistances are computerized and then are put on the printed circuit board.

The summing amplifier builts the command voltage and applies it to the varactor diode.

The die has been designed in order to minimize the number of components for the TCXO, to reduce external interconnections and to make the automatic trimming easier. The circuit is 5 V power supplied and the power consumption is 1.4 mA in all the temperature range. Varactor diodes takes 0.4 mm^2 and oscillator capacitances have an area of 1.5 mm^2 on the chip.

NUMBER OF TRANSISTORS

- Oscillator and output buffer	NPN :	12
	PNP :	5
- Band Gap generator	NPN :	17
	PNP :	9
- Polynomial generator	NPN :	33
	PNP :	16
- Amplifiers		
• 3 output buffer for polynomial generator	NPN :	3 x 5
	PNP :	3 x 7
• summing amplifier	NPN :	5
	PNP :	7
TOTAL	NPN :	82
	PNP :	58

Technology : f_T NPN = 3.5 GHz
f_T PNP = 10 MHZ

TRIMMING AND TEST BENCH

An automatic trimming and test bench has been designed in order to get an industrial and low cost caracterization for each TCXO. A computer controls several test cards equipped each with 64 TCXO positions. The bench takes in charge first the fitting caracterization in all the temperature range and after the final TCXO test.

RESULTS

The TCXO is housed in a miniature surface mount package of 0.6 cm^3. The major characteristics are described below :

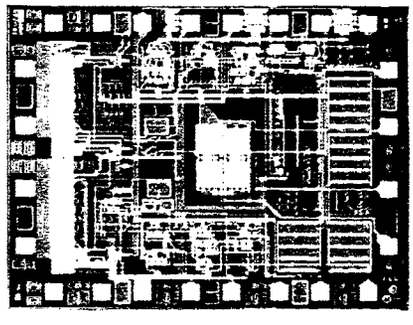
- Nominal frequency	:	[1 MHz, 30 MHz] range
- Power supply	:	5 V +/- 5%
- Load	:	10 KΩ // 10 Pf +/- 10%
- Temperature range	:	[-20°C, 70°C]
- Power consumption	:	1.4 mA
- Output voltage	:	1.5 Vpp (typ.)
- Frequency stability	:	
• with temperature	:	+/- 1 ppm MAX
• with power supply	:	+/- 0.6 ppm MAX
• with load	:	+/- 0.1 ppm MAX
- Phase noise (10 KHz from carrier	:	-135 dBc/Hz
- Volume	:	11.5 mm x 12.5 mm x 3.8 mm.

CONCLUSIONS

An analog bipolar integrated circuit has been developped in order to built a low cost Temperature Compensated Crystal Oscillator in a small volume. The TCXO product needs only a few trimming resistances and a cheap and standard quartz crystal around the ASIC to operate. An automatic trimming and test bench has also been designed to get a low cost industrialization. Electrical performances, volume and final cost makes the TCXO competitive on the market.

Reference : "An improved method of temperature compensation of crystal oscillators".
James S. Wilson - STC Components plc
Harlow - England

TCXO CHIP



FREQUENCY STANDARDS AND NOISE

A low noise 100 MHz distribution amplifier for precision metrology

M. Siccardi*, S. Römisch*, F. W. Walls†, and A. De Marchi*.

*Politecnico di Torino, Torino, Italy.

†National Institute of Standards and Technology, Boulder, USA.

Abstract

We have realized a new distribution amplifier with 0 dB direct gain. The amplifier handles up to 13 dBm at 100 MHz with low harmonic distortion and has a reverse gain smaller than -100 dB. The input and output return losses are tipically more than 25 dB. Each channel consists of one common collector NPN and two common base PNP BJT stages with a channel current of 40 mA. The power dissipation is approximately 1 W/channel using a 20 V single supply. A white PM noise floor of $S_\phi(10\,\text{kHz}) = -173\,\text{rad}^2\,\text{Hz}^{-1}$ and flicker phase noise of $S_\phi(10\,\text{Hz}) = -163\,\text{rad}^2\,\text{Hz}^{-1}$ have been measured. We have also measured a temperature coefficient of phase delay smaller than 0.4 ps/K. Circuit design phylosophy and performances will be described in detail.

1 Introduction

Distribution amplifiers are widely used in measurement systems to create independent copies of reference signals, as show in Figure 1.

Basic characteristics of this device are high outputs-to-input isolation and negligible crosstalk between outputs. Low phase noise, low distortion, good impedance match and low environmental sensitivities are required [1]. In addition, it is necessary to preserve spectral purity characteristics of distributed signals. Finally, low power dissipation is necessary for distribution system with multiple channels.

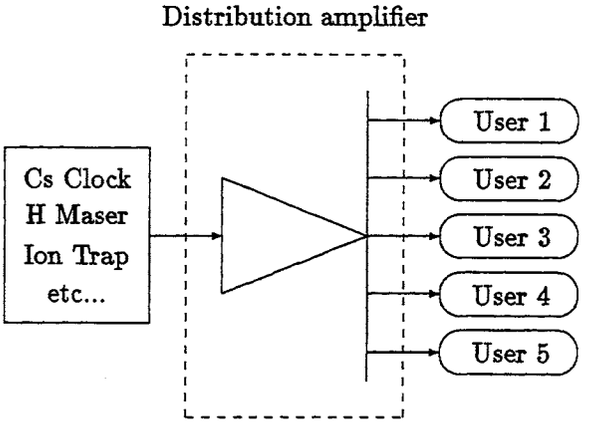


Figure 1: Distribution amplifier operation.

Distribution amplifiers suitable for 5 MHz to 10 MHz applications were reported in [3]. New distribution amplifiers for 100 MHz applications are now required by research laboratories to improve timing and frequency accuracy. The advantages of an increase in reference signal frequency in measurement systems are show in [1].

We have made a 100 MHz distribution amplifier with unitary gain, high linearity and low distorsion for input powers up to 13 dBm. The circuit design philosophy follows the same guidelines successfully used for similar devices [2][3], in which the signal is carried by current through a high isolation chain made by common-base BJT

This work was partially founded by:
- Consiglio nazionale delle Ricerche,
- Agenzia Spaziale Italiana,
- Istituto Nazionale di Fisica della Materia.

Contribution of US Government not subject to US copyright.

stages. Each chain is completed by input and output stages to build an isolating channel. A distribution amplifier is obtained by connecting channels in parallel.

2 The isolation channel

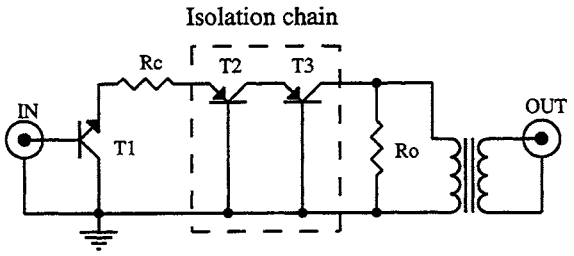


Figure 2: Basic structure of a channel.

The channel structure is shown in Figure 2. A chain of two common base BJT stages, T_2 and T_3, is current driven by the common collector BJT voltage buffer T_1 and the resistor R_c. R_c operates like a voltage to current converter. The chain made by T_2 and T_3 is a unidirectional current buffer. Its direct gain is almost unity and its reverse attenuation is very high. Resistor R_o converts the collector current of T_3 which is the chain output current, into voltage. The chain current gain is almost unity. Finally the output match is made by a transformer.

The use of BJTs with $f_T > 4$ GHz, at the operating frequency of 100 MHz, and a DC current of approximately 40 mA yields a theoretical channel reverse gain smaller than -120 dB in our configuration. This derives from the very small capacitive coupling between collector and emitter, the low impedance due to the high standing current, and the shielding between collector and emitter due to the grounded base configuration of T_2 and T_3. To improve ground connection efficency, low base resistance BJTs were used.

The common collector input stage makes the channel input impedance independent of the amplitude of the current signal going to the isolation chain. The channel input impedance has a design value of 250 Ω in order to set the device input impedance to 50 Ω. The voltage-to-current conversion network is made by the series of buffer (T_1) output impedance, resistor R_c and chain (T_2) input impedance. The latter is a function of quiescent current inside BJT T_2 and it can be viewed as a virtual short to ground. The voltage buffer output impedance is a function of the quiescent current inside T_1 and can be neglected. An admittance of 20 mS is required to obtain a 20 mA rms chain current from an input signal of 13 dBm.

The value of resistor R_o (200 Ω) and the transformer ratio (2 to 1) allow a theoretical output signal of 13 dBm for the nominal input power. Impedance match of the output is also obtained by this approach.

Signal losses in BJTs and transformer are compensated by increasing the value of resistor R_o to 302 Ω. The result is a channel gain of 0 dB and output match to 50 Ω.

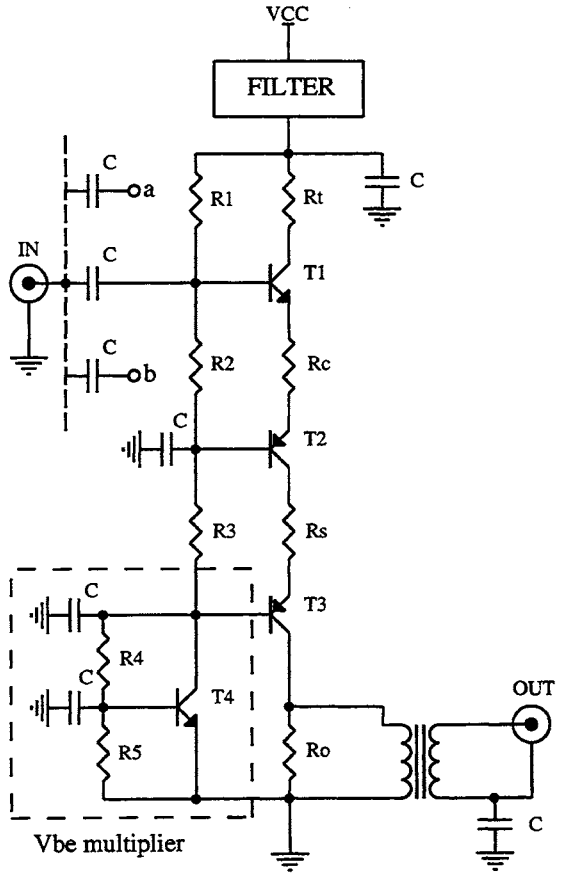


Figure 3: Schematics of a biased channel: on the dashed line more channels can be connected in parallel. As an example 'a' and 'b' can be viewed as the channel entry point (after the input's capacitor) for two of these.

By choosing an NPN BJT for T_1 and a PNP BJT for T_2 and T_3 the channel is biased by a

single primary current that flows through all active devices, as shown in Figure 3. This current starts from a positive voltage supply and ends at ground. In this way the output stages are not connected to any DC supply. Consequently the channel-to-channel crosstalk through supply lines is greatly reduced.

The quiescent current coming from the isolation chain is grounded by the primary winding of the output transformer. This reduces the DC voltage needed at the collector of T_3 to avoid cutoff when the load is open and hence reduces the power dissipation. As a result, harmonic distortion may be introduced due to saturation in the transformer's core by high values of DC current. Thus, the harmonic distortion is increased by increasing DC current. On the other hand, the BJTs quiescent collector current must be greater than the peak signal current. The harmonic distortion due to the non-linear behavior of BJTs increases by decreasing the quiescent current. Measurements on our particular choice have shown that a quiescent current of 38 mA DC minimizes the global harmonic level functions, which therefore is the design value.

The ratio of bias voltage across emitters of T_1 and T_2 and value of the resistor R_c set the primary quiescent current. This current changes with emitter bias voltage due to the dispersion of h_{fe} values in BJTs. To reduce this problem a V_{be} multiplier is used instead of a classical resistor to connect the base of BJT's T_3 to ground. The V_{be} multiplier is realized by transistor T_4. This approach decreases the current dispersion from channel to channel to less than 5%.

To avoid possible DC ground loops in measurement systems that include this device, the output ground is connected only by capacitor to the device main ground.

3 Realized devices and measurements

A distribution amplifier is made by connecting the inputs of several isolating channels in parallel. The input impedance of the device is, therefore, the parallel of several channel's input impedances. Different versions with different number of channels were made. In devices with a number of channels lower than five, a shunt resistor was connected at the input node to emulate

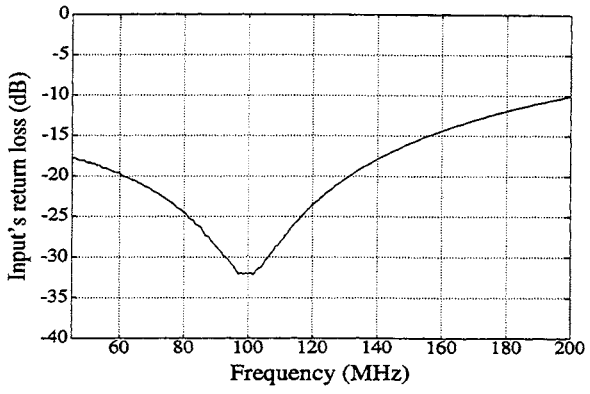


Figure 4: Plot of the device's input return loss.

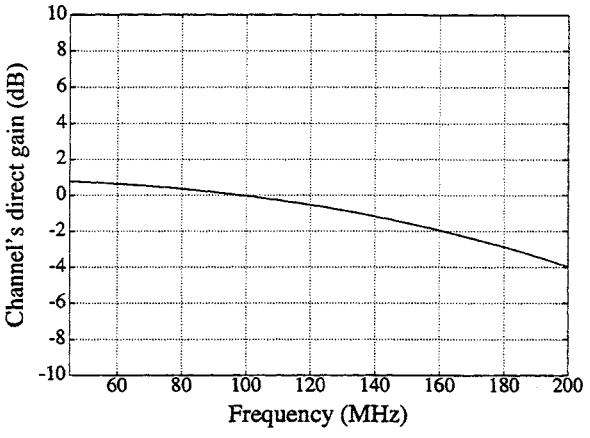


Figure 5: Plot of a channel's direct gain.

the omitted channels.

The reverse channel gain at 100 MHz is smaller than -100 dB. Channel to channel crosstalk is smaller than -100 dB.

Harmonic distortion is -37 dBc at 200 MHz and -40 dBc at 300 MHz.

Phase noise was measured with a crosscorrelation technique [4]. The noise of the measurement system was always 5 dB below the noise in measured devices.

Measurement results are shown in Table 1. The phase delay change with temperature was

$f(Hz)$	10	100	1000	10000
$S_\phi(\text{rad}^2\,\text{Hz}^{-1})$	-163	-170	-172	-173

Table 1: Phase noise at different Fourier frequency.

about 0.3 ps K^{-1}.

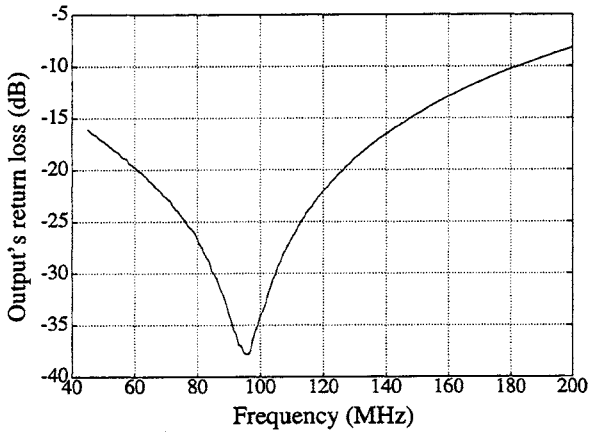


Figure 6: Plot of a channel's output return loss.

The measurements of the scattering parameters are shown in Figure 4, 5 and 6.

4 Conclusion

Measurements show exellent phase noise, isolation, and temperature coefficient of phase delay for the new 100 MHz distribution amplifier. The philosophy of carrying signal via current seems to be a winning way for designing distribution amplifiers [3][2].

5 Acknowledgments

We thank Andrew Novick for help with the printed circuit boards. We also thank Franklin Ascarrunz and Craig Nelson for useful talk about design and realization of this device. Finally we thank dr. Borsero, dr. Brunetti and dr. Novero for the loan of a network analizer for the final scattering parameter measurements.

References

[1] L. M. Nelson, and F. L. Walls, "Environmental effects in mixers and frequency distribution systems," in *Proc. IEEE Freq. Contr. Symp, pp. 831-837*, (1992).

[2] A. De Marchi, F. Mussino, M. Siccardi, "A high isolation low noise amplifier with near unity gain up to 100 MHz," in *Proc. IEEE Freq. Contr. Symp, pp. 216-219*, (1993).

[3] C. W. Nelson, F. L. Walls, M. Siccardi, A. De Marchi, "A new 5 and 10 MHz high isolation distribution amplifier," in *Proc. IEEE Freq. Contr. Symp, pp. 567-571*, (1994).

[4] W. F. Walls "Cross-correlation phase noise measurements," in *Proc. IEEE Freq. Contr. Symp, pp. 257-261*, (1992).

LOW FREQUENCY NOISE REDUCTION IN DMTD SYSTEMS

Mihai P. Dinca

Faculty of Physics, University of Bucharest, PO BOX MG-11,
Bucharest-Magurele, 76900 ROMANIA

Abstract - We present a numerically study on the effect of the low frequency rejection of the additive noise upon the noise floor of the DMTD frequency stability measuring system. White, flicker and random walk noise and several kinds of RF filtering transfer functions were investigated.

1. Introduction

The effect of the different noise sources upon the noise floor of the dual mixer time difference (DMTD) system, used to measure the frequency stability in the time domain, was evaluated by Barillet [1] and the importance of the beat signals filtering before the zero-crossing detector has been emphasized. Although, the approximation of the RF filtering by a phase filtering apply only for low-pass transfer functions, provided that the cut frequency is sufficiently high. Since the analytical formulae given by Tremblay for RF filtering [2] do not refer to flicker and random walk noise filtered by pass-band transfer function, in order to study the effect of the low frequency rejection, we investigated numerically the contribution to measured Allan variance, $\sigma_y^2(\tau)$, produced by the white, flicker and random walk noise, for several kinds of filters which can be used in the DMTD systems. Both cases, when the noise is superimposed over the RF signal (additive noise), and when the noise is a phase noise of the oscillator itself, were considered with the aim to improve the capability of the system to characterize less noisily oscillators.

2. Effect of the RF filtering in the presence of additive noise.

Each of the beat signals can be expressed by

$$v_o(t) = A_o\left[1+\varepsilon_o(t)\right]\cos\left[2\pi\nu_o t+\varphi_o(t)+\theta_o\right] \qquad (1)$$

where A_o is the signal mean amplitude, $\varepsilon_o(t)$ is its relative amplitude fluctuations, ν_o is its mean frequency, $\varphi_o(t)$ is its phase fluctuations and θ_o is its initial phase, considered as uniformly distributed.

If the signal is perturbed by an additive noise and the amplitude noise, phase fluctuations and additive noise are low, the composite signal can be put in a form similar to (1), i.e.

$$v_{co}(t) = v_o(t)+n_o(t) = A_o[1+\varepsilon_{co}(t)]\cos[2\pi\nu_o t+\varphi_{co}(t)+\theta_o] \qquad (2)$$

Then, for the spectral densities we have

$$S^{TS}_{\varepsilon co}(f) = S^{TS}_{\varepsilon o}(f)+\frac{1}{A_o^2}S^{TS}_{no}(f)\otimes N(f)$$

and (3)

$$S^{TS}_{\varphi co}(f) = S^{TS}_{\varphi o}(f)+\frac{1}{A_o^2}S^{TS}_{no}(f)\otimes N(f)$$

the convolution product being performed with the function

$$N(f) = \delta(\nu_o+f)+\delta((\nu_o-f)\,. \qquad (4)$$

When the composite signal $v_{co}(t)$ is passed through a linear and time invariant filter, with the transfer function $H(f)$, the output signal can be expressed, in its turn, in a form similar to (1)

$$v_C(t) = A_o\left|H(\nu_O)\right|\left[1+\varepsilon_C(t)\right] \times\cos\left[2\pi\nu_O t+\varphi_O(t)+\theta_O+\psi(\nu_O)\right] \qquad (5)$$

where $|H(\nu_o)|$ and $\psi(\nu_o)$ are the magnitude and phase of $H(\nu_o)$, respectively.

If the amplitude and phase fluctuations, $\varepsilon_o(t)$ and $\varphi_o(t)$ of the initial beat signal are supposed as uncorrelated, the spectral densities of the filtered fluctuations have the form:

$$S^{TS}_{\varphi c}(f) = |H_c(f)|^2 S^{TS}_{\varepsilon o}(f)+|H_d(f)|^2 S^{TS}_{\varphi o}(f) +\frac{1}{A_o^2|H(\nu_o)|^2}\left[S^{TS}_{no}(f)|H(f)|^2\right]\otimes N(f)$$

and (6)

$$S^{TS}_{\varepsilon c}(f) = |H_d(f)|^2 S^{TS}_{\varepsilon o}(f)+|H_c(f)|^2 S^{TS}_{\varphi o}(f) +\frac{1}{A_o^2|H(\nu_o)|^2}\left[S^{TS}_{no}(f)|H(f)|^2\right]\otimes N(f)$$

Here, the transfer functions $H_d(f)$ and $H_c(f)$ are the symmetrical and, respectively, anti-symmetrical part of $H(f)$ around ν_o:

$$H_d(f)=\frac{1}{2}\left[\frac{H(\nu_o+f)}{H(\nu_o)}+\frac{H^*(\nu_o-f)}{H^*(\nu_o)}\right]$$

and (7)

$$H_c(f)=\frac{i}{2}\left[\frac{H(\nu_o+f)}{H(\nu_o)}-\frac{H^*(\nu_o-f)}{H^*(\nu_o)}\right] .$$

The Allan variance of the fractional frequency fluctuations for the filtered composite signal can be computed from $S_{\varphi c}^{TS}(f)$ as

$$\sigma_y^2(\tau)=2\int_{-\infty}^{+\infty}S_{\varphi c}^{TS}(f)\frac{\sin^4(\pi\tau f)}{(\pi\tau\nu_o)}df \quad . \tag{8}$$

Due to the fact that the variance is measured only for averaging time

$$\tau=nT_o=\frac{n}{\nu_o} \tag{9}$$

with n integer, the effect of the convolution on the additive noise spectral density disappears, and, finally, the Allan variance can be expressed in terms of the one sided spectral densities, by

$$\sigma_y^2(\tau)=2\int_0^{\infty}\left[S_{\varphi o}(f)|H_d(f)|^2+S_{no}(f)\frac{1}{A_o^2}\left|\frac{H(f)}{H(\nu_o)}\right|^2\right]\times\frac{\sin^4(\pi nf/\nu_o)}{(\pi n)^2}df \tag{10}$$

provided that the contribution of the amplitude noise can be neglected.

3. Numerical calculation of the Allan variance

The spectral density of the additive noise was supposed of the form

$$S_{no}(f)=h_{n\beta}f^{\beta} \tag{11}$$

and β = 0, -1 and -2, i.e. white, flicker and random walk were considered. The same model was used for the phase fluctuations

$$S_{\varphi o}(f)=h_{\varphi\beta}f^{\beta} \tag{12}$$

with the same values for β. Thus, the white phase, flicker phase and random walk phase (or, equivalently, white frequency) noises were investigated.

We computed the integrals of the form

$$I=\int_0^{\infty}\frac{|H(f)|^2}{f^{\beta}}\sin^4(\pi nf/\nu_o)df \tag{13}$$

on two domains of the integration variable, separately. Between f = 0 and $f=f_1=\nu_o/n$ the integral was performed by the Simpson method. Over the frequency f_1, the identity

$$\sin^4\alpha=\frac{3}{8}-\frac{1}{2}\cos 2\alpha+\frac{1}{8}\cos 4\alpha \tag{14}$$

was used to decompose the integral in three terms, separately computed [3].

The first term was integrated by the Simpson method over n_d decades starting with f_1 The number n_d is each time selected large enough to may approximate the squared magnitude of the transfer function, for the remainder of the integration domain, by a power law and the integral was analytically evaluated.

The last two terms lead to oscillatory integrals which were performed, over the n_d decades, starting from f_1, by using the Filon method [4].

4. Results

From eq. (10) results that in obtaining the Allan variance of the fractional frequency fluctuations, $\sigma_y^2(\tau)$, a certain filter with transfer function $H(f)$ acts upon the phase fluctuations of the signal itself and upon the additive noise in different manners. Since, generally, the RF filtering affects the both terms in (10), and the aim of the frequency stability measurement is to characterize the signal itself, despite the additive noise, we introduced the filtering merit function F_β by

$$F_\beta(\tau)=\frac{\sigma_{y\varphi}^2(\tau)h_{n\beta}}{\sigma_{yn}^2(\tau)h_{\varphi\beta}} \tag{15}$$

where $\sigma_{y\varphi}^2(\tau)$ and $\sigma_{yn}^2(\tau)$ are the two terms in (10) and $h_{n\beta}$ and $h_{\varphi\beta}$ have the meaning given by (11) and (12). In the next we assume, for simplicity, an unity value for $h_{n\beta}$ and $h_{\varphi\beta}$ and, then, the merit function reduces to

$$F_\beta(\tau)=\frac{\sigma_{y\varphi}^2(\tau)}{\sigma_{yn}^2(\tau)} \quad . \tag{16}$$

For power spectral densities of phase fluctuations with β = -2, -3 and -4, i.e. random walk phase (white frequency), flicker frequency and random walk frequency fluctuations, the Allan variance is independent of the filtering transfer function, provided that $2\pi f_c\tau>>1$, where f_c is the high cut frequency of the transfer function $H_d(f)$. This condition is satisfied in the

range of τ where these kinds of noise are dominant and, thus, the measured variance produced by these noises is not affected by the filtering process. The ability of the measuring system to evaluate them depends only on the system noise floor, determinated by the contribution of the additive noise. From this point of view, it is necessary only to diminish the value of $\sigma_{ym}^2(\tau)$.

On the contrary, if phase fluctuations with $\beta = 0$ and $\beta = -1$ have to be measured (i.e. white phase and flicker phase noise) the attempt to decrease the additive noise contribution will decrease in the same time the quantity to be measured. For this situation, the filtering merit function must be taken into consideration and its value increased by a proper selection of the filter transfer function. Therefore, several kinds of filter were studied by computing the dependence of $\sigma_{y\varphi}^2(\tau)$ and $\sigma_{ym}^2(\tau)$ for white, flicker and random walk noise. The nominal frequency ν_O was set 1 Hz and unity values were considered for $h_{\varphi\beta}$, $h_{n\beta}$ and A_O.

a) First order low-pass filter.

A filter with the transfer function

$$H_1(f) = \frac{f_c}{if + f_c} \tag{17}$$

was assumed to process the beat signals. in the frequency stability measuring system. For $f_C >> \nu_O$ we have

$$|H_{d1}(f)|^2 \cong |H_1(f)|^2 \tag{18}$$

and the filtering merit function equals unity. The decreasing of the cut frequency f_C proportionally reduces the value of $\sigma_{ym}^2(\tau)$ for white noise. The reduction is much lower for the flicker noise for which $\sigma_{ym}^2(\tau)$ vary with ln f_C, and no reduction is obtained for random walk noise.

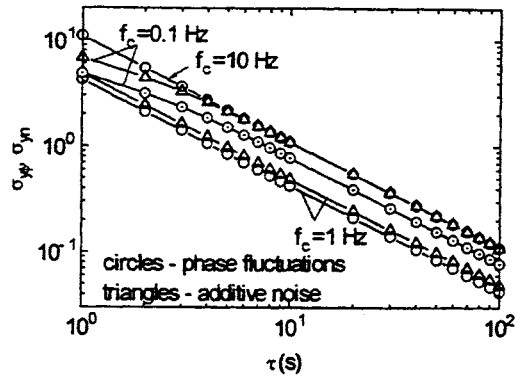


Fig. 1. White noise filtered by a first order low-pass transfer function.

If the cut frequency is decreased to values comparable to ν_O, the eq. (18) is no more valid and the merit function is reduced bellow unity. However, the both transfer functions exhibit the same low-pass behaviour which produces similar dependencies upon τ of $\sigma_{y\varphi}^2(\tau)$ and $\sigma_{ym}^2(\tau)$. The lowest $\sigma_{ym}^2(\tau)$ is obtained when $f_C = \nu_O$ [2]. Fig. 1 presents the results obtained in the above mentioned conditions for white noise and several values of f_C.

If the cut frequency is further reduced, $\sigma_{ym}^2(\tau)$ begin to rise and the filtering merit function continue to decrease. Again, for flicker noise the obtained variations are much lower than those for white noise.

b) Non resonant second order band-pass filter.

The transfer function we consider here is

$$H_2(f) = \frac{if}{if + f_1}\frac{f_2}{if + f_2} \tag{19}$$

which, for $f_1 = 0$ reduces to the previous case.

The corresponding function $H_{d2}(f)$, being its symmetrical part around ν_O, has a low pass behaviour and, then, the processing of the phase fluctuations and additive noise are always different.

The value of f_2 was 1 Hz and for the low cut frequency we considered the values 0, 0.01, 0.1 and 1 Hz in order to study the effect of the low frequency rejection. The low cut frequency affects very slightly the transfer function $H_{d2}(f)$ and, therefore, $\sigma_{y\varphi}^2(\tau)$ is practically independent of f_1. On the contrary, the function which filters the additive noise is modified in the low frequency range.

For white noise, due to the flat distribution of the power through the spectrum, low variations of the $\sigma_{ym}^2(\tau)$ values were observed when f_1 is changed and negligible improvement of the filtering merit function can be obtained.

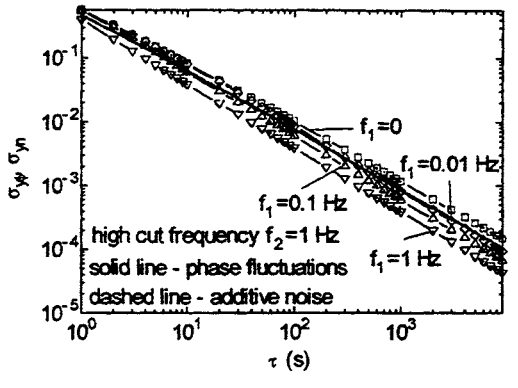


Fig. 2. Flicker noise filtered by a non resonant second-order band-pass transfer function.

In the case of flicker noise, as one can observe in Fig. 2, the presence of the low cutting slightly changes the slope of $\sigma_{ym}^2(\tau)$ dependence and the noise floor is reduced

with the decreasing of the filter pass-band. The filtering merit function increases with τ and it is enhanced when the pass-band become more narrow.

The existence of the low cutting produces for random walk noise the doubling of the $\sigma_{yn}^2(\tau)$ slope. From Fig. 3 result that the noise floor is reduced inverse proportionally to f_1 for values of f_1 below ν_o. Consequently, appreciable improvements are obtained for the merit function.

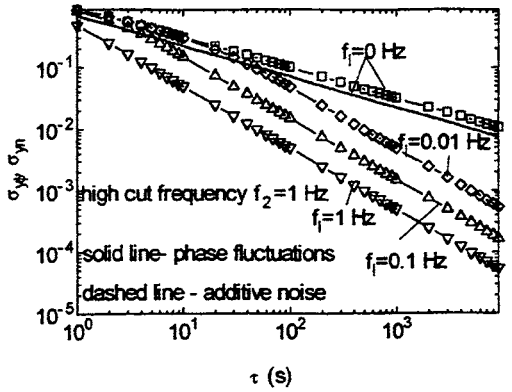


Fig. 3. Random walk noise filtered by a non resonant second-order band-pass transfer function.

c) Resonant second order band-pass filter.

Further reducing of the noise floor and enhancement of the merit function are produced if a filter with transfer function

$$H_3(f) = \frac{if\nu_o}{if\nu_o + Q(\nu_o^2 - f^2)} \qquad (20)$$

is used. Values of 1, 5 and 10 were considered for Q.

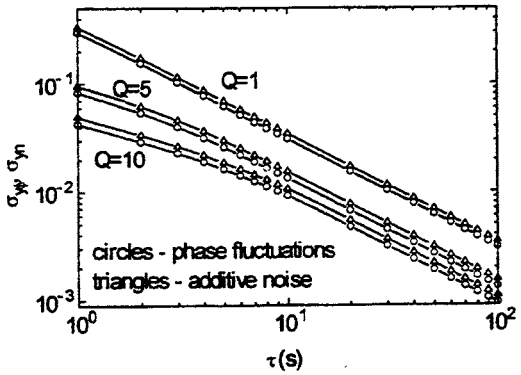


Fig. 4. White noise filtered by a resonant second order band-pass transfer function.

For white noise $\sigma_{y\varphi}^2(\tau)$ and $\sigma_{yn}^2(\tau)$ are reduced, the change in the merit function continuing to be negligible (Fig. 4). It can be observed in Fig. 5 than the noise floor is reduced for the flicker noise, too. As $\sigma_{y\varphi}^2(\tau)$ is practically unaffected by the filtering, the merit function is increased up to values around 10.

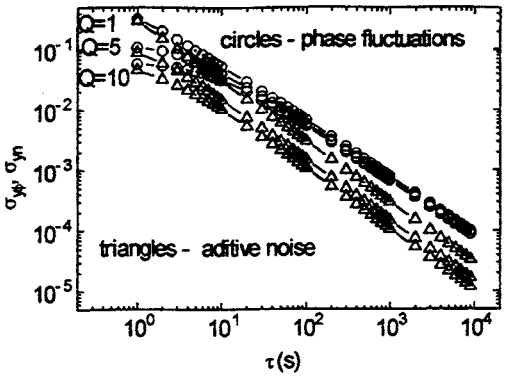


Fig. 5. Flicker noise filtered by a resonant second-order band-pass transfer function.

The random walk noise exhibits a behaviour similar to that obtained for the non resonant bad-pass filter, but the improvements are greater.

5. Conclusions

The resonant pass-band filter with high Q appears to be the most efficient in reducing the noise floor of the measuring system. The greatest improvements are obtained for random walk noise where the slope of $\sigma_{yn}^2(\tau)$ is doubled. A little change in the slope is observed for flicker noise, while for the white noise the slope remains unchanged. Moreover, for all kinds of noise we considered, the noise floor was reduced by the decreasing of the equivalent noise band.

However, the ability to measure $\sigma_{y\varphi}^2(\tau)$, characterized by the filtering merit function, can not be improved for white phase fluctuations. In the case of flicker phase fluctuations, the merit function can be obtained several times greater than unity. The effect is greater for the random walk phase fluctuations, for which the merit function can be multiplied by a factor up to one thousand.

References

1. R. Barrilet, "Comparateur de phase ultra faible bruit pour les futurs etalons de frequence", Proc. 3rd European Frequency and Time Forum, March 1989, Besancon.

2. P. Tremblay, "Etude de l'effect du filtrage sur la caracterization de la stabilite de la frequence"' PhD dissertation, Universite Laval, Quebec, Canada, 1985.

3. F. L. Walls et al., "Time domain frequency stability calculated from the frequency domain: an update", 4th European Frequency and Time Forum, March 1990, Neuchatel.

4. P. Davis and P. Rabinowitz, "Numerical integration", Blaisdell Publ. Co., Waltham, Massachusetts, 1967.

CUT-OFF FREQUENCIES AND NOISE POWER LAW MODEL OF SPECTRAL DENSITY : ADAPTATION OF THE MULTI-VARIANCE METHOD USING THE STRUCTURE FUNCTION APPROACH

F. Vernotte, M. McHugh, G. Zalamansky
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

Abstract - The noise levels, the cut-off frequencies and the drift coefficients are objective parameters for time stability characterization. In order to estimate these parameters, variances may be used. The multi-variance analysis, which uses several variances over the same signal, yields very accurate measurements of these parameters if the choice of variances is appropriate to the signal. Using the structure function approach, it is possible to select an optimized set of variances.

1. INTRODUCTION

Variance measurements are often confused with a direct estimation of the time instability of an oscillator. These two quantities are strongly correlated but they are not strictly equal. For example, two variances may exhibit different responses for the same noise. On the other hand, the noise levels given by spectrum analysis, depend only on the oscillator. The noise levels may then be considered as objective parameters. Variance measurements are easy to translate into noise level measurements. Thus, the variances may be used as estimators of the power spectrum.
In order to completely characterize the time instability of an oscillator, the cut-off frequencies should also be estimated as well as the drifts which may also be measured by variances. However, the determination of all these parameters cannot be performed with only one variance : it is necessary to define a set of variances adapted to each problem.

2. TIME INSTABILITY CHARACTERIZATION : OBJECTIVE PARAMETERS

2.1. The power law model of spectral density

The Power Spectral Density (PSD) of time error of an oscillator may be modeled as a sum of power laws :

$$S_x(f) = k_0.f^0 + k_{-1}.f^{-1} + k_{-2}.f^{-2} + k_{-3}.f^{-3} + k_{-4}.f^{-4} \qquad (1)$$

As the instantaneous normalized frequency deviation y(t) is the derivative of the time error x(t), the PSD of y(t) may be modeled as :

$$S_y(f) = h_{+2}.f^{+2} + h_{+1}.f^{+1} + h_0.f^0 + h_{-1}.f^{-1} + h_{-2}.f^{-2} \qquad (2)$$

with :

$$h_\alpha = 4\,\pi^2\,k_{\alpha-2} \qquad (3)$$

Thus, the 5 noise levels h_α are objective parameters for the time instability characterization since they depend only on the oscillator.
In some cases (e. g. the study of millisecond pulsars), very low frequency noise levels (h_{-3} and h_{-4}) may be added.

2.2. Cut-off frequencies

2.2.1. High cut-off frequency

For the high frequency noises (white phase noise and flicker phase noise), a high cut-off frequency f_h must exist since the power is finite. This high cut-off frequency may be due to the measurement device or to the noise itself (it is the smaller of these two cut-off frequencies).
For a white phase noise, for instance, the effects of spectral aliasing completely modify the estimation of the PSD : it is not possible to distinguish between a very high cut-off frequency with a low noise level and a lower high cut-off frequency with a higher noise level. Only the product $k_0.f_h$ may be measured.
The high cut-off frequency is then a characteristic feature of the signal and must be measured.

2.2.2. Low cut-off frequency

On the other hand, for low frequency noises, a low cut-off frequency f_l must exist.
In the case of a low cut-off frequency smaller than the inverse of the duration of the measurement sequence, the very long term fluctuations may be confused with deterministic drifts (see figure 1). Hence, the low cut-off frequency should also be determined.

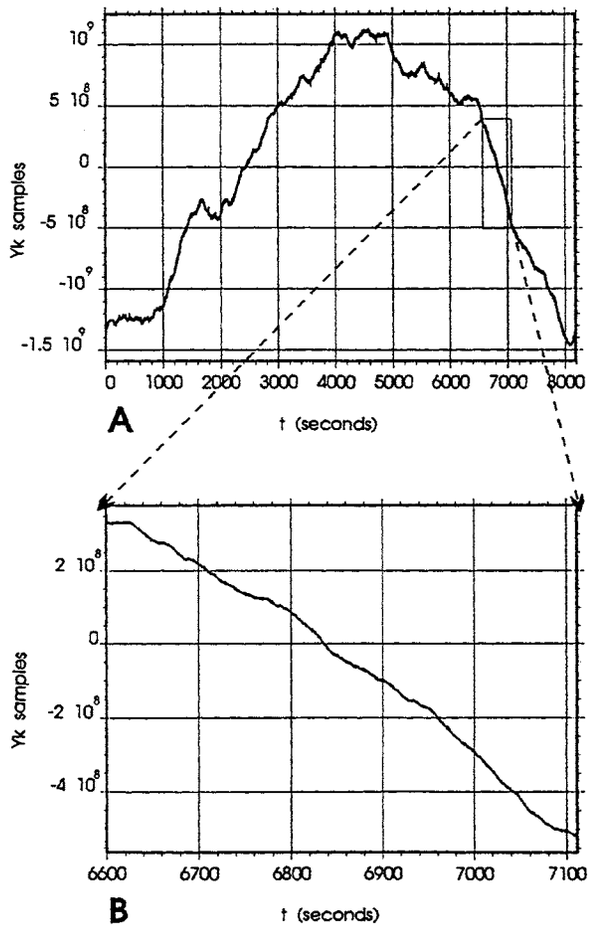


Figure 1 : Figure 1.A shows a f^{-3} noise over a time interval equal to the inverse of its low cut-off frequency. The enlargement of figure 1.B shows that the long term random fluctuations may be confused with a linear drift.

2.3. Drifts

Oscillators may be affected by linear phase drift and quadratic phase drift, i. e. constant and linear frequency drifts.
These deterministic drifts must be distinguished from the "artificial drift" due to very long term fluctuations (see figure 1). It is possible to estimate a limit for the coefficient of artificial drift :

- for a f^2 frequency noise : $C_1 < 4\pi\sqrt{h_{-2}}$ (4)
- for a f^3 frequency noise : $C_1 < 4\pi\sqrt{\frac{h_{-3}}{f_l}}$ (5)

Thus, if the measured coefficient C_1 of linear frequency drift is greater than this limit, it is a deterministic drift, whereas, in the opposite case, it may be considered as a part of a very long term random fluctuation.

2.4. Objective parameters

Thus, in order to completely characterize an oscillator (a classical one as well as a pulsar), the list of objective parameters could be :

- 7 noise levels : h_{+2}, h_{+1}, h_0, h_{-1}, h_{-2}, h_{-3} and h_{-4}
- 2 cut-off frequencies : f_h and f_l
- 2 drift coefficients : C_0 and C_1.

3. ESTIMATION OF THE OBJECTIVE PARAMETERS

3.1. Variances

The determination of the parameters (especially the noise levels) is usually performed by using variances. The variances are defined in the time domain by their measurement sequence. For example, the measurement sequence of the Allan variance is :

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle\left(\overline{y_2} - \overline{y_1}\right)^2\right\rangle \quad (6)$$

The principle of a variance measurement consists of calculating the variance with different integration times τ and then plotting these measurements versus τ on a log-log graph. Since the variances exhibit different slopes for the different types of noises, the identification of each slope yields the identification of each type of noise. For a f^α noise, a variance exhibits a τ^μ slope, and generally $\mu = -1 - \alpha$. The h_α coefficients can then be computed from Table I.

	Variance	
	Allan	modified Allan
$S_y(f)$	$\sigma_y^2(\tau)$	Mod $\sigma_y^2(\tau)$
$h_2.f^2$	$\frac{3 h_2 f_h}{4\pi^2\tau^2}$	$\frac{3 h_2}{8\pi^2\tau^3}$
$h_1.f^1$	$\frac{h_1[1.038 + 3\ln 2\pi\tau f_h]}{4\pi^2\tau^2}$	$\frac{h_1[24\ln 2 - 9\ln 3]}{8\pi^2\tau^2}$
$h_0.f^0$	$\frac{h_0}{2\tau}$	$\frac{h_0}{4\tau}$
$h_{-1}.f^{-1}$	$2\ln 2\, h_{-1}$	$\frac{h_{-1}[27\ln 3 - 32\ln 2]}{8}$
$h_{-2}.f^{-2}$	$\frac{2\pi^2 h_{-2}\tau}{3}$	$\frac{11\pi^2 h_{-2}\tau}{20}$

Table I : Responses of the Allan variance and of the modified Allan variance for different types of noises.

However, variances present several drawbacks :

- some variances depend on the high cut-off frequency for high frequency noises (Allan variance)
- some variances cannot distinguish between f^{+2} and f^{+1} frequency noises (τ^{-2} slope for both noises in the case of the Allan variance)
- some variances depend on the low cut-off frequency for very low frequency noises such as f^3 or f^4 frequency noises (Allan variance, modified Allan variance, ...)
- some variances cannot distinguish between a f^3 frequency noise and a linear frequency drift (τ^2 slope in both cases for the Allan variance and the modified Allan variance).

Nevertheless, a variance does not generally present all these drawbacks simultaneously. It is then possible to use a set of complementary variances.

3.2. Multi-variance method

The multi-variance analysis should be considered as an estimator of the PSD of y(t) rather than a classical variance measurement. Since the PSD of y(t) may be modeled as a sum of power laws (from f^{-2} to f^{+2} frequency noise), the goal of this method is to measure each noise level. The principle consists in using several variances over the same signal. These different variances must be chosen in order to get an accurate measurement for low frequency noises (Allan variance, Picinbono variance) as well as for high frequency noises (modified Allan variance, Time variance). Moreover, the drifts may also be measured by this method.
The relevant parameters (noise levels and drift coefficients) are determined by a weighted least square method : knowing the response of each variance for each type of noise (see Table I), and weighting each variance measurement by the inverse of its confidence interval, the noise levels and drift coefficients are calculated as the most probable parameters in the sense of the least squares.
This method is more sensitive than a single variance measurement since the accuracy of the measurement is at least as good as the most accurate variance of the set of variances. The choice of the variance thus determines the sensitivity of the multi-variance analysis : the best set of variances would be a set containing variances specially designed for each relevant parameter. But this choice may be adapted to each case (see section 4.2).
The second improvement of this method concerns the dynamics. The dynamics may be defined as the range in which a noise level or a drift coefficient can vary without swamping the other levels or without itself being swamped by the other levels. The dynamics depends on the number and on the type of the variances chosen, but is always greater than the dynamics of each variance of the set.
The main advantage of the multi-variance analysis is the determination of the uncertainty domain over each parameter measurement. The confidence intervals are estimated from the dispersion of the variance results by using a singular value decomposition. Moreover, since the assessment of the precision leads to the association of an eigen vector with each noise level or drift coefficient, it is possible to calculate the angle between each eigen vector. A small angle between two eigen vectors implies that the parameters corresponding to these vectors yield similar effects over the variance responses. In this connection, the separability may be define as the aptitude of a variance or a set of variances to distinguish between two noises. Thus, the determination of the angles between the eigen vectors corresponding to the parameters yields a quantitative estimation of the separability of the multi-variance analysis and provides a criterion to estimate the measurement reliability. Only the types of noises and drifts that actually exist in the signal are then measured.
For a full description and measurement results of the multi-variance analysis see [1] and [2].

4. DEFINITION OF SUITABLE VARIANCES USING THE STRUCTURE FUNCTION APPROACH

4.1. The structure functions

Thus, the choice of the set of variances is crucial. In this connection, the structure function approach [3] may be useful in determining the variance (i. e. the structure function) which is optimized for a type of noise and for an order of drift. The concept of the structure function is an extension of the variance approach to n^{th} difference operator [3, 4].

Figure 2 shows the measurement sequences of several structure functions, from first differences to third differences (from top to bottom) and belonging to the first two families. These measurement sequences are given for x(t) time error data (left) and for $\overline{y_k}$ instantaneous normalized frequency deviation samples (right).

The measurement sequence of the structure function corresponding to n^{th} difference and m^{th} family (that we will denotes in the following $\sigma_{n,m}^2(\tau)$) is obtained by applying the measurement sequence for $\overline{y_k}$ samples of $\sigma_{n+1,m-1}^2(\tau)$ over the x(t) data. The difference between the measurement sequence of $\sigma_{n+1,m-1}^2(\tau)$ for $\overline{y_k}$ samples and the measurement sequence of $\sigma_{n,m}^2(\tau)$ for x(t) data comes from the nature of the data : the x(t) data are instantaneous values whereas the $\overline{y_k}$ samples are averaged over a time interval τ (this integration time corresponds to the length of the steps of the measurement sequence applied to $\overline{y_k}$ and to the interval between two data in the measurement sequence applied to x(t)). Thus, it is possible to calculate, step by step from the first family, any structure function for any order of difference and for any family.

The number of alternations of the measurement sequences depends on the order of difference : the measurement sequence of $\sigma_{n,m}^2(\tau)$ for $\overline{y_k}$ samples exhibits n+1 alternations. This yields two important properties :

- a structure function with a high order of difference is insensitive to a high order of drift : $\sigma_{n,m}^2(\tau)$ is insensitive to frequency drifts up to (and including) $t^{(n-1)}$ (i. e. for phase drift up to t^n)
- a structure function with a high order of difference converges for very low frequency noises without taking into account the low cut-off frequency : $\sigma_{n,m}^2(\tau)$ converges for frequency noises up to (and including) $f^{(-2n)}$ (i. e. for phase noises up to $f^{(-2n-2)}$).

On the other hand, the order of family is linked to the convergence for high frequency noises. The higher the order of family is, the smoother the measurement sequence is. This means that $\sigma_{n,m}^2(\tau)$ converges for frequency noises up to (and including) $f^{(+2m-2)}$ (i. e. for phase noises up to $f^{(+2m-4)}$), without taking into account the high cut-off frequency (if the sampling rate satisfies the conditions of the sampling theorem).

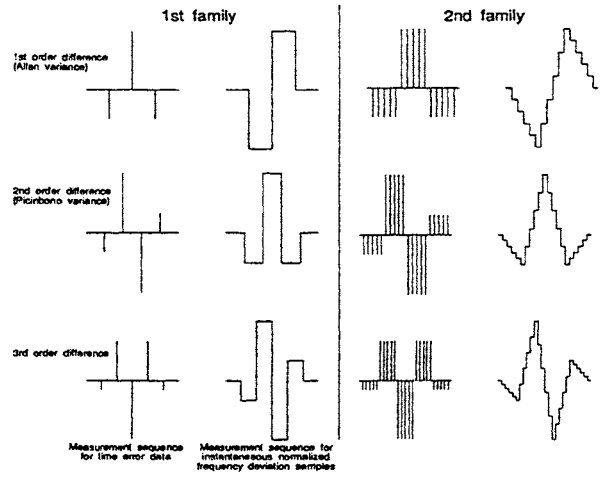


Figure 2 : The first and the second family of structure functions.

These properties may be summarized by :

- $\sigma_{n,m}^2(\tau)$ is insensitive to frequency drift up to $t^{(n-1)}$
- $\sigma_{n,m}^2(\tau)$ converges for low frequency noises up to $f^{(-2n)}$
- $\sigma_{n,m}^2(\tau)$ converges for high frequency noises up to $f^{(+2m-2)}$.

Moreover, since a high order of difference provides a measurement sequence with many alternations and a high order of family yields a very smooth measurement sequence, the use of structure functions may be very close to the wavelet analysis.

4.2. Choice of an optimized variance

We will illustrate this purpose with an example : we want to characterize the time instabilities of a millisecond pulsar.

The timing measurements should be affected by high frequency noises (poor short term stability) as well as linear frequency drift and, although millisecond pulsar exhibits a great long term stability (over time interval between 1 and 10 years), very low frequency noises (f^3 or f^4 frequency noises). Moreover, the low cut-off frequency may be estimated as the inverse of the time of flight of the light from the pulsar to the earth.

An ideal variance should be able to measure high frequency noises (f^{+2}) as well as very low frequency noises (f^4) and distinguish between very low frequency noises and linear frequency drifts. Unfortunately, a variance (structure function) which converges for very low frequency noises (at least 2^{nd} difference) is insensitive to drifts.

Hence, we must simultaneously use variances of first difference (such as the Allan variance $\sigma_y^2(\tau) = \sigma_{1,1}^2(\tau)$ or the modified Allan variance Mod $\sigma_y^2(\tau) = \sigma_{1,2}^2(\tau)$) in order to measure the linear frequency drift, and variances of second difference (such as the Picinbono variance $\sigma_{2,1}^2(\tau)$, or $\sigma_{2,2}^2(\tau)$) in order to measure the very low frequency noises.

On the other hand, in order to distinguish between f^{+1} and f^{+2} frequency noises, variances of the second family (such as modified Allan variance $\sigma_{1,2}^2(\tau)$ or $\sigma_{2,2}^2(\tau)$) must be used.

The variance of second difference and second family $\sigma_{2,2}^2(\tau)$ seems to be interesting for this purpose since it converges for f^{+2} to f^4 frequency noises (see Table II). But, this variance is insensitive to linear frequency drift.

Thus, the multi-variance analysis may be used with a set of 4 variances :

- the Allan variance $\sigma_{1,1}^2(\tau)$
- the Picinbono variance $\sigma_{2,1}^2(\tau)$
- the modified Allan variance $\sigma_{1,2}^2(\tau)$
- $\sigma_{2,2}^2(\tau)$.

$S_y(f)$	$\sigma_{2,2}^2(\tau)$
$h_{+2} \cdot f^{+2}$	$\frac{5\, h_{+2}}{2\, \pi^2\, \tau^3}$
$h_{+1} \cdot f^{+1}$	$\frac{18(2\ln 2 - \ln 9) h_{+1}}{\pi^2\, \tau^2}$
$h_0 \cdot f^0$	$\frac{4\, h_0}{3\, \tau}$
$h_{-1} \cdot f^{-1}$	$2(27\ln 3 - 40\ln 2) h_{-1}$
$h_{-2} \cdot f^{-2}$	$\frac{4\, \pi^2\, \tau\, h_{-2}}{3}$
$h_{-3} \cdot f^{-3}$	$\frac{4\, \pi^2\, \tau^2 (416\ln 2 - 243\ln 3)\, h_{-3}}{15}$
$h_{-4} \cdot f^{-4}$	$\frac{1208\, \pi^4\, \tau^3\, h_{-4}}{315}$

Table II : Response of $\sigma_{2,2}^2(\tau)$ (2^{nd} order difference, 2^{nd} family) for the different types of noises.

$(\sigma_{1,2}^2(\tau)$; $\sigma_{2,2}^2(\tau))$ are used for measuring high frequency noises, $(\sigma_{2,1}^2(\tau)$; $\sigma_{2,2}^2(\tau))$ are used for measuring low frequency noises, $(\sigma_{1,2}^2(\tau)$; $\sigma_{2,2}^2(\tau))$ are used for measuring the linear frequency drift. Then, if the linear frequency drift is smaller than the limit given in (5), it can be neglected.

5. CONCLUSION

The multi-variance analysis may be an accurate tool if the variances are adapted to each problem. The only restriction concerns the irregularly spaced timing data (millisecond pulsars, GPS data, etc...). However, the lowest mode estimators [5, 6] may be considered as an extension of the structure functions in the case of irregularly spaced timing data. In a next paper, we will develop an adaptation of the multi-variance analysis to the case of irregularly spaced timing data by using several lowest mode estimators.

6. REFERENCES

[1] F. Vernotte, E. Lantz, J. Groslambert, J. J. Gagnepain, "Oscillator noise analysis : multi-variance measurement", *IEEE Trans. on Instrum. and Meas.*, vol. IM-42, no. 2, April 1993, pp. 342-350.

[2] T. Walter, "A multi-variance analysis in the time domain", *Proc. of 24th PTTI*, 1992.

[3] W. C. Lindsey, C. M. Chie, "Theory of oscillator instability based upon structure functions", *Proceedings of the IEEE*, vol. 64, December 1976, pp. 1652-1666.

[4] J. Rutman, "Characterization of phase and frequency instabilities in precision frequency sources : fifteen years of progress", *Proceedings of the IEEE*, vol. 66, no. 9, September 1978, pp. 1048-1075.

[5] J. E. Deeter, P. E. Boynton, "Techniques for the estimation of red power spectra. I. Context and methodology", *The Astrophysical Journal*, vol. 261, October 1982, pp. 337-350.

[6] J. E. Deeter, "Techniques for the estimation of red power spectra. II. Evaluation of alternative methods", *The Astrophysical Journal*, vol. 281, June 1984, pp. 482-491.

IS THE FREQUENCY NOISE OF AN OSCILLATOR OF DETERMINISTIC ORIGIN ?

M. Planat, V. Giordano, G. Marianneau, F. Vernotte*, M. Mourey,C. Eckert*** and J. Miehe*****

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
* Observatoire de Besançon
** LCEP-ENSMM - Besançon
*** Groupe d'Optique Appliquée - Centre de Recherches Nucléaires - Strasbourg

ABSTRACT :

We report on a new multifractal type approach of low frequency noise in oscillators. The time sequence is interrogated by means of a multiscale local stability exponent which acts as a sliding filtering window. Its discontinuities are reflected into a binary coding of which the correlations are analysed. The mapping of binary data into a devil's staircase is used to emphasize the correlated origin of frequency fluctuations in quartz oscillators.

1 - INTRODUCTION

After decades of efforts the low frequency noise of an oscillator is still unsatisfactory understood [1,2]. A widely used measure of the random frequency deviation of an oscillator with respect to a reference clock is the difference $\Delta y(\tau) = |y_2(\tau) - y_1(\tau)|$ between adjacent relative frequency measurements, each sampled over a time interval τ. The so-called Allan variance follows from the mean squared value $\sigma^2_y(\tau) = (1/2) < (\Delta y(\tau))^2 >$. The graph of the variance versus the sampling time obeys power laws τ^m (m integer) over a restricted range of time scales and always takes the form of a valley. A typical oscillator has $m > 0$ (stability) over short times, $m = 0$ (flicker noise floor) at intermediate times and $m < 0$ (instability) at large times. If one applies the laws of statistical physics to the modeling of oscillator noise one is in a position to only account for the stable part : for instance : the thermal noise in a quartz oscillator or a maser leads to the exponent $m = 2$ and the shot noise in a cesium clock leads to the exponent $m = 1$. The origin of the floor and drift is still unexplained.

A dual measure of frequency noise is the power law which is followed by the power spectral density of frequency fluctuations $S_y(f) \sim f^p$ (p integer). Fortunately, in the hypothesis of stationarity, the time and frequency domains are related with a simple law $m = p + 1$ (if $-3 < p \leq 1$) and $m = 2$ (if $p \geq 1$). Therefore shot noise in an atomic clock leads to white frequency noise : $p = 0$. The flicker floor : $m = 0$ corresponds to the scaling $S_y(f) \sim 1/f$ in the power spectrum. This low frequency noise is present in a large number of natural processes (solid state physics, biologic, geographic, economic and even artistic data). It has been postulated that a self-organizing principle is brought into operation [3].

Two problems have been left aside with this simple approach. First, there is an ambiguity at $m = 2$ on the type of noise in frequency domain. More dramatically at $m < 0$ the frequency deviation diverges in time. These problems are conventionaly removed by averaging over N samples so that the effective bandwith of the measurement decreases from $\Delta f = 1/\tau$ down to $1/N\tau$ and better estimation of frequency deviations may be performed [2].

2 - BASIC MULTISCALE MULTIFRACTALS CONCEPTS

We have developed an alternative characterization of the frequency noise of an oscillator which is in the spirit of contemporary work on multifractals [4, 5,6]. The local stability has been measured from a scaling exponent $\beta^{(N)}$ defined by

$$\Delta^{(N)} y(\tau) = (N\tau)^{-\beta^{(N)}} \qquad (1)$$

where the frequency deviation is expressed as

$$\Delta^{(N)} y(\tau) = (1/N) \left| \sum_{i=N+1}^{2N} y_i - \sum_{i=1}^{N} y_i \right| \qquad (2)$$

The above multiscale deviation contrasts with the standard averaging over frequency deviations so that the discontinuities of the time series may be smoothed out by increasing the number of samples.

The multifractal dimension $g(\beta)$ is calculed from the equation :

$$g(\beta^{(N)}) = log\,(n(\beta^{(N)}))/\,log\,(N\tau) \qquad (3)$$

in order to tranfer the scaling effect from the local exponent to the probability density $n(\beta)$ [4].

The method was used to characterize a 10 MHz quartz oscillator with crystal quality factor $Q \sim 10^6$. Oscillations were sustained from a Colpitts type set-up. Temperature fluctuations were controled down to 10^{-2} C over time scales ranging from 1 to 1 000 s. The frequency was recorded from the 500Hz beat of the oscillator with respect to a reference oscillator. When plotted versus the sampling time the standard frequency deviation showed the conventional valley like graph with a flicker floor at $\sigma_y(\tau) \sim 10^{-9}$.

In Fig. 1 are shown 4 500 consecutive data of the local stability exponent $\beta^{(1)}$ with the sampling time $\tau = 1$s. The inset shows the corresponding multifractal density $g(\beta^{(1)})$.

The curve grows steeply up to a maximum value at $\beta = \beta_{max}$ and then decreases slowly approximately as $1/\beta^{(1)}$. This kind of signature is typical of all oscillators whatever their origin and Q-factor : they present a widerange of local stabilities extending from $\beta \sim 1$ up to high β values.

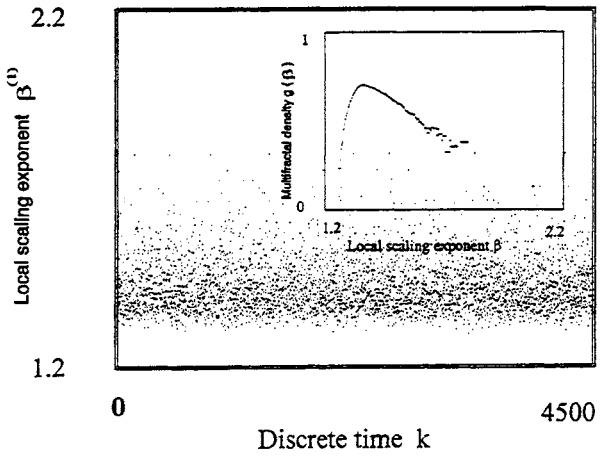


Fig. 1. *Time series of the local stability exponent for a Colpitts type quartz oscillator. The sliding window has unit width (N = 1) and the sampling time is τ = 1 sec. The inset shows the corresponding multifractal density.*

The frequency deviation and the local scaling exponent resulting from the averaging over N = 250 samples and the step by step sliding of the the time window are shown on Fig. 2. We observe the smoothing of the curves mainly in the poorly stable part at $\beta < \beta_{max}$. Conversely the stable part at $\beta > \beta_{max}$ remains very discontinuous.

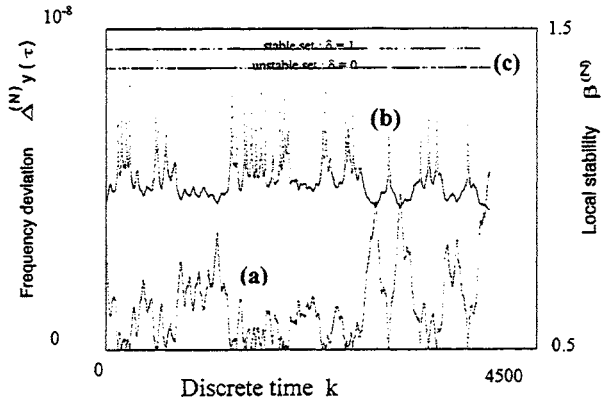


Fig. 2. *The smoothing of time series resulting from a wide sliding window (N = 250)*
(a): Frequency deviation
(b): Local scaling exponent
(c): Binary sequence

Similar plots were drawn versus the window size N at fixed time i = N (fig. 3). It is observed that the frequency deviation (curve a) quickly converts into a complex oscillating regime until at $N \sim 10^3$ divergence occurs. Correspondingly the number of stability bursts (curve b) decreases versus the window size until a constant value $\beta \sim 0.88$ is reached. This new stability measure contrasts with the smoothly diverging Allan variance (curve c).

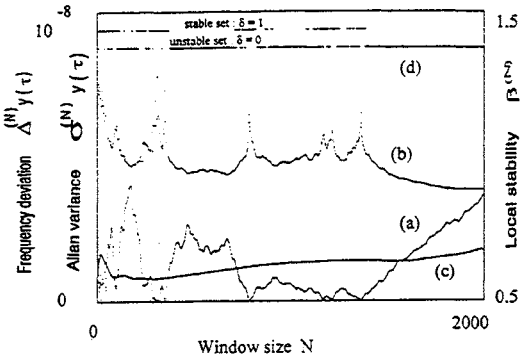


Fig. 3. *The smoothing of time series versus the window size N*
(a) Frequency deviation
(b) Local scaling exponent
(c) Allan variance
(d) Binary sequence

3 - BINARY CODING AND CORRELATION ANALYSIS

More structure has been drawn out by introducing a measure $\delta^{(N)}$ of the discontinuities which are suffered by the local exponent $\beta_k^{(N)}$ from the relation :

$$\delta^{(N)} = |1 - \beta_{k+1}^{(N)} / \beta_k^{(N)}| \qquad (4)$$

It is observed that if $\beta_{n+1} \sim \beta_n$ the discontinuity is smoothed out and the fractal deviation $\delta^{(N)}$ approximately vanishes ($\delta \sim 0$). This happens mainly at the points located at the left hand side of the peak β_{max} in the multifractal set. At the opposite discontinuity of two consecutive data implies $\delta > \delta o > 0$ where δo is significantly larger than 0. This happens mainly at the points located at $\beta > \beta_{max}$ in the multifractal set. These data can be translated into a binary sequence by coding the former points as $\delta = 0$ (unstable set) and the latter points as $\delta = 1$ (stable set).

The problem has turned into that of analysing the ergodicity of a binary sequence. The time correlations in this sequence have been found as follows. To each step k the subsequent bits have been used to attach the real number δ_s with the same binary expansion. Secondly the bits have been mapped into a real number δ_c from the continuous fraction :

$$\delta_c = n_o + 1/(n_1 + 1/(n_2 + 1/(n_3 + ...))) \qquad (5)$$

where the n_i means $n_o = 0$, n_1 -1 times 0, followed by n_2 times 1, n_3 times 0 and so on. The randonness of a binary sequence can be appreciated from the properties of the resulting graph δ_c versus δ_s.

It has been demonstrated that if δ_s is choosen continuous the graph δ_c (δ_s) is also continuous. It is a monotone increasing curve : the so-called slippery devil's staircase. Step are found each time δ_s gets close to a rational number when there is a large integer n_i in the continuous fraction [6]. In between the steps a significative slope in the staircase means small integers n_i in the continuous fraction.

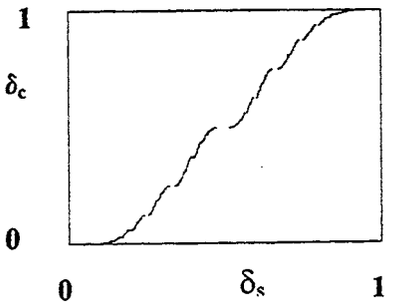


Fig. 4. *The mapping of an uncorrelated binary sequence into a devil's staircase with unoccupied steps. δ_s is the real number with the same binary development and δ_c results from the continuous fraction expansion*

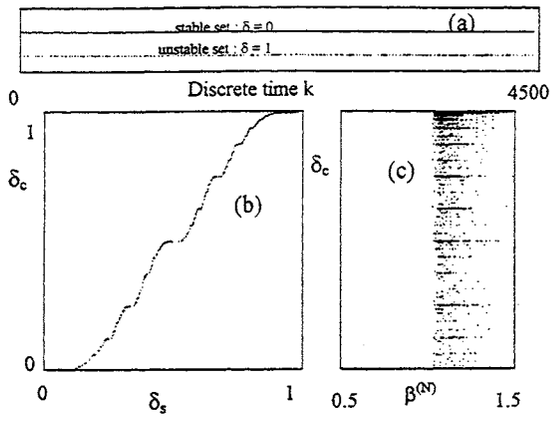


Fig. 5.
(a) Binary coding of the oscillator time series by using the fractal deviation exponent δ (N = 25)
(b) The mapping of the oscillator time series into the devil's staircase.
The filling of resonant steps attached to $\delta_s = p/q$ with $\delta_c = 1 - p/2^{q-1}$
(p and q integers) emphasizes the correlated origin of frequency fluctuations
(c) It is shown that the filling of steps arises from a wide range of scaling exponents $\beta^{(N)}$

A binary uncorrelated sequence shows a staircase with gaps at the steps locations (Fig. 4). Conversely, Fig. 5 (N = 25 samples) gives credit to a correlated interpretation of low frequency noise. The noise data have been split into the disjoint stable and unstable sets by using the fractal deviation δ as explained above. Small averages provide a partition between a full measure stable set and a nowhere dense unstable set (Fig. 2c, Fig. 3d and Fig. 5a). This corresponds to the filling of steps attached at rational numbers $\delta_s = p/q$ and $\delta_c = 1 - p/2^{q-1}$ (with p and q integers, $p \leq q - 1$) in the staircase, with p even for the most populated ones. Each occupied step is found to arise from a wide range of stability exponents (Fig. 5b and 5c). By increasing N more points are transfered to the unstable set and the filling of steps in the lower part of the staircase (p odd) increases. The above low frequency data supply in a way the complement to the random high frequency data shown in Fig. 4.

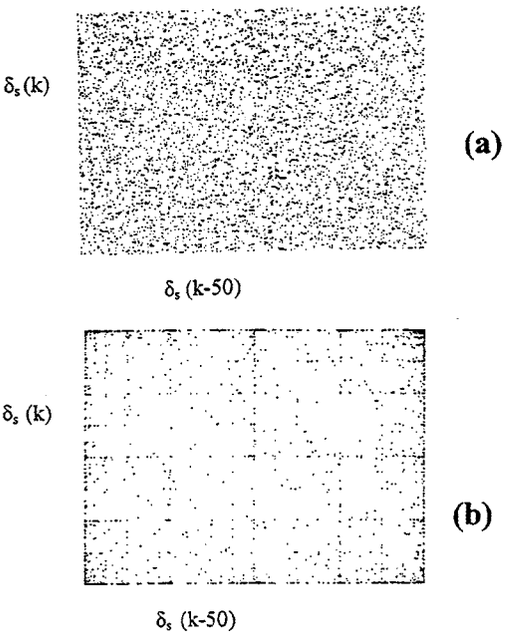


Fig. 6. *A two dimensional plot δ_s (k) versus δ_s (k - 50) of ergodic properties in a time series*
(a) The binary random sequence of Fig. 4
(b) The binary correlated sequence of Fig. 2

Long range correlations can also be detected by plotting δ_s and δ_c at remote times. For a random binary sequence the two dimensional resulting distribution is uniform (Fig. 6a) while for the binary sequence arising from the quartz oscillator data most points are located at the even resonances.

4 - CONCLUSION

In summary a three steps procedure : time series → local scaling exponents → binary coding → devil's staircase has been implemented to derive correlations in the low frequency spectrum of a Colpitts type quartz oscillator. Similar results were also observed in quartz and cesium oscillators.

Future work should focus on a letter understanding of the filling of levels in the staircase. The above stable-unstable dynamics may be used as an accurate sensor for temperature, power, magnetic, or accelerometric perturbations or may help designers to improve the performance of these oscillators.

REFERENCES

[1] D.W. Allan, IEEE Trans. on Ultras. Ferroelect. and Freq. Control, UFFC-34, 647 (1987)

[2] J. Vanier and C. Audoin, The Quantum Physics of Atomic Frequency Standards. (Adam Hilger, Bristol and Philadelphia, 1989)

[3] P. Bak, C. Tang and K. Wiesenfeld, Phys. Rev. A. 38, 364 (1988)

[4] B.B. Mandelbrot. The Fractal Geometry of Nature. (Freeman, New York, 1982)

[5] M. Planat and P. Koch, Fractals 1, 727 (1993) ; and EFTF Proceedings : 1992, p. 333 ; 1993, p. 349 ; 1994, p. 821.

[6] M.C. Gutzwiller and B.B. Mandelbrot, Phys. Rev. Lett., 60, 673 (1988)

[7] M. Planat, The discontinuities of metrological time : Poincare's heritage ; to appear in Chaos, Solitons and Fractals.

A miniaturized Rubidium Physics Package of Atomic Clock for Space, Air and Industrial applications

A.Gevorkyan, V.Khutorschikov, G.Smirnova

Russian Institute of Radionavigation and Time

Abstract

Problem of improving the stability of atomic gas-cell frequency standards and of their miniaturization is analyzed based upon the theory of optical pumping and experimental investigations. It is carried out an analysis of tendencies of changes of parameters for devices produced in Russia and other countries. It is discussed earlist forecast and made new forecast of estimated parameters for 2000 year.It is shown that in spite of the advance made in that direction in the last few years there are considerable reserves to improve the performance of the instruments of this class. The miniaturized rubidium physics package and atomic gas-cell frequency standard with it based upon the ultra miniaturized rubidium cells is described.

- Atomic frequency standard, physics package, miniaturized cell

Predicting standard accuracy parameters and its analysis

Rubidium-cell frequency standards are of wide use as reference oscillators in radionavigation and doppler-systems, digital/cellular communication systems and other systems as well.

However, over last years, alongside with development of conventional atomic frequency standards research of new stabilization frequency methods is actively and successfully being conducted. That is why, the utility of expensive work on improvement of rubidium atomic clock has became disputable. Therefore, authors carried out an analysis of the tendencies of change of parameter for devices produced in USSR (Russia) and other countries [1]. This analysis has shown a steady growth of its operational and accuracy characteristics and has allowed to make the forecast of estimated parameters for 1995-2000 years. An expected characteristics (instability of frequency - $(1.5\text{-}20)\cdot10^{-13}\tau^{-1/2}$, volume - (0.3-5) dm^3, life time - 50000-90000 hours) in the lately forecast were clarified (fig.1,2). The rate of change instability of frequency was reduced whereas the tendency for a decrease of volume of the atomic frequency standards was conserved.

But in order to prove the validity of extrapolations for long-term prospect, it should be convinced that the investigated characteristics are yet far from limit values to be achieved. For their account it should be mentioned, that the dispersion of frequency of a crystal oscillator σ^2, stabilizied on frequency of atomic transition, is stipulated by parameters of radio-optical resonance and conditions of its supervision - σ^2_1, and instability of a resonant transition under action of external perturbations - σ^2_2:

$$\sigma^2 = \sigma^2_1 + \sigma^2_2 . \qquad (1)$$

The accuracy of set-up of a crystal oscillator under frequency of atomic transition depends on parameters of radiooptical resonance (slope of the discriminator S_d, spectral density of noise on an output of the discriminator g_d and on parameters of a system of a lock of frequency) [2,3]:

$$\sigma^2_1 = \frac{g_d}{2\,T\,N^2\,S^2_d} + \sigma^2_\omega \frac{\Omega_0 T}{\Omega_0\,T + N^2\,S^2_d\,S^2_y} , \qquad (2)$$

where T is a constant system of a loop; S_y - slope of a managing element; σ^2_ω - dispersion of frequency crystal oscilliator, N - frequency of multiplication; the relation $g^{1/2}_d/S_d$ can be named as a threshold of sensitivity of a radio-optical resonance [2]. Instability of frequency responsible due the external perturbations of resonant transition, the fluctuations of intensity of light and light shifts, the fluctuations of temperature of environment and temperature factor of the physics package, mechanical effects and etc. in conditions of the statistical independence of perturbations:

$$\sigma^2_2 = \frac{1}{N^2}\Sigma\sigma^2_i = \frac{1}{N^2}\sum_i \int_0^\infty g_i(\omega)\,\left|k^2(i\omega)\right| \frac{\sin^2\omega\pi}{(\omega\pi)^2}\,d\omega,$$

Where $g_i(\omega)$ is spectral density of i-th perturbation, $k^2(i\omega)$ - factor of transfer of the physics package of i-th perturbation on frequency of resonant transition. At time of measurement less than 10 s in (1) first member prevails, at time more than $3\cdot10^5$ s - second member.

Short-term stability

In case, when the instability of the device is determined by a threshold of sensitivity of physics package (first member in (2)), one can evaluate the limit threshold of sensitivity amplitude of the fist harmonic of a radio-optical resonance (3) by equating to a square root of spectral density of noise on its output. And the amplitude of the fist harmonic can be calculated as a result of Fourier transformation of an exact contour of a absorption line, received on the basis of the quantum theory of radiation for the case of thin cell [3]:

$$I_1(\delta) = \frac{4}{27} G_1 \frac{1}{W+\Gamma} , \qquad (3)$$

W - speed of optical pumping; Γ - "dark" width of a spectral line; G_1 -coefficient depending on photocurrent, parameters of radio-optical resonanse,spin of a nucleus - J . Spectral density of noise in centre of a line can be submitted [4]:

$$g(\Omega,0) = g_L + g_{RF} + S^2 k_\nu(\Omega/2)\, g_\nu(\Omega/2) + \\ + S^2 k_\nu(2n\Omega)\, g_\nu(2n\Omega), \qquad (4)$$

where g_L, g_{RF}, g_ν are spectral densities respectively of peak fluctuations of light, of RF field and phase fluctuations of RF field on frequency Ω. The third member corresponds to transformation of phase fluctuations RF field from area Ω/2 in peak fluctuations of light on frequency Ω, fourth - similar process from areas corresponding to even harmonics of freguency of modulation [4]. In limiting case of prevalence of shot noise of a source of light and assumption, that light of pumping does not contain nonresonant lines threshold of the sensitivity for the case of thin cell [3,1]

$$\frac{\delta}{f_0} = \frac{27}{4} \frac{[J(J+1)W + (2J+1)\Gamma]^2}{C_1 W^{3/2} k l} \tau^{-1/2},$$

where k is coefficient of absorption; l - length of a cell; C_1 -constant; τ - time of measurement.

In [5] the limit threshold of sensitivity for "thick" cell is considerated. The threshold of sensitivity depending on radius of a cell for opticaly "thin" and "thick" cell is indicated on fig. 3. When the radius of a cell is equal to 1.6 sm the limiting sensitivity for opticaly "thin" cell will be $3\cdot10^{-13}\tau^{-1/2}$ [3], and for opticaly "thick" cell- $5\cdot10^{-14}\tau^{-1/2}$ [5]. In [6] for case of optical pumping by the laser the limiting sensitivity was appreciated by value $5\cdot10^{-15}\tau^{-1/2}$ (that close to results [5], its recalculation on case of the laser gives limiting sensitivity less than $1\cdot10^{-14}\tau^{-1/2}$). Thus, the limit threshold of sensitivity of radio-optical resonance for rubidium cell standard is $(0.5\text{-}5)10^{-14}\tau^{-1/2}$. On the other hand, from fig.3 it is visible, that even at the smaller sizes of a cell it is possible to receive rather low the threshold of the sensitivity of the physics package. So,if the radius of a cell is equal 0.65 sm (V=1 sm^3) the threshold of sensitivity equals $4\cdot10^{-12}\tau^{-1/2}$, that exceeds of the conventional sensitivity of the cells of volume 30-100 sm^3. By use of cells of volume 0.12 sm^3 (R=0.3 sm) the threshold of the sensitivity can be evaluated by value $(0.15\text{-}1.5)\cdot10^{-11}\tau^{-1/2}$.

Long-term instability

Researches of fluctuations of the frequency depending on the different factors (intensity of light of pumping, pressure, the input voltage, the temperature, etc.) have shown, that the determining factories at the time > 10^5 s are the fluctuations of the temperature and the intensity of light of pumping.

Fluctuations of sources of light can be approximate as it is accepted at the analysis of frequent fluctuations by decomposition in a series:

$$g_L(\omega) = \sum_{i=-\infty}^{\infty} h_i \omega^i. \qquad (5)$$

Being limited the members, appropriate to the white noise h_0, flicker noise h_{-1}, and random walk h_{-2} in temporary area we have

$$\sigma^2_L(\tau) = (2\pi)^2 h_{-2}\tau/6 + 2\ln 2 h_{-1} + h_0/2\tau,$$

where h_{-2}, h_{-1}, h_0 are coefficiens from (5).

On fig. 4 experimental dependencies of fluctuations of light on time of measurement are indicated. The most important established fact is dependence of fluctuations type on a mode of operation of source.

Important restriction of sensitivity can be a fluctuations of RF field (3 and 4 members in (4)). These fluctuation will be discriminator transformed proportionally to a slope of the discriminator curve [4], and it completely change the approach to choice of mode of operation of the physics package. Namely, the increase of a slope makes sense up to a moment, at which the peak fluctuations of light, stipulated by transformation of frequency-phase fluctuations on a spectral line, became equal to shot noise of a photon current. At achievement of this condition the threshold of sensitivity, so and stability of the device don't depend on a slope and is completely determined by the fluctuation characteristics of the RF field. Thus, in considered case for achievement probably smaller threshold of the sensitivity it should be aimed to decrease a level of fluctuations of RF field, i.e. crystal oscillator, multiplier of frequency, synthesizer of frequency. These fluctuations due to light shifts influences also the long term stability. And at given parameters of RF field is appropriate to reduce the dimensions of the physics package and reduce the intensity of light in comparison with optimum intensity.

These results alongside with experimental research of a radio-optical resonance in small-sized cells of volume 1 and 0.5 sm^3 [3,4,5] have resulted in development of the small-sized physics package (fig.5), having following values of characteristics:

- Sensitivity — $3\cdot10^{-11}\tau^{-1/2}$
- Operating Temperature, °C — -60 - +75
- Temperature Coefficient, 1/°C — $5\cdot10^{-13}$
- Aging monthly — $8\cdot10^{-12}$
- Warm-up time to $1\cdot10^{-10}$, min
  at 25 °C — 5
  at -60 °C — 15
- Volume, sm^3 — 140
- Life time, hour — 55000

This physics package is already production type.

At present, there is a design version of physics package with volume less than 40 cm^3 ensuring accuracy characteristic. While using any of these physics package the volume of the frequency standard is defined by electronics and can be reduced up to the value less than 200 cm^3 with maintaining high accuracy characteristics.

References

1. A. Gevorkyan, V .Khutorschikov, G. Smirnova. Proc. of All-Union meetings "Quantum metrology and fundamental physical constants " -Leningrad. 1985 - P.27.

2. V. Kushnir. Quantum standards of frequency. P.2. - Leningrad. 1991.

3. G. Smirnova. Thesis. Lenindrad. 1985.

4. V. Khutorschikov. Radio-optical method of measurement of pressure of gases. - Moskow, 1992.

5. M. Gornyi and other. Zurnal technicheckoi fisiki. 1987. V.57, v.4, p.740-746.

6. Camparo M.C., Frueholz R.P. 40th Ann. Frequency Control Symp. - 1986. - P.447-449.

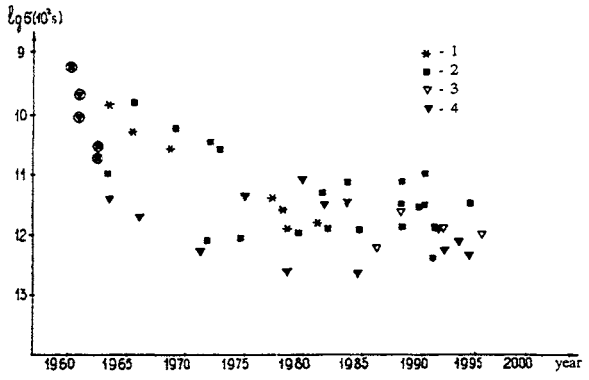


Figure 1. Short-term instability of rubidium standards mode by the various firms of the world:
1 - beginning of research;
2 - document's data;
3 - prior information;
4 - experimental data; o - $\tau = 1$s.

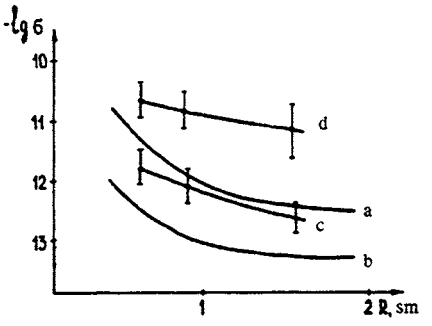


Figure 3. The threshold of sensitivity:
calculated: a - for "thin" cell;
b - for "thick" cell;
experimental: c - the highest;
d - for production types.

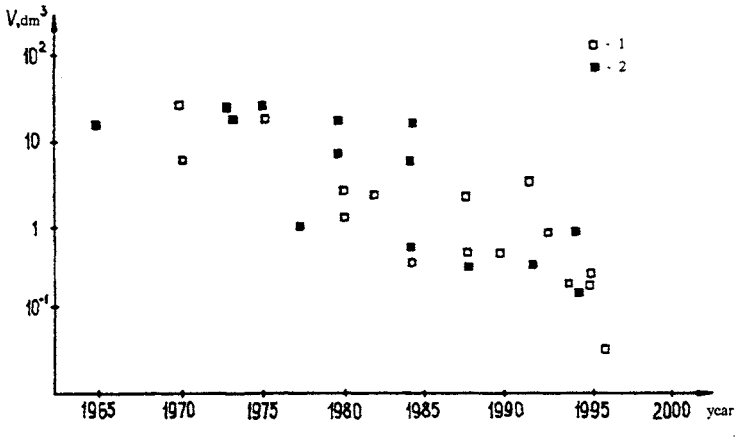


Figure 2. The dimentions of rubidium standards mode by the various firms of the world:
1 - prior information;
2 - document's data.

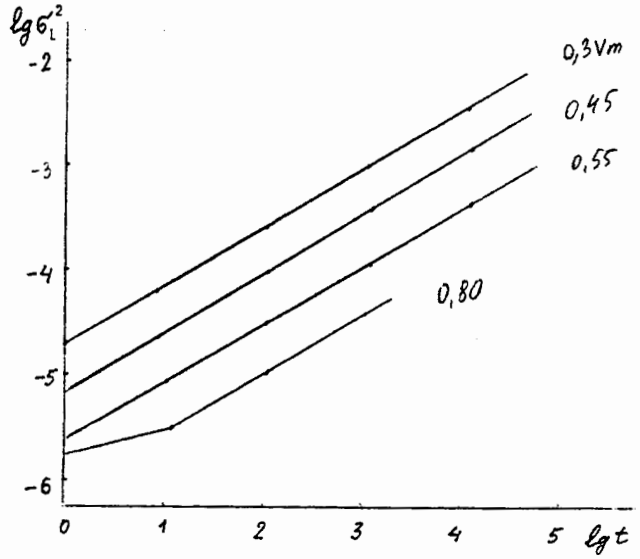


Figure 4. Fluctuations of light depending on time and operation mode.

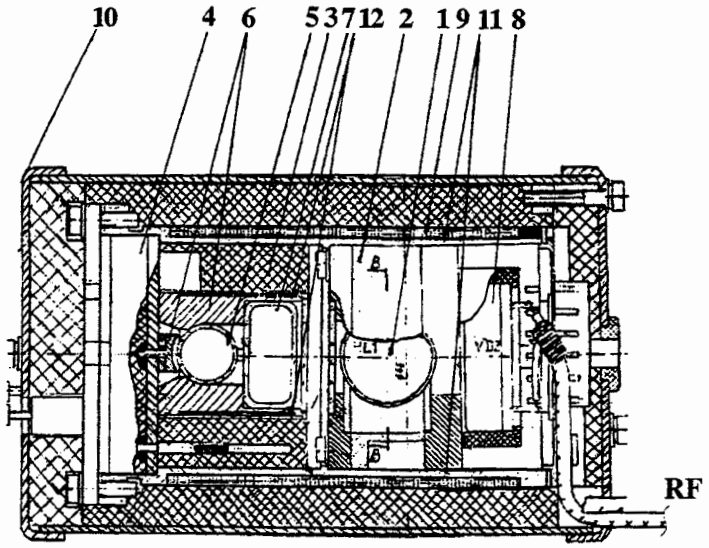


Figure 5. Physicks package:
1 - gas-cell; 2 - cavity; 3 - source of light; 4 - hf-generator;
5 - lamp; 6 - condensator; 7 - filter; 8 - photo-reciver; 9 - coil;
10 - magnetic screen; 11 - gas-cell oven; 12 - lamp oven.

Operation of optically pumped clock with high spectral purity laser diodes

G. Sélo, B. Boussert, P. Cérez, and N. Dimarcq

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

Introduction.

The short-term frequency stability of an optically pumped caesium beam resonator is directly related to the fluorescence signal-to-noise ratio (S/N) in 1 Hz bandwidth in the detection region. The latter depends in particular on the optical configuration for the state preparation and the detection of the fluorescence signal, and on the spectral purity of the excitation laser. Improvement of these two factors must greatly increase the stability of an atomic resonator.

On one hand, in order to obtain a clock signal as good as possible, pumping and detection processes must be optimised, by exciting the best transitions in both cases. Previously, two laser diodes were used and tuned to the required frequencies. Our experiment operates with a single laser which emission frequency is shifted by an acousto-optic modulator between the pumping and detection regions.

On the other hand, the commercial Fabry-Pérot laser diodes linewidths are about 30 MHz, which is quite large compared to the caesium natural linewidth of 5 MHz. This poor spectral purity greatly limits the S/N ratio achievable with such excitation lasers. Narrower linewidths are currently available with DBR diodes, of about 3 MHz linewidth. In order to obtain linewidths lower than 100 kHz, external cavity laser structures are required. We have made two different ones: a selfaligned extended cavity laser and a DBR diode in presence of weak optical feedback.

I. Narrow linewidth laser structure.

I.1. Self-aligned extended cavity laser.(SAECL)

The large linewidth of semiconductor laser diodes is due to the shortness of the diode cavity, since $\Delta\nu$ is inversely proportional to the square of the cavity length. A well-known solution to reduce the laser linewidth is to extend the cavity with a passive section. An usual configuration is presented on figure 1[1].

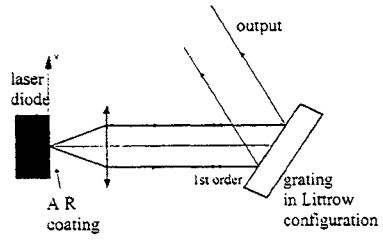


Fig.1: Extended cavity laser with selective frequency feedback

The external cavity, few centimetres long, is closed by a grating in Littrow configuration, which selects the frequency of the beam reflected inside the diode. Linewidths of the order of 30 kHz have been obtained with 10 cm long extended cavity laser (ECL)[1].

The main disadvantage of this structure is its critical sensitivity to the adjustment of the reflected beam perpendicular to the active layer. Any mechanical perturbation must be strictly avoided to maintain the laser running in extended cavity. Self-alignment techniques are then very useful to obtain a long term stability of the ECLs [2]. Our ECL is self-aligned with the help of an unidimensional cat's eye (Fig.2), which returns in the diode the whole beam reflected by the grating at a specific wavelength.

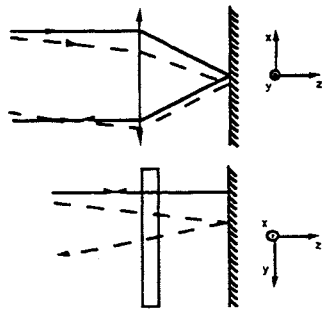


Fig.2: Principle of a 1D cat'eye

Indeed, in the (xOz) plane, the cat's eye returns a beam parallel to the incident one, whatever the direction of the latter; on the contrary, the grating diffracted lines in the (yOz) plan are not returned inside the diode. Many configurations of SAECL exist; one example is presented on figure 3:

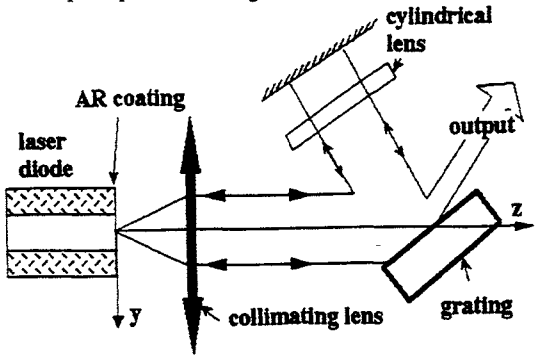


Fig.3: Example of a SAECL

SAECLs have been experimentally tested; their low sensitivity to vertical misalignment has been verified.

These lasers join a high spectral purity and a very good mechanical stability: no settings are needed for months.

I.2. DBR diode in presence of weak optical feedback.

Weak optical feedback is known to modify appreciably the characteristics of a semi-conductor laser. In particular, it narrows or broadens the laser diode linewidth, depending both on the phase of the reflected light and on the strength of the feedback, according to the following expression [3]:

$$\Delta\upsilon = \frac{\Delta\upsilon_0}{[1 + C\cos(2\pi\upsilon\tau_{ext} + \tan^{-1}\alpha_H)]^2},$$

where :

$\Delta\upsilon_0$ is the laser diode spectral linewidth without feedback,

τ_{ext} is the round trip delay through the external cavity,

α_H is the phase-amplitude coupling factor,

C is a feedback parameter.

According to the above formula, the narrowest linewidth is achieved when the phase of the reflected light is equal to $-\tan^{-1}\alpha_H$:

$$\Delta\nu_{min} = \frac{\Delta\nu_0}{[1+C]^2}$$

This structure is an alternative to the current SAECL. The compound cavity is driven by the diode; its emission line is fixed by the integrated grating structure of the DBR diode.

We have tried to reduce the 3 MHz linewidth of a DBR diode to a few hundred kHz by a weak reflection on a anti reflection coated microscope lame, as presented on figure 4. The phase of the reflected light must be carefully controlled to achieve a monomode emission with an actual linewidth narrowing.

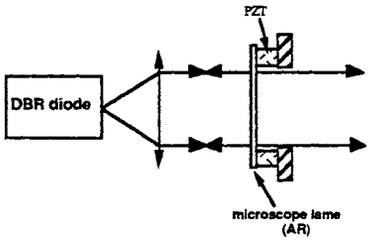


Fig. 4: DBR diode with weak optical feedback

Experiments are presently running to characterise the actual spectral performances of our laser. In order to improve it, we plan to use a 2D-self-alignment technology as shown on figure 5.

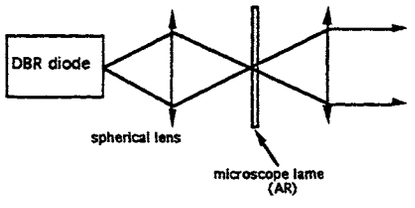


Fig. 5: Self-aligned DBR diode with weak optical feedback

II. Choice of the optima optical transitions.

II.1. State preparation.

In the first atoms-light interaction region, the laser beam creates a population inversion between the $|F = 3, m_F = 0\rangle$ and $|F = 4, m_F = 0\rangle$ hyperfine state levels of the fundamental $6^2S_{1/2}$ level. The highest population difference is obtained with the $4 \rightarrow 4'\sigma$ polarised transition of the caesium D_2 line which leads to $\Delta n = -15.6\%$ [4], where Δn is defined as :

$$\Delta n = \frac{n_{|F=4,m_F=0\rangle} - n_{|F=3,m_F=0\rangle}}{n}.$$

The laser spectral linewidth does not affect the fluorescence signal in the detection region as long as the pumping is completely achieved in the first region. Nevertheless, it may be uncomplete if the laser intensity is not high enough to compensate a lack of spectral purity. Then, the unpumped atoms propagate the laser frequency fluctuations in the detection region, strongly decreasing the S/N ratio [5].

II.2. Detection.

The highest fluorescence signal is obtained when the exciting light is tuned to the $4 \rightarrow 5'\sigma$ transition [6]. During its transit in the laser beam, each atom on the ($^2S_{1/2}$, F = 4) hyperfine state level emits a large number of fluorescence photons. The average number of emitted photons per atom depends on the intensity and spectral purity of the light, and on the interaction duration [5].

Unlike pumping transition, a cycling one is very sensitive to laser frequency noise. The fluorescence S/N ratio decreases with increasing laser linewidth : it is limited by the atomic shot-noise at narrow laser lines while at larger ones the frequency noise is the main limiting factor. Figure 6 shows that the highest fluorescence S/N ratio is achievable with laser linewidth less than 1 MHz.

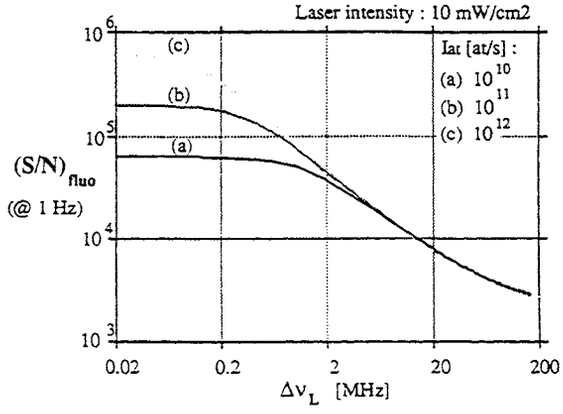


Fig. 6: Fluorescence S/N ratio vs the laser spectral linewidth [5]

It is noteworthy that the fluorescence S/N ratio is related to the number of emitted photons in the detection region ; it is distinct from the clock S/N ratio which refers to the photo detector current.

Nethertheless, in the case of cycling transition, the detection noise contribution is negligible compared to the others noise sources and does not degrade considerably the S/N ratio. The clock signal-to-noise ratio, which occurs in the expression of the atomic resonator short-term frequency stability, is then of the same order of magnitude than the fluorescence signal-to-noise ratio [5].

III. Experimental set-up.

As explained before, pumping and detection beams are derived from a single high spectral purity laser. The light frequency is shifted by 250 MHz between the two interaction regions so as to be tuned on the optima transitions - $4 \rightarrow 4'\sigma$ and $4 \rightarrow 5'\sigma$ respectively – for both pumping and detection processes (see figure 7).

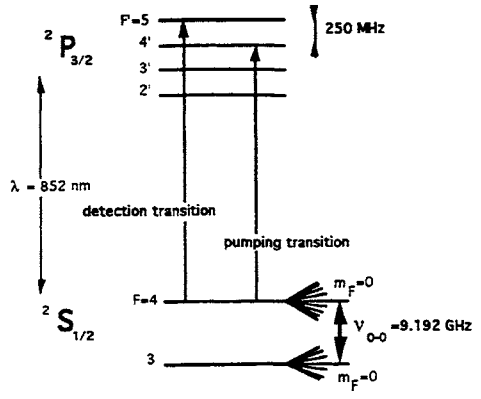


Fig.7: The Caesium state levels

The experimental set-up is shown on figure 8. The high spectral purity laser may be either a SAECL or a DBR with feedback. The light is σ-polarised, i.e. perpendicular to the magnetic field in the light-atom interaction regions.

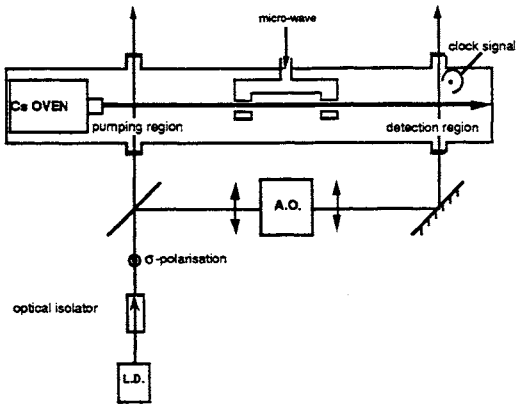


Fig.8: Experimental set-up

The short-term frequency stability of a resonator is inversely proportional to the clock (S/N) ratio. Our experimental configuration is expected to give a S/N ratio as good as 40 000 in 1 Hz bandwidth, leading to a short-term frequency stability of about 1.10^{-12} in one second.

Conclusion

The very good short-term frequency stability that might be achieved from our set-up results from the high spectral purity of the laser for both pumping and detection processes.

Indeed, the 4→ 5' transition is too sensitive to the excitation light frequency noise to give good results with laser diodes with linewidths greater than 1 MHz. That is why previous experiments were conducted with the 3 → 3'σ transition for both state preparation and detection, using a commercial laser diode of about 30 MHz linewidth. The detected signal-to-noise ratio was about 20 000 [7].

We hope to improve this (S/N) ratio by a factor two, and get a frequency stability of about 10^{-12}.

References

[1] M. de Labachelerie, C. Latrasse, P. Kemssu and P. Cérez, "The frequency control of laser diodes", *J. Phys. III France* **2** (1992) 1557-1589

[2] M. de Labachelerie, Dissertation, University of Paris-Sud, Orsay (1988)

[3] K. Petermann, Laser diode modulation and noise. Dordrecht, The Netherlands: Kluwer Academic, 1988

[4] P. Cérez, G. Théobald, V. Giordano, N. Dimarcq, and M. de Labachelerie, "Laser diode optically pumped cesium beam frequency standard investigations at LHA", *IEEE Transactions on Instrument and Measurement* **40**, 2 (April 1991) 137-141

[5] N. Dimarcq, V. Giordano, P. Cérez, "Statistical properties of laser-induced fluorescence signals", *Applied Physics B* **59** (1994) 135-145

[6] V. Giordano, Dissertation, University of Paris-Sud, Orsay (1987)

[7] V. Giordano, A. Hamel, P. Petit, G. Théobald, P. Cérez, and C. Audoin, "New design for an efficient optically pumped cesium beam tube", *43rd Annu. Frequency Control Symp.*, Denver, CO, May 1989, 130-131

Signal-to-Noise Considerations in Trapped-Atom Spectroscopy

V. Rieger, J.H. Müller*, D. Bettermann*, F. Ruschewitz, J.L. Peng, K. Sengstock, and W. Ertmer

Institut für Quantenoptik der Universität Hannover, Welfengarten 1, D-30167 Hannover, Germany
* Institut für Angewandte Physik der Universität Bonn, Wegelerstr. 8, D-53115 Bonn, Germany

Abstract

The achievable signal-to-noise ratio for spectroscopy on laser-trapped atoms is influenced by atomic number fluctuations. We present a theoretical model for these fluctuations and provide comparisons with experimental results. Theory and experiment can be used to determine important trap parameters in a non-destructive way. By incorporating an additional, spectroscopic loss-mechanism into the model, the signal-to-noise ratio for high-resolution spectroscopy on laser-trapped atoms can be calculated. In a pulsed experiment, different measurement strategies are shown to lead to different signal-to-noise ratios.

Introduction

Laser-trapped and cooled atoms promise to bring significant improvements in future optical atomic frequency standards. Extremely narrow optical resonances have already been observed with very high accuracy [1,2,3]. For the purpose of laser frequency stabilization onto such atomic lines the achievable signal-to-noise ratio (S/N) is a factor of crucial importance. Since the fluorescence from the almost stationary ensemble can be collected with high efficiency, S/N is mainly determined by the noise properties of the atoms, i.e., the atomic shot noise. The dynamics of the trapping process introduce long-term correlations to the fluctuations of the number of atoms. This leads to a non-white spectral density of the atomic shot noise. The correlations occur at time scales that are comparable to typical integration times. Therefore the exact shape of the noise spectral density has to be taken into account when considering the contribution of S/N to the achievable short-term frequency stability of trapped-atom frequency standards.

A simple model serves to calculate the spectral density of the atomic shot-noise fluctuations and allows comparisons with experimental data. The model elucidates the central role played by the lifetime of the trap for the noise behavior. In return, it is possible to use an analysis of experimental S/N to determine important trap parameters.

Understanding the influence of the trap dynamics on the spectroscopically observed signal points the way to optimizing the trap parameters for high-resolution spectroscopy and frequency stabilization. In a pulsed experiment it is possible to use a different measurement strategy to overcome the atomic shot-noise limitations.

Theoretical model of S/N

Both for applications in frequency metrology and for ease in comparisons with experimental data, it is advantageous to work with a signal-to-noise ratio in the time domain, dependent on the integration time τ. We consider a fluorescence count rate R(t) of the trapped atoms on the trapping transition. The number of photons detected at time t during the integration time τ is then given as

$$I(t,\tau) = \int_{t}^{t+\tau} R(t')\,dt' \quad . \tag{1}$$

In a straightforward manner we define the signal $S(\tau)$, the noise $\sigma(\tau)$ and the signal-to-noise-ratio $S/N(\tau)$:

$$S(\tau) := \langle I(t,\tau)\rangle_t \tag{2}$$

$$\sigma^2(\tau) := \langle \Delta I(t,\tau)^2\rangle_t = \langle (\int_{t}^{t+\tau} \Delta R(t')dt')^2\rangle_t \tag{3}$$

$$S/N(\tau) := S(\tau)/\sigma(\tau) \tag{4}$$

These definitions can be converted to the frequency domain by using the standard box function $h_\tau(t)$ of width τ and its Fourier transform $H_\tau(\omega)$ [4].

$$\sigma^2(\tau) = \langle (\int_{-\infty}^{+\infty} \Delta R(t')\, h_\tau(t-t')\,dt')^2 \rangle_t \tag{5}$$

$$\sigma^2(\tau) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} P_{\Delta R}(\omega)|H_\tau(\omega)|^2 d\omega \tag{6}$$

Thus the power spectral density $P_{\Delta R}(\omega)$ of the fluctuations of the fluorescence count rate is the important quantity in theoretically determining the signal-to-noise ratio.

Theoretical model of power spectral density

We consider, as in fig. 1, individual atoms entering the trapping region at random times t_j and being randomly ejected after a trapping time θ. Again using the box function $h_\theta(t)$, the number N(t) of trapped atoms at time t is given by:

$$N(t) = \sum_{\theta}\sum_{j=1}^{M_\theta} h_\theta(t-t_j) \quad . \tag{7}$$

At first only looking at the group of atoms (M_θ strong) that remain a given time θ within the trap, the power spectral density $P_{\Delta N\theta}$ is calculated using Carson's theorem [5,6]:

$$P_{\Delta N_\theta}(\omega) = R_{N_\theta}|H_\theta(\omega)|^2 \quad . \tag{8}$$

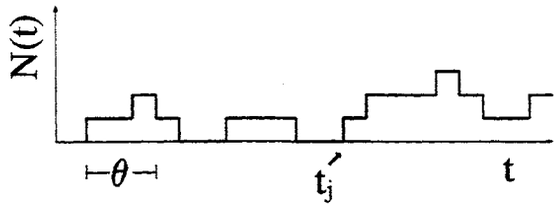


Fig. 1 Schematic representation of individual atoms entering and leaving the trapping region (see text)

Here $R_{N\theta}$ is the capture rate for such atoms. If we assume an exponential distribution $p(\theta)$ of trapping times θ with the mean value given by the trap life time θ_T, the total spectral density $P_{\Delta N}$ of the number fluctuations can be calculated [5,6]:

$$(9) \qquad P_{\Delta N}(\omega)=\int_0^\infty P_{\Delta N_\theta}(\omega)\,d\theta=R_N\int_0^\infty p(\theta)\,|H_\theta(\omega)|^2 d\theta \quad .$$

The result is a Lorentzian power spectrum:

$$(10) \qquad P_{\Delta N}(\omega) = 2R_N\theta_T \frac{1/\theta_T}{\omega^2+(1/\theta_T)^2} \quad ,$$

where R_N is the mean capture rate ($R_N=<N(t)>_t/\theta_T$). The Lorentzian line shape leads to deviations from a white-noise behavior at Fourier frequencies corresponding to the inverse of the trap lifetime θ_T. In the time domain this leads to a modification of the $\sqrt{\tau}$-behavior usually found for shot noise.

S/N of trap fluorescence:

The signal-to-noise ratio for fluorescence detection of the trapping radiation is determined from three major contributions [7]:

- atomic shot noise: $\eta^2\gamma^2\beta^2 P_{\Delta N}(\omega)$
- photon shot noise: $\eta^2\gamma^2\beta\,\langle N(t)\rangle_t$
- partition noise: $\eta\gamma(1-\eta\gamma)\beta\,\langle N(t)\rangle_t$

Here β denotes the photon scattering rate per atom ($\beta \gg 1/s$), γ the solid angle of detection in units of 4π, and η the quantum efficiency of photomultiplier. These three factors will be combined in a single parameter $\alpha:=\eta\gamma\beta$, describing the collection efficiency. Using $P_{\Delta N}$ from eq. (10), $\sigma^2(\tau)$ can be computed according to eq. (6):

$$(11) \qquad \sigma^2(\tau) = 2R_N\alpha^2\tau\theta_T\left[\frac{\tau}{\theta_T}-1+e^{-\frac{\tau}{\theta_T}}\right]+\alpha R_N\theta_T\tau$$

This leads to a signal-to-noise ratio $S/N(\tau)$:

$$(12) \qquad S/N\left(\frac{\tau}{\theta_T}\right) = \frac{1}{\sqrt{2}}\sqrt{\langle N\rangle_t}\,\frac{\tau/\theta_T}{\sqrt{f\left(\frac{\tau}{\theta_T}\right)+\frac{1}{2\alpha\theta_T}\frac{\tau}{\theta_T}}} .$$

where $f(x)=x-1+e^{-x}$.

This result is graphically presented for realistic experimental conditions in fig. 2. One can clearly see the deviation from the asymptotic $\sqrt{\tau}$-behavior. With atomic shot noise only, S/N would tend towards a constant value for short integration times; however, even with excellent collection optics, sooner or later photon shot noise will become dominant.

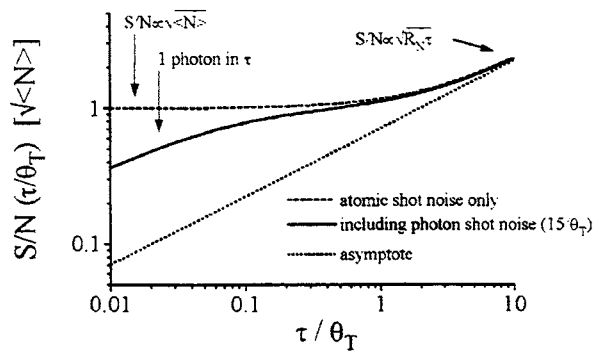


Fig. 2 Theoretical curves for S/N as a function of integration time τ scaled by the trap lifetime θ_T. The limiting cases of pure atomic shot noise and purely white noise are indicated

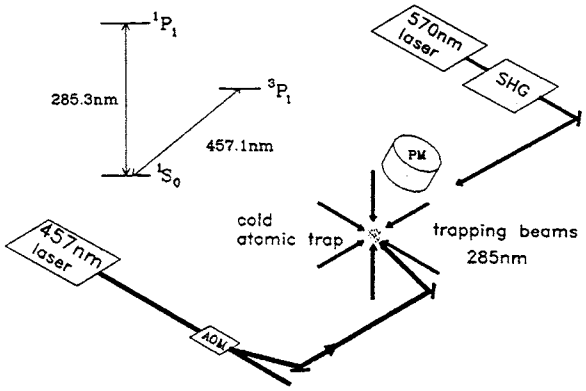


Fig. 3 Experimental setup showing the MOT, the UV laser system for cooling and trapping, and the 457 nm dye laser spectrometer. The Mg level scheme indicates the cooling transition and the narrow intercombination line

Comparison with experiment

Our experiment consists of a magneto-optical trap (MOT) for magnesium atoms and has been described in detail previously [1,2]. A brief overview is given in fig. 3. The fluorescence of the trapping transition is monitored with a photomultiplier (PM). We have observed the trap fluorescence for different integration times τ, different mean numbers of trapped atoms $\langle N\rangle$ and different collection efficiencies α. Typical results are presented in fig. 4. The theoretically predicted behavior is clearly visibly. Fitting eq. (12) to the data using $\langle N\rangle$, $\alpha\theta_T$ and θ_T as parameters yields these important trap constants with high accuracy. We can determine these values in a completely independent way based on the observation of single trapped atoms [8] and find good agreement.

S/N in the case of a spectroscopic trap loss

The first important application of the above theory is to consider the influence of spectroscopic trap loss. During excitation of the weak intercombination transition atoms are shelved at a rate ρ in the metastable $^3P_1$-level [1,9]. Since they no longer absorb the UV radiation, they are lost from the trapping region. This leads to a reduction in trap lifetime and in the number of trapped atoms. Equally important is a modification of the S/N behavior as indicated in fig. 5. Both experimental and theoretical curves are presented. One can see that for strong additional losses the typical shape of $S/N(\tau)$ of fig. 2 becomes less and less pronounced. Thus the model can be successfully used to predict signal-to-noise ratios for different levels of spectroscopic excitation.

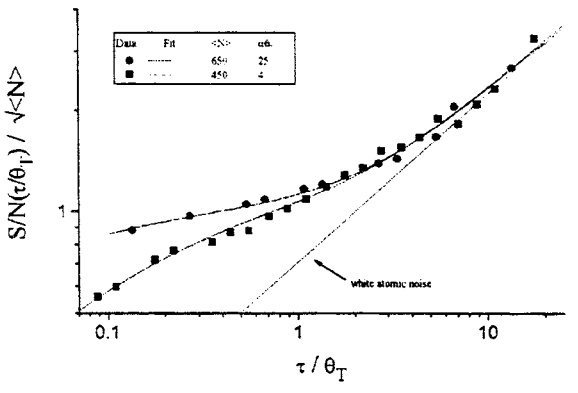


Fig. 4 Experimentally observed S/N. The solid line is a fit of the model to the data. The resulting trap parameters are indicated in the box

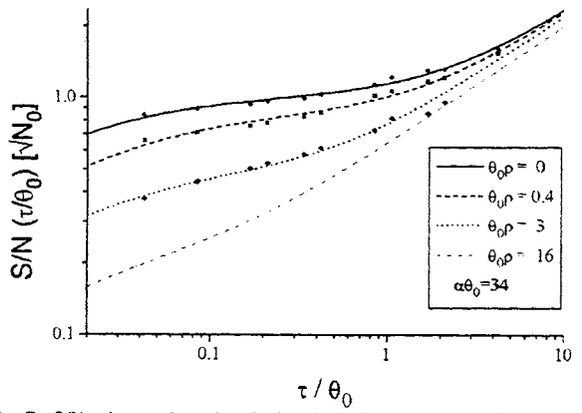


Fig. 5 S/N observed and calculated during cw excitation on the intercombination line. Different excitation rates ρ are indicated

S/N in the case of pulsed spectroscopy

For frequency standard's applications the important case to consider is pulsed excitation of the clock transition [1,2,3,9]. As indicated in fig. 6, pulsed excitation leads to a modified time behavior of the number of trapped atoms [1,9]. After a few trap lifetimes the average number of atoms decreases from its undisturbed value $\langle N_0 \rangle$ to a new equilibrium $\langle N_P \rangle$. However, the repeatedly occurring pulsed excitation P leads to a cyclic modulation of the number of trapped atoms with period t_{cyc}. In this case, two alternative measurement strategies are available to determine the influence of the weak line excitation. Strategy [A] is to integrate the trap fluorescence over a time $\tau_{[A]} \gg t_{cyc}$, and then compare the average values $\langle N_0 \rangle$ and $\langle N_P \rangle$:

$$(13)\quad [A] \qquad S(\tau_{[A]}) = 1 - \frac{\left\langle \int_t^{t+t_{[A]}} \alpha N_P \, dt' \right\rangle_t}{\left\langle \int_t^{t+t_{[A]}} \alpha N_0 \, dt' \right\rangle_t}$$

Strategy [B] is to integrate the fluorescence for short time periods $\tau_{[B]} \ll t_{cyc}$ before and after the pulsed excitation P, and then to average over several cycles quotients of such photon counts:

$$(14)\quad [B] \qquad S(\tau_{[A]}) = 1 - \left\langle \frac{\int_t^{t+\tau_{[B]}} \alpha N \, dt'}{\int_{t-\tau_{[B]}}^{t} \alpha N \, dt'} \right\rangle_{M=\frac{\tau_{[A]}}{t_{cyc}}}$$

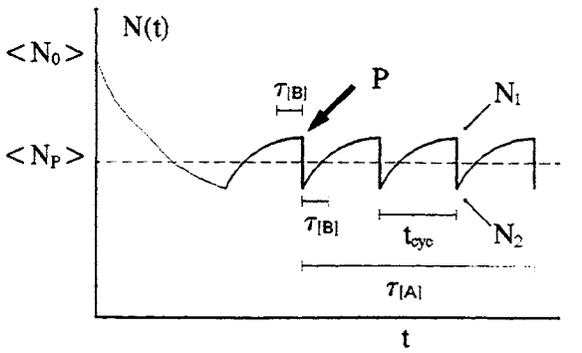


Fig. 6 Sketch of the time behavior of the number of atoms in the trap under pulsed excitation. After an initial decrease a new mean value is reached together with a cyclic modulation

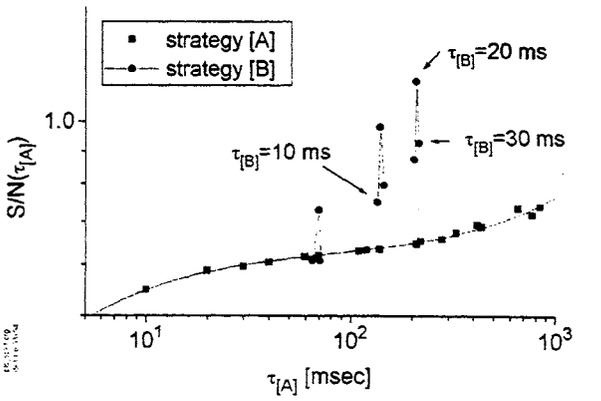


Fig. 7 Observed S/N for the two different measurement strategies [A] and [B]. Strategy [B] surpasses the atomic shot-noise limitations

Depending now on a further parameter $\tau_{[B]}$, different S/N values can be obtained.

Figure 7 shows experimental comparisons for S/N obtained from the same set of data. One can see that strategy [A] leads to the general form known from cw excitation. Depending strongly on $\tau_{[B]}$, strategy [B] can yields significant improvements in signal-to-noise ratio. Qualitatively, this can understood as follows: If the integration time $\tau_{[B]}$ is chosen too short, photon shot noise will degrade S/N, and if it is too long, the averaging effect of the cycles will already be felt. The development of a full theory of S/N for this case is necessary to adapt the detection parameters to the specific experimental conditions. Work on this is in progress.

Conclusions

A theoretical model for the signal-to-noise ratio of fluorescence from laser-trapped atoms has been presented and compared with experiments. For the case of no additional excitation, noise analysis can yield important data of the trap behavior. The spectroscopically relevant cases of cw and pulsed excitation on a weak clock transition lead to a modification of the signal-to-noise behavior that can also be understood within the present model.

The model points out the limitations imposed by atom number fluctuations on the maximum obtainable signal-to-noise ratio for high resolution spectroscopy. Since trapping very large number of atoms is often not practical for various reasons, a different measurement strategy can be used during pulsed excitation to overcome this limit. First experiments in this direction have been presented.

This work was supported by the Deutsche Forschungsgemeinschaft.

References:

[1] K. Sengstock et al.: Appl. Phys. B **59**, 99 (1994)
[2] V. Rieger et al.: Proc. 8th EFTF (1994), p. 33
[3] Th. Kisters et al.: Appl. Phys. B **59**, 89 (1994)
[4] C. Audoin in: *Metrology and Fundamental Costants* (Proc. Int. Fermi School of Physics), F. Milone, P. Giacomo, S. Leschiutta, eds., (North-Holland, Amsterdam, 1980) p.181
[5] A. Yariv *Optical Electronics* (HBJ Publishers, Fort Worth, 1991) p. 362
[6] Ch. Heiden: Phys. Rev. **188**, 319 (1969) and references therein
[7] V. Giordano et al.: Opt. Comm. **67**, 287 (1988)
[8] D. Bettermann et al.: to be published
[9] K. Sengstock et al.: Opt. Comm. **130**, 73 (1993)

SPECTROSCOPIC STUDY OF THE CaII LONG-LIVED 3D-DOUBLET IN VIEW OF A FREQUENCY STANDARD APPLICATION

Martina KNOOP, Joëlle ROCHER, Michel VEDEL and Fernande VEDEL

Université de Provence, PI2M (U.R.A. CNRS 773), Case C21,
F 13397 Marseille Cedex 20, France, Tel. 91 28 81 45, Fax 91 28 87 45

The electrical quadrupole transitions of Ca^+-ions are very well suited for metrological applications in the optical domain using the ion storage technique. We studied the behavior of the metastable 3D-doublet of CaII under the influence of collisions with particles of the background pressure. Quenching and j-mixing processes could be observed and quantified. We measured the corresponding rate constants for ionic energies between 0.3 and 1.3eV. The experiments have been carried out using direct laser excitation of the forbidden transitions, leading to very precise results.

One of the most interesting proposals for a frequency standard in the optical domain concerns the electrical quadrupole transition $3D_{5/2}$-$4S_{1/2}$ at 729nm of Ca^+-ions confined in a radio-frequency trap [1]. The long natural lifetime of the 3D-doublet ($\tau \approx 1s$) of Ca^+ allows to estimate a potential line quality $\nu/\Delta\nu$ superior to 10^{15}.

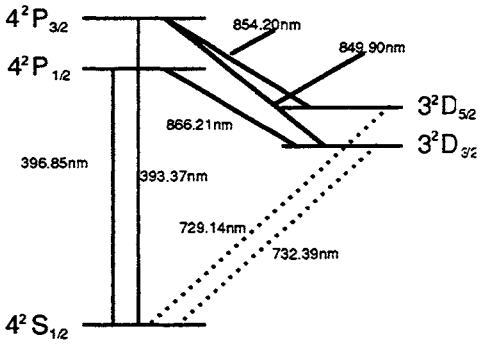


Figure 1: First energy levels of CaII .

Maybe the main advantage of the Ca^+-ion is the fact, that all the wavelengths involved in the experiment can be generated by diode lasers, either directly, using commercially available semiconductors (850 and 730nm), or by frequency-doubling of an appropriate high-power diode (397nm). In first order, the clock frequency will be magnetic field independent, since calcium has an isotope with odd nuclear spin (^{43}Ca), providing the opportunity to work on $m_F=0 \rightarrow m_{F'}=0$ hyperfine transitions. The confinement of the ions in a radiofrequency trap guarantees long observation times, the use of a miniature ion trap will be necessary for accessing the Lamb-Dicke regime where the main contribution of line broadening (the first order Doppler effect) becomes negligible.

Under extreme conditions, collisions with the background pressure play an important role for the broadening of the clock transition [2]. To explore their influence, we investigated the behavior of the metastable 3D-doublet in the course of elastic and inelastic collisions [3], [4]. The transfer of internal energy during a collision can give rise to a de-excitation of the metastable state towards the ground state. In addition, mixing of the two fine-structure levels is expected to be even more probable, as the energy difference is much smaller. We have investigated both processes under the influence of different neutral buffer gases.

In our experiment, Ca^+ ions are confined in a medium-size Paul trap (r_0=7.2mm) which allows long observation times with high resolution [3]. The ring of the trap is made from stainless steel, whereas both endcaps are formed from molybdenum mesh of high transmission. An aspheric lens placed near one of the endcaps collimates the fluorescence light of the ion cloud onto a bluepeaked photomultiplier (Hamamatsu H4730), run in the photon counting mode. Calcium ions are created by electron bombardement of an atomic beam produced by a small oven. The background pressure can be monitored by a quadrupole mass spectrometer (Balzers QMG 064) as well as a nude Bayard-Alpert jauge. After two days of bake-out, residual pressures inferior to $1x10^{-9}$mbar are attained. Different neutral gases can be introduced into the ultrahigh vacuum vessel by a needle valve avoiding contamination by impurities. For spectroscopic investigations we use a cw Titanium-Sapphire laser (Coherent 899-21), pumped by an Argon ion laser (Coherent Innova 310). Intra-

cavity frequency doubling by a LIO_3 crystal furnishes more than 15mW of power in the ultraviolet (397-393nm). The fundamental wave of the solid-state laser delivers up to 1W at 729nm, necessary for the direct excitation of the forbidden transition. The wavelengths at around 850nm are generated by grating-stabilised diode lasers with powers around 1mW. All the laser beams cross the trap diagonally (Figure 2).

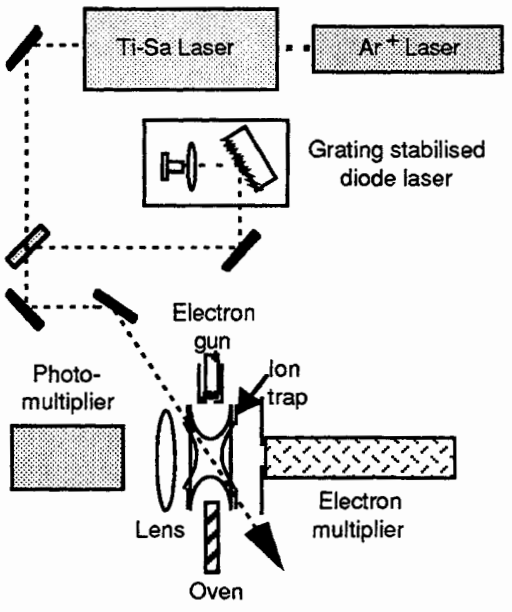


Figure 2: Experimental set-up

The measurements of the natural lifetime of the metastable levels as well as quenching and fine-structure mixing rates have been carried out using two methods (Figure 3).

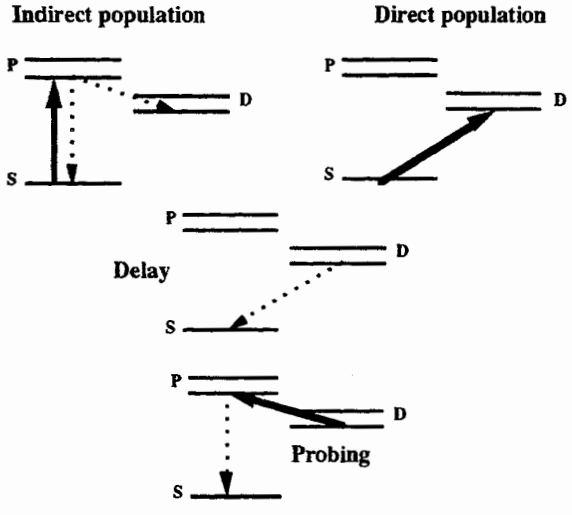


Figure 3: The different methods used to measure the effective lifetime of the 3D-levels.

The first method consists of populating the metastable state by excitation of the very short-lived P-level which decays into the D-state due to a high branching ratio (~5%, [5]). This technique has been used elsewhere with an ion cloud at some fractions of eV [6] and with a single laser-cooled ion [7].

In the second method, the metastable level is populated by direct excitation of one of the electric quadrupole 4S — 3D transitions (732 and 729 nm). In both cases the D-state population is probed via excitation of the corresponding 4P—3D transition using a diode laser and observation of the fluorescence on the resonance line. The decay of the D-state population is measured by varying the delay between excitation and probing of the metastable state (Figure 4).

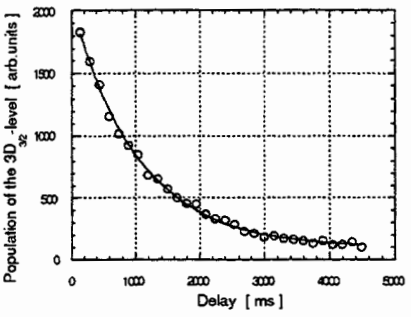


Figure 4: Decay of the $D_{5/2}$ state at a buffer gas pressure $p=1x10^{-7}$ mbar.

The latter scheme has been found to be preferable, since only the investigated level is populated. This allows better definition of the initial conditions, which is important because of the fine-structure mixing.

Measurements are performed on about $5x10^4$ ions in the pressure regime between 10^{-8} and 10^{-5} mbar, in the presence of different buffer gases (Figure 5). The data have been fitted by a χ^2-minimization using a multilinear equation. This is important because every single measured lifetime is described by a set of partial pressures which all contribute to the collisional de-excitation of the metastable state in a different way. In fact, they are given by

$$\frac{1}{\tau_{eff}} = \frac{1}{\tau_{nat}} + \sum_i \left(n_B{}^i \Gamma_Q{}^i + n_B{}^i \gamma_{mix}{}^i \right) \qquad (1),$$

where τ_{eff} and τ_{nat} designate the effective and the natural lifetimes, n_B is the density of a partial pressure denoted i, and Γ_Q and γ_{mix} denote the corresponding quenching and j-mixing rate constants. In our adjustment of the quenching influence, the j-mixing process is neglected, since it is known to take place on a different time scale [8].

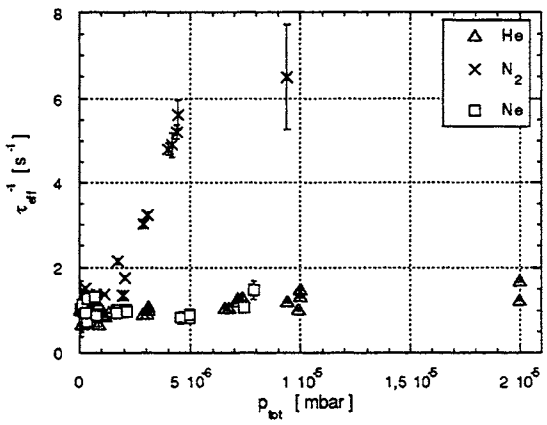


Figure 5: Effective lifetimes of the $3D_{5/2}$-level. as a function of the total buffer gas pressure.

Extrapolation to zero-pressure yields lifetimes of 1111 ± 46 ms ($3D_{3/2}$) and 994 ± 38 ms ($3D_{5/2}$) which are in very good agreement with theoretical calculations [9], [10], [11] and previous experimental values ($\tau(D_{5/2})$=1054±61ms, $\tau(D_{3/2})$ =1113±45ms [6], $\tau(D_{5/2})$=1080±220ms [7]).

The multilinear least-squares fit allowed us to determine the quenching rates for Helium (Γ_Q(He)=(1.05±0.42)x10^{-12}cm^{3}s^{-1}), Neon (Γ_Q(Ne) =(0.9±0.7)x10^{-12}cm^{3}s^{-1}) and Nitrogen ($\Gamma_Q(N_2)$ = (1.7±0.2)x10^{-10}cm^{3}s^{-1}). Careful bakeout of the apparatus has made the pressures of all other gases negligible. The quantification of relaxation velocities is very important for elimination of collisional broadening. Quenching rate constants for collisions with rare gas are particularly interesting in astrophysics where the electric quadrupole lines of CaII are used as calibration lines in stellar spectra [12], [13].

Fine-structure mixing of the two metastable states has been estimated qualitatively as a function of population difference using the method of sensitized fluorescence. The collisional mixing of the neighbouring levels is then monitored as a function of buffer gas pressure (Figure 6).

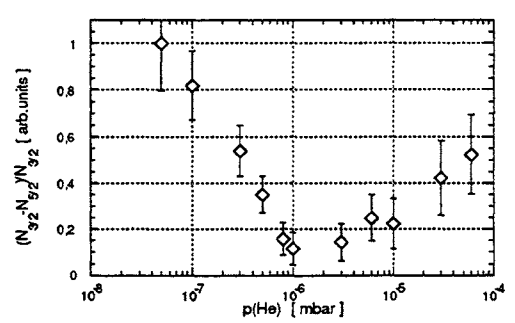


Figure 6: Qualitative behaviour of the fine-structure mixing as a function of buffer gas pressure (helium).

Time resolved measurements of populating dynamics allows the determination of the fine-structure mixing rates for different gases (Figure 7). To obtain this curve, the $D_{3/2}$-level was kept empty by a diode laser, tuned to the corresponding P-D transition, while the $D_{5/2}$-level was directly populated by excitation at 729nm. After a variable delay during which both lasers were turned off, the $D_{3/2}$-level was probed.

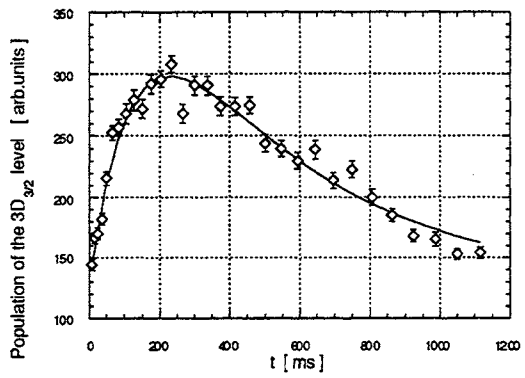


Figure 7: Temporal evolution of the $D_{3/2}$-state population at p(He)=3x10^{-7} mbar

The temporal evolution of the $D_{3/2}$-population shows, that the level is rapidly populated by collisions, before the influence of quenching and radiative de-excitation towards the ground state becomes important. The dynamics of the D-states due to fine-structure mixing, quenching and the finite lifetime is described by

$$N_{5/2}(t) = N_0 \frac{\gamma_{35}}{\gamma_{53} + \gamma_{35}} \left[\exp\left(-\frac{\Gamma_{fond} t}{2}\right) - \exp\left(-\frac{\Gamma_{dès} t}{2}\right) \right] \quad (2),$$

where $\Gamma_{fond} = \Gamma_{nat} + \Gamma_Q n_B$
and $\Gamma_{dès} = \Gamma_{nat} + (\Gamma_Q + 2\gamma_{35} + 2\gamma_{53}) n_B$,
γ_{35} and γ_{53} are the mixing rates from the $D_{3/2}$ to the $D_{5/2}$ level and vice versa, Γ_{nat} is the inverse of the natural lifetime.

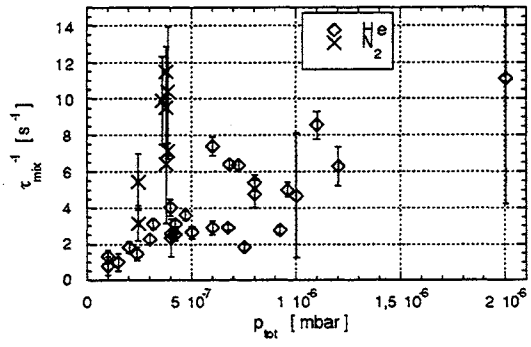


Figure 8: Effective mixing time constants as a function of buffer gas.

The application of equation (2), taking into account the principle of detailed balancing ($\gamma_{35}/\gamma_{53}\approx1.5$, for our experimental conditions), allows to determine the effective mixing time constants (Figure 8). Mixing rate constants for the different buffer gases have been determined by a multilinear least-squares fit, and are the following:

	γ(He) [cm^3 s^{-1}]	$\gamma(N_2)$ [cm^3 s^{-1}]
$D_{3/2}\rightarrow D_{5/2}$	$(2{,}24\pm0{,}1)\cdot10^{-10}$	$(2{,}8\pm0{,}3)\cdot10^{-9}$
$D_{5/2}\rightarrow D_{3/2}$	$(1{,}2\pm0{,}7)\cdot10^{-10}$	$(1{,}26\pm0{,}1)\cdot10^{-9}$

As there exists almost no data for fine structure mixing in the case of alkaline earth ions with neutral gases, we compared our values to the measurements which have been performed on alkaline atom-rare gas systems. This comparison is straightforward because of the similar atom structure of the colliding particle. Our rate constant values agree well with the measurements made by [14] and [15]. In addition, we found that the j-mixing rate constants can be described quantitatively by the semiclassical approach of the Landau-Zener model [4].

This work was supported by the Direction des Recherches et Etudes Techniques, M.K. thanks the European Community (HC&M program).

[1] G.WERTH, Proc. Summer School "Appl. Laser Spectr.", San Miniato, W.Demtröder and M.Inguscio (eds.), Springer, Berlin 1989
[2] D.J.WINELAND, J.C.BERGQUIST, J.J. BOLLINGER, W.M.ITANO, F.L.MOORE, J.M.GILLIGAN, M.G.RAIZEN, C.S.WEIMER, C.H.MANNEY, Proc. Intern. School of Physics "Enrico Fermi", Course CXVIII, E.Arimondo, W.D.Phillips, F.Strumia (eds.), North-Holland, Amsterdam 1992
[3] M.KNOOP, M.VEDEL, F.VEDEL, J.Phys.II (France) **4**, 1639-1650 (1994)
[4] M.KNOOP, M.VEDEL, F.VEDEL, submitted to Phys.Rev.A.
[5] A.GALLAGHER, Phys.Rev. **157**, 24-30 (1967)
[6] F.ARBES, M.BENZING, T.GUDJONS, F.KURTH, G.WERTH, Z.Phys.D **29**, 159-161 (1994)
[7] S.URABE, K.HAYASAKA, M.WATANABE, H.IMAJO, R.OHMUKAI, R.HAYASHI, Appl.Phys.B **57**, 367-371 (1993)
[8] D.A.MILLER, L.YOU, J.COOPER, A. GALLAGHER, Phys.Rev.A **46**, 1303-1309 (1992)
[9] C.GUET, W.R.JOHNSON, Phys.Rev.A **44**, 1531-1535 (1991)
[10] N.VAECK, M.GODEFROID, C.FROESE FISCHER, Phys.Rev.A **46**, 3704-3716 (1992)
[11] T.BRAGE, C.FROESE FISCHER, N.VAECK, M.GODEFROID, A.HIBBERT, Phys.Scripta **48**, 533-545 (1993)
[12] L.M.HOBBS, A.M.LAGRANGE-HENRI, R.FERLET, A.VIDAL-MAJAR, D.E. WELTY, Astrophys.J. **334**, L41-L44 (1988)
[13] C.ZEIPPEN, Astron.Astrophys.**229**, 248-252 (1990)
[14] L.KRAUSE, Appl.Opt. **5**, 1375-1382 (1966)
[15] M.PIMBERT, J.Phys.(Paris) **33**, 331-343 (1972)

Influence of Spurious Microwave Fields in Cesium Clocks

B. BOUSSERT, G. THEOBALD and N. DIMARCQ

LABORATOIRE DE L'HORLOGE ATOMIQUE
Unité Propre de Recherche du CNRS associée à l'Université Paris-Sud
Bâtiment 221 - Université Paris-Sud - 91405 Orsay Cédex

Introduction/Abstract.

The frequency evaluation of our miniature cesium beam resonator required both the knowledge of the long-term frequency stability and the accuracy. Only the short and medium term stability matched our hopes but the accuracy did not. The reason of this discrepancy was a high sensitivity of the output frequency of our resonator to the microwave power injected in the Ramsey cavity [1].

We have suspected the presence of spurious microwave fields in the vicinity of the cavity. Such microwave leakages are known to induce unexpected and large frequency offsets.

As a follow-up to the first investigations of Andreas Bauch [2], we have tried to develop a simple model to predict the behaviour of the atomic system travelling in an ideal Ramsey cavity surrounded by stray microwave running waves. This approach is based on the density matrix formalism and led to frequency offset as a function of the amplitude of these waves. Moreover, we are able to determine the domains :

- between an optical zone and an arm of the cavity,
- within the drift space of the cavity,

where this perturbation is the most detrimental.

In parallel to this calculation, for a given beam hole, we run experiments devoted to find an empirical relation between the level of microwave leakage detected at the output of the cut-off waveguide to the geometrical characteristics of this latter. For this purpose, we have realized several $\frac{3\lambda g}{2}$ test cavities differing one from the other by the length of the cut-off rectangular waveguide. The results have been gathered in a table. Such measurement have been done with a high sensitive spectrum analyser (- 130 dBm).

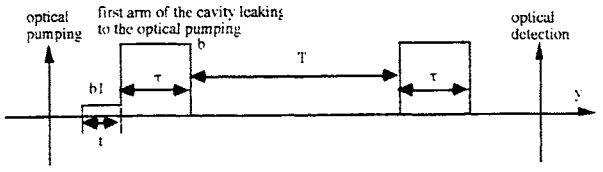
1. Theoretical approach.

1.1 set of the problem

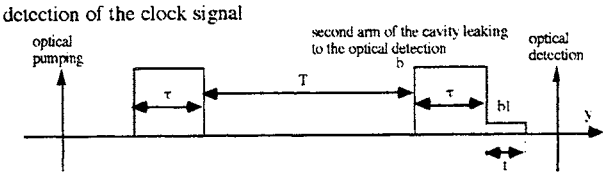
The aim of this approach is the evaluation of the frequency offset induced by the presence of microwave leakages around the Ramsey cavity. We have restricted our study to the emission of microwaves along the atom beam by a beam hole of the cavity. These latter are described by a free running sine wave of constant amplitude.

During this evaluation we have considered a Ramsey cavity where the microwave magnetic induction interrogating the atoms is the sum of the usual standing wave process of amplitude b and a free-running wave process of amplitude b_1, as shown on the two figures below :

- microwave leakage of amplitude b_1 (counter-propagating wave) between the optical pumping and the first arm of the cavity with microwave amplitude b



-microwave leakage of amplitude b_1 between the second arm of the cavity with microwave amplitude b and the optical detection of the clock signal



However, the determination of the transition probability in these two cases requires the characterization of the interaction of the atom with the free-running wave of constant amplitude b_1.

1.2 pre-requisite calculation

Our aim here is to find a matrix characterizing the interaction of a two-level atomic system with a microwave magnetic induction using the formalism developed in [2,3]. This latter is described by a sinusoidal function of time t and position y, $B_y(t)$:

$B_y(t)=B_1 \cos(\omega t - ky)$

B_1 is a constant amplitude, k is the wavelength vector and y is the direction of the atom beam.

The evolution of density operator obeys to the equation :

$$i\hbar \frac{d\rho}{dt} = [H_0 + V(t), \rho],$$

with H_0 the unperturbed hamiltonian and V(t) the perturbation proportional to $B_y(t)$.

The resolution of the above differential equation system in the referential of the atom leads to the following 3*3 matrix.

$$\begin{pmatrix} \text{coskvt} & .\pm \text{sinkvt} & \pm\frac{b_1}{kv}(1 - \text{coskvt}) \\ -+ \text{sinkvt} & \text{coskvt} & \frac{b_1}{kv}\text{sinkvt} \\ \pm\frac{b_1}{kv}(1 - \text{coskvt}) & -\frac{b_1}{kv}\text{sinkvt} & 1 \end{pmatrix}$$

the sign ± is related to direction of propagation of the perturbing running wave. t is the duration of the perturbation, b_1 is the Rabi pulsation of the perturbation magnetic field leaking from a beam hole of the cavity, v is the atom velocity.

1.3 expression of the frequency offset

Since each part of the Ramsey cavity is now described with a 3*3 matrix, we can evaluate the transition probability in the two cases described in §1.1 with the formalism developed in [2,3]. The frequency offset is then derived by differentiating the transition probability versus the frequency in the vicinity of the cesium frequency transition.

In the first case, the frequency offset writes down :

$$\omega - \omega_0 = -\frac{2\, b\tau}{kv\ \tau^2\left(\frac{T}{\tau} + \frac{2}{b\tau}\right)}\ \kappa,$$

where κ is the ratio of the amplitude b_1/b and ω_0 is the cesium transition frequency

with numerical values : $b\tau = \pi/2$; $\frac{T}{\tau} = 16$; $k = 193\ m^{-1}$

$$v = 200\ ms^{-1}$$

$$\omega - \omega_0 = 1888\ \kappa$$

How important this effect can be on the accuracy of the output frequency of our resonator ?

$\kappa=10^{-4}$ gives $\omega - \omega_0$=0.18 Hz or $\frac{\omega-\omega_0}{\omega_0} = 2\ 10^{-11}$

$\kappa=10^{-5}$ gives $\omega - \omega_0$=0.018 Hz or $\frac{\omega-\omega_0}{\omega_0} = 2\ 10^{-12}$

$\kappa=10^{-6}$ gives $\omega - \omega_0$=0.0018 Hz or $\frac{\omega-\omega_0}{\omega_0} = 2\ 10^{-13}$

Let take the case $\kappa=10^{-5}$ and determine the equivalent microwave leakage detected. Using the same numerical values as above the power equivalent to b = 31000 s^{-1} is about $1.2 10^{-12}$ W ; thus by adapting κ to power it comes that P_{leak} = -110 dBm.

In the second case, the frequency offset writes down :

$$\omega - \omega_0 = -\frac{2\, b\tau}{kv\ \tau^2\left(\frac{T}{\tau} + \frac{2}{b\tau}\right)}\ \kappa,$$

with the same numerical values as above :

$$\omega - \omega_0 = -1888\ \kappa$$

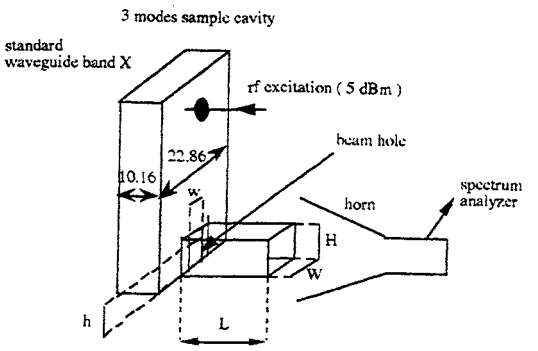
The treatment of more complex situations, like the leakage of two beam holes, requires a numerical computation.

2. Experimental data.

In order to design the right cavity, we need previously to answer the following question : given a beam hole of section w*h (see figure below), what should be the minimal length L of the cutoff waveguide assuring a minimum level of microwave outside of the cavity.

Instead of plunging in a tough electromagnetic description of the fields in the vicinity of the beam hole, we have decided to measure the level of microwave power leaking from the output of the cutoff waveguide of a sample cavity. Given this latter is fed by a signal of power 5 dBm.

For this purpose, we have designed a set of several $\frac{3\lambda g}{2}$ cavities tuned at the cesium transition frequency and differing one from the other by the length of the cutoff rectangular waveguide and the beam hole section as shown on the figure below.



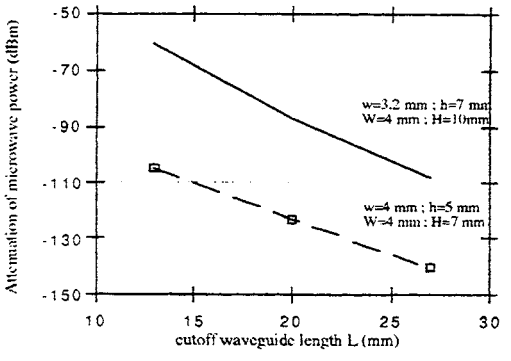
The accuracy of our measurement required a perfect shielding of the excitation chain of the cavity.

The results of this measurements have been collected in the following table.

cut-off waveguide length L (mm)	cut-off waveguide section W*H (mm^2)	beam hole w*h (mm^2)	level of leakage P_{leak}(dBm)
27	4 x 10	3.2 x 7	- 108 RBW 30 Hz
20	4 x 10	3.2 x 7	- 87 RBW 30 Hz
13	4 x 10	3.2 x 7	- 60 RBW 30 Hz
27	4 x 7	4 x 5	< - 125 RBW 30 Hz
20	4 x 7	4 x 5	- 123 RBW 30 Hz
13	4 x 7	4 x 5	- 105 RBW 30 Hz

Given these experimental measures, we have plotted the attenuation of the microwave power by the cutoff waveguide versus its length. The attenuation A is here given by the relation :

$$A(dBm) = P_{leak}(dBm) - P_{injected}(dBm)$$



From the curves, we note a linear dependence of the attenuation versus the length of the cutoff waveguide whatever the power fed to the cavity. We can evaluate this dependence to -3 dBm/mm.

Conclusion.

For a given beam hole section in the walls of the waveguide, the length of the cutoff waveguide must be chosen such as the level of microwave within this latter died away steeply.

Moreover since this frequency offset is inversely proportional to T, time of flight between the two microwave interaction regions, its effect on the accuracy of the output frequency is then worsened with a short cavity.

From above measurements and the previous calculations, we can derive that a sensitivity of 10^{-12} of the frequency to microwave leakage leads for the two cutoff waveguide configuration to the following :

- beam hole section 3.2*7 a length larger than 27 mm is required

- beam hole section 4*5 a length larger than 14 mm is required

References.

[1] P. Petit, V. Giordano, P. Cérez, B. Boussert, C. Audoin, N. Dimarcq. Proc. 8th European Frequency and Time Forum (1994), pp. 517-522.

[2] Andreas Bauch. "Majorana-Übergänge in Caesium-Atomstrahl-Frequenznormalen". PTB-Bericht, Februar 1986, pp. 20-21.

[3] J. Vanier, C. Audoin. "The Quantum Physics of Atomic Frequency Standards". Bristol and Philadelphia I.O.P. Publishing 1989.

THE M8 AND M9 HYDROGEN MASERS AS THE NATIONAL FREQUENCY STANDARD AT THE NATIONAL INSTITUTE OF METROLOGY - BUCHAREST

*Octav C.GHEORGHIU, Liviu C.GIURGIU**, Bogdan M.MIHALCEA, Dragoș M.CACICOVSCHI, Anca NICULESCU****

Inst. of Phys. and Techn. of Rad.Devices, Bucharest, ROMANIA
*)University of Bucharest, Faculty of Physics,
**)National Institute of Metrology, Bucharest.

Our paper describes the performances and characteristics of the M8 and M9 masers; which are part of the frequency standard at the Romanian National Institute of Metrology, in Bucharest. We have outlined the original solutions, which we have used.

The Frequency and Time Standard at the Romanian National Institute of Metrology, in Bucharest, consists of two Hydrogen Masers (M8 and M9) together with a GPS standard (SATSYNC II 425).

The maser frequencies are continuously compared (on 100 kHz) with the GPS standard frequency, then monitored and recorded every 3 minutes with a 10^{-11} sensibility, very inferior to the maser stability (10^{-14} for 100 s).

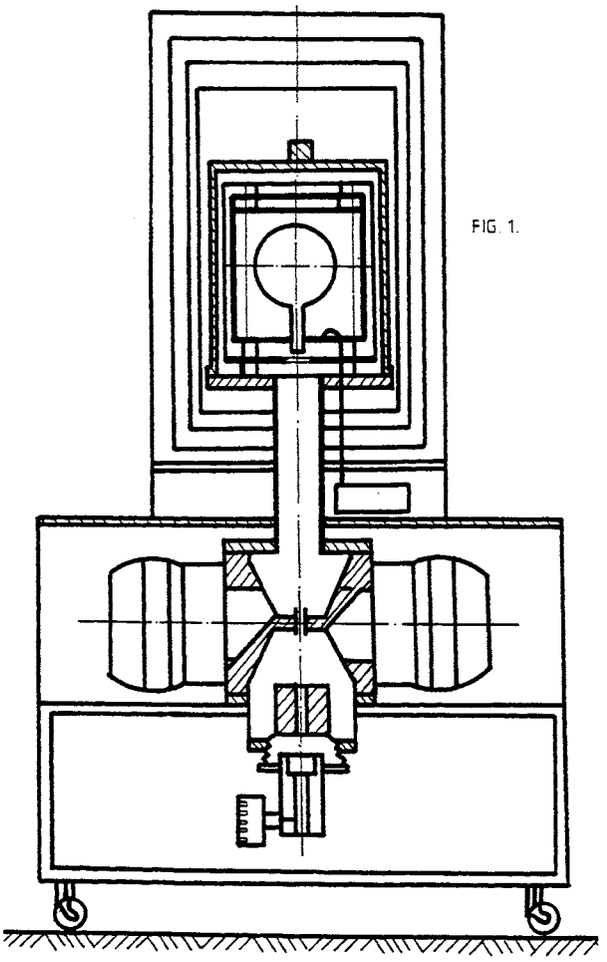


FIG. 1.

In the present paper, we describe a few of the features characterizing our masers.

Figure 1 shows the general arrangement. There are three ionic pumps mounted around an inox cube.

Four Permalloy shields (1.5 mm thick) form the magnetic shielding. On the aluminium Bell jar is situated the degaussing connection (max 1000 A). The receiver is under the magnetic shields.

The thermostatation is made with five heating windings. The cavity thermal isolation is increased by means of a thermal shield. The thermistor C in fig. 2 allows the cavity temperature determination.

The cavity temperature thermal fluctuations are $\pm 2 \times 10^{-5}$ Kh^{-1} /1/.

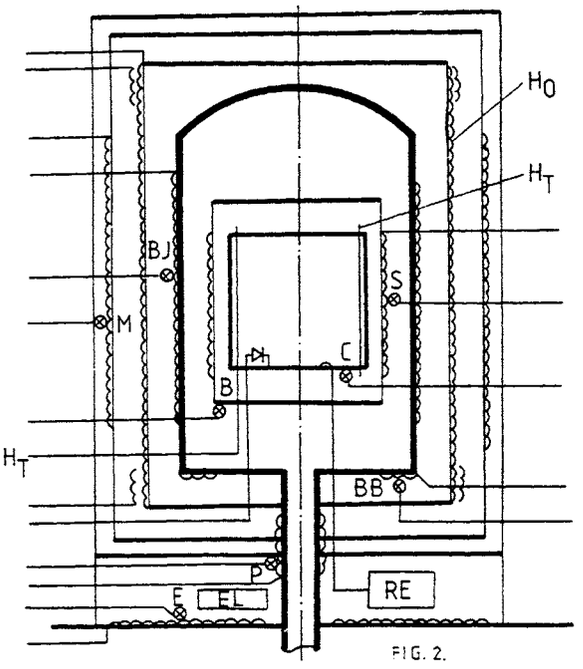


FIG. 2.

In the molecular hydrogen source, the Palladium leak is located very close to the discharge tube. Due to the pressure thermistor gauge and Palladium leak reduced time constant, the pressure servo loop time constant is very short (about 3,5 s) /2/.

The molecular dissociator is a discharge tube located at the end of a coaxial cavity tuned on 400 MHz (Fig. 3) /3/. Due to the system high efficiency, only 5 W of power is sufficient for the maximum possible dissociation degree /4/.

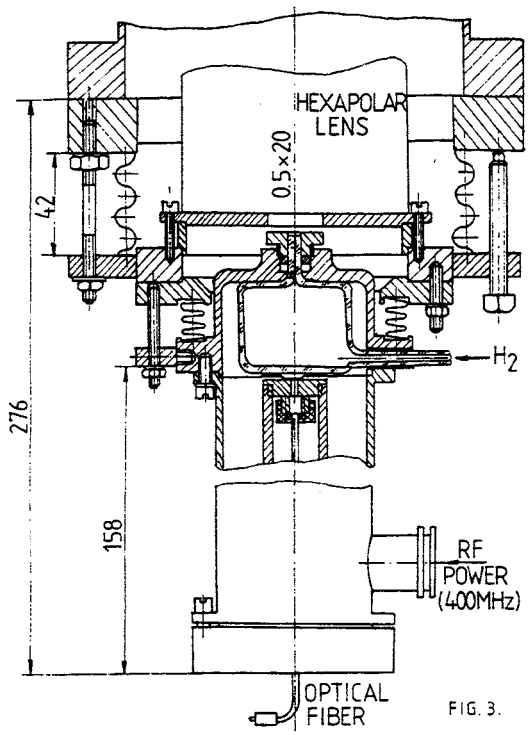


FIG. 3.

The coaxial cables for the coupling to the resonant cavity and thermal control loops (Filotex TU 165) together with the thermistors and transversal magnetic field wires, are isolated towards the vacuum vessel by nonmagnetic passages made of brass with teflon isolation (Fig. 4).

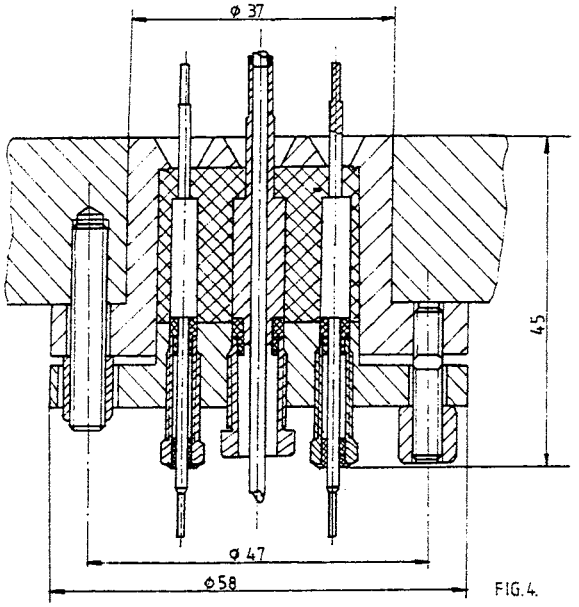


FIG. 4.

In the receiver there are four frequency changes. The synthesized signal for the phase comparator γ = 6A, BCD [Hz] is obtained by frequency divisions of the 10 MHz quartz signal (Fig. 5).

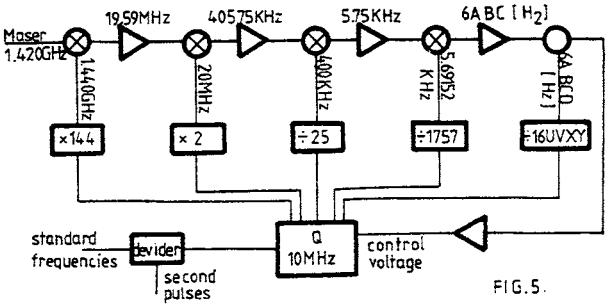
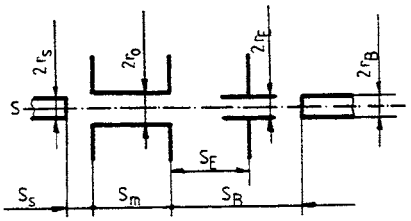


FIG. 5.

The hexapolar magnet, made of ALNICO 5 and ARCO, is designed in the usual mode /5/. The field is 0,8T at the pole tips. The diagram in Fig. 6 enables the determination of the bulb aperture ($2r_B$) focused atoms ($\mu_{ef} < 0$) velocity spectrum and also of the outward deflected atoms ($\mu_{ef} > 0$) spectrum.



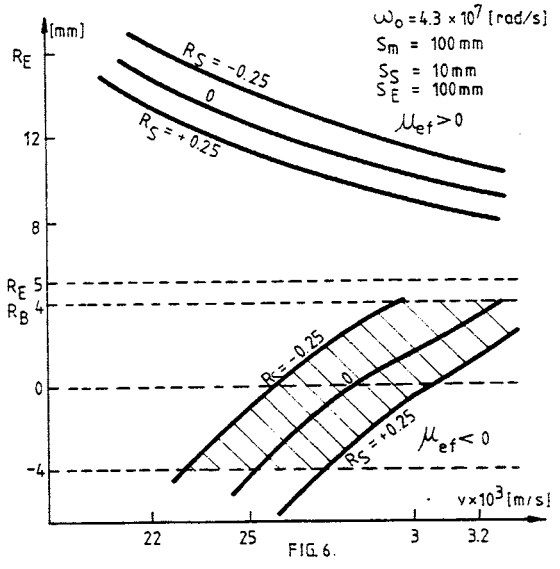


FIG. 6.

For the homogeneous H_o field measurement, a set of four rods (φ = 2 mm) is provided inside the resonant cavity, parallel with its axis and located near the walls. The rods are external to the cavity ends, connected as to form a two loops coil. The system is very efficient, and a signal higher as 5 mV kills the oscillation. External coils are not recommended because of the aluminium cavity walls thickness (10 mm).

The maser amplitude variation against the transversal field frequency shows a minimum at the π transition frequency (F = 0, m = 0) → (F = 1, m = 1). From the frequency minimum the H_o intensity results (Fig. 7). The minimum is thinner as the degaussing gets better /6/.

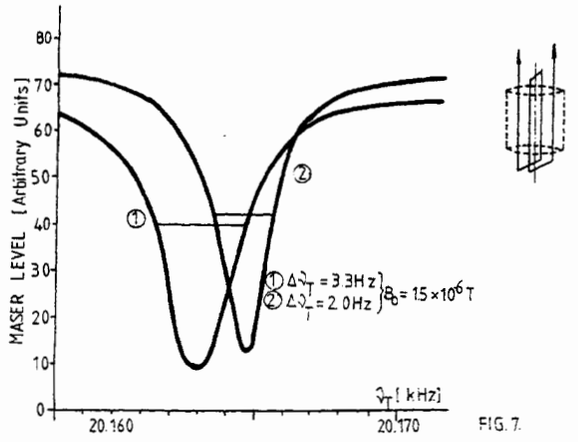


FIG. 7.

The maser frequencies biases against the ν_o = 1420405751.773 Hz /7/ unperturbed frequency are:

	M8	M9
Cavity temperature T_c	319.58K	316.54K
DOPPLER effect	-0.0625Hz	-0.0619Hz
Magnetic field H_o	+0.8668Hz	+1.2601Hz
Wall shift	-0.0279Hz	-0.0289Hz
Total bias	+0.7764Ez	+1.1693Hz
Operating frequency	1420405752.549Hz	1420405752.942Hz

In the near future the frequency standard at the National Institute of Metrology will be completed with the M3 maser submitted to a modernization process at the present time.

Within a relatively short period of time, all the masers (M3, M8 and M9) will be modified as to work, also, as passive hydrogen masers. After this process the life of the ionic pumps will be longer and we obtain an improved very long term frequency-stability.

BIBLIOGRAPHY

/1/. L.C.GIURGIU, Teza, Bucharest University, Physics Faculty (1988)

/2/. L.C.GIURGIU, M.P.DINCA, O.C.GHEORGHIU 7-th European Frequency and Time Forum, Neuchatel - 16 - 18 March, 1993

/3/. O.C.GHEORGHIU, V.N.GHEORGHE Int.J.Electronics 39, 329 (1975)

/4/. O.C.GHEORGHIU, C.M.MANDACHE CR Acad.Sci.Paris 205, 131 (1982)

/5/. C.AUDOIN J.Phys.Appl. 26A, 71 (1965)

/6/. L.C.GIURGIU, O.C.GHEORGHIU Ann.University, Iasi, sub tipar

/7/. P.PETIT, M.DESAINT FUSCIEN, C.AUDOIN Metrologia 16, 7 (1980)

CZECH NATIONAL TIME AND FREQUENCY GROUP STANDARD

Otokar Buzek, Jan Čermák, Blanka Čemusová

Institute of Radio Engineering and Electronics, Czech Academy of Sciences

Summary

Since 1993 the Time Laboratory of the Institute of Radio Engineering and Electronics (IREE) of the Czech Academy of Sciences, Prague, has been responsible for the Czech National Time and Frequency Standard. The Group Standard is currently based on four commercial cesium clocks operating at two different sites. Two clocks (one of them defined as master) are placed at IREE itself and other two clocks are operating at the International Telephone Exchange (ITE), part of the Czech Telecom, located about 10 km from IREE.

The two sites are interconnected via four fiber optic links consisting of a 1300 nm single-mode fiber and the necessary transmitting and receiving electronics. The optical carrier of each link is modulated on-off by a 5 MHz signal derived from the respective clock.

This paper is a report on present metrological parameters of the Standard. The main interest concerns the performance of the fiber optic link which determines the internal traceability within the system. To assess the link performance, hourly phase differences have been measured simultaneously at IREE and ITE for roughly a year. Thus, for instance, frequency instability of the link shows $\sigma_y(\tau) = 2.5 \times 10^{-15}$ for $\tau = 10$ days. Also seasonal variations of about ±10 ns have been found which are correlated with the outdoor temperature shifted by 15 days. These are due to the fiber buried under ground. Compensation for the temperature sensitivity can be achieved by using temperature coefficients computed for each link with the aid of the least-squares estimation technique.

1. Introduction

The IREE Time Laboratory, which is responsible for the Czech National Time and Frequency Standard, has recently passed through significant changes. Besides overall modernization, such as implementation of GPS common-view comparisons, installation of a new HP5071A cesium clock, and full automation of measurements, a collaboration of great importance was established with the ITE, Prague, which is currently the most exigent user of standard frequency in this country.

In the past ITE had been employing a 100 kHz standard frequency signal transmitted from IREE to ITE via a permanent telephone line. So when in 1992 ITE pondered a question of possible frequency sources for its newly constructed digital exchange, IREE offered its standard frequency as an alternative of the third backup frequency source for the exchange.

Though the 100 kHz signal via telephone conformed with the CCITT G.811 Standard, the telephone link was found neither reliable nor perspective. Therefore it was decided to build up a more stable link for which the most appropriate choice was a fiber optic link. An agreement was signed between IREE and ITE obliging IREE to design the whole ITE frequency reference system [1] and to take the technical responsibility for its operation in the future.

The system was implemented by the end of 1993 and it became fully operable in January, 1994. The whole project turned out to be advantageous for both partners. For Czech time and frequency metrology it provided the direct access to ITE's cesiums with the consequent traceability to the BIPM. In addition, the system became an interesting subject for extensive studies of the time transfer via a fiber optic link.

2. Description of the System

Fig. 1 shows a block diagram of the Czech National Time and Frequency Standard. Each of the four cesium clocks is generating its individual atomic time scale denoted by the clock serial number. The master clock is the IREE's HP5071A whose frequency is steered to generate the Czech official time scale called UTC(TP) (TP stands for *Tempus Pragense*).

The clocks at IREE are sited in a special underground cellar with temperature control within 22.5 ± 0.5°C. The clocks at ITE are placed in a rack separated from the exchange by a wire barrier to prevent unauthorized personnel from interfering. The room is both temperature and humidity controlled within 22.3 ± 1.5°C and 43% ± 10%, respectively.

There are four fibers laid between IREE and ITE. Three of them are connected as illustrated in Fig.1. They ensure the transfer of UTC(TP) from IREE to ITE, and AT(154) and AT(163) from ITE to IREE. The fourth fiber has been installed only recently and is currently spare.

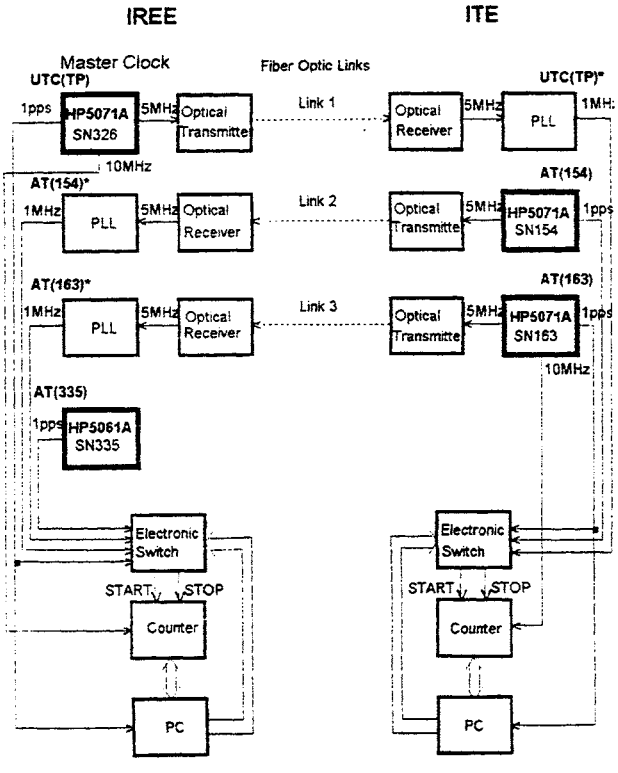


Fig.1 Block diagram of the Group Standard

As shown in Fig.1, each fiber optic link consist of a transmitter, a fiber, a receiver, and a phase locked loop (PLL).

In the transmitter the optical carrier of 1300 nm is modulated with TTL pulses derived from the zero crossings of the 5 MHz signal supplied from the corresponding clock. The optical radiated power is 4 μW when the signal is „on" and 0.4 μW when it is „off". The „on" power may be reduced by 50%.

The optical signal received at the remote end is demodulated at the optical receiver. The 5 MHz signal obtained is amplified first in an aperiodic preamplifier and then in an amplifier tuned at 5 MHz with a bandwidth of 450 kHz.

The PLLs are based on temperature controlled crystal oscillators and they are used to filter the white phase noise and to ensure the continuity of the 5 MHz signal in the case of a link failure (this is particularly important at ITE). The loop-filter transfer function, which can be modeled by

$$H(s) = \frac{P_1(s)}{sP_2(s)}, \qquad (1)$$

where $\lim_{s\to 0} P_i(s) = const = 0$, enables the PLL to maintain a constant phase difference at the phase detector for a linear change in frequency. The loop gain can be set to provide the effective bandwidth from 1 Hz to about 0.01 Hz. A coherent detector indicates the presence of the incoming 5 MHz signal which gives information on operation continuity. It is worth noting that so far no discontinuity due to a failure of the link has been observed.

The fiber optic cable is 10050 m long and on its way it passes through two telephone exchanges and through eight optical connectors. Attenuation between the end points reads 9.6 dB. The interconnection is schematically illustrated in Fig.2. Cable *a*, between IREE and the first telephone exchange, is buried about 0.5 m under ground while the longest part, cable *b*, is laid in the urban communication cable ducts some 2 to 3 m under ground. Finally, the shortest part, cable *c*, goes through the ITE building up to the digital exchange.

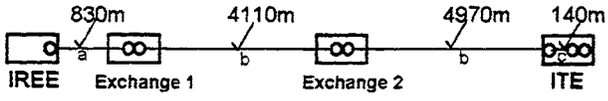


Fig.2 Fiber optic connection between IREE and ITE

The link electronics at ITE is placed in a rack next to the ITE clocks. Thus the above mentioned environmental control also applies to this electronics. At IREE, however, the link electronics had to be placed in a room without any temperature control so it is exposed to much larger temperature changes.

The measurement system at the two sites is identical both in hardware and software. It consists of a one-nanosecond counter, a fast electronic switch, and an associated PC driven in a real time.

3. Internal Traceability

In this paper only the transfer of relative time (i.e. the transfer of frequency) will be of concern. In this sense the internal traceability is determined mainly by the link instability caused by variations of its group delay and by noise.

The long-term effects can be assessed from the simultaneous measurements of phase differences that are carried out at both sites on an hour-to-hour basis. Assessing the short-term instability, however, requires a special measurement because of the poor resolution of the counters involved. In this case, the sensitivity was increased with the aid of a frequency deviation multiplier.

3.1 Long-term Instability

The hourly simultaneous phase-difference routine measurements began in February, 1994, and so far they have provided enough data to estimate the variations of group-delay over a large range of intervals. Together with these phase differences, temperature has also been measured at the critical points of the link since the influence of temperature was assumed predominant [2].

The group delay of the *i*-th link (in our case $i = 1,2,3$) can be modeled by a simple sum

$$X_i = x_o + x_i , \qquad (2)$$

where x_o denotes the constant part of the overall group delay and
x_i denotes its variable part which depends predominantly on temperature.

As noted above, the constant delay x_o will not be of interest and therefore speaking about delay, we will understand the variable part only. This can be modeled in terms of a sum of partial delays

$$x_i = x_{Ti} + x_{Fi} + x_{Ri} + x_{Pi} \qquad (3)$$

caused by the optical transmitter (x_{Ti}), optical fiber (x_{Fi}), optical receiver (x_{Ri}), and PLL (x_{Pi}). Each of these partial delays depends on local environmental conditions.

From the three following phase differences measured at ITE (here asterisk denotes the phase at the remote end, see also Fig.1),

UTC(TP)-AT(154)*, *UTC(TP)*-AT(163)*, and *AT(154)-A T(163)*,

and the corresponding differences taken simultaneously at IREE

*UTC(TP)-AT(154)**, *UTC(TP)-AT(163)**, and *AT(154)*-AT(163)**,

we can form the relations

$$Xs(1,2) = [UTC(TP)^* - AT(154)] - [UTC(TP) - AT(154)^*] \qquad (4)$$

$$Xs(1,3) = [UTC(TP)^* - AT(163)] - [UTC(TP) - AT(163)^*] \qquad (5)$$

$$Xd(3,2) = [AT(154) - AT(163)] - [AT(154)^* - AT(163)^*] \qquad (6)$$

where $Xs(1,2) = x_1 + x_2$ (7)

$$Xs(1,3) = x_1 + x_3 \qquad (8)$$

$$Xd(3,2) = x_3 - x_2 . \qquad (9)$$

The value $Xd(3,2)$ represents the differential delay between links 2 and 3, which makes it possible to assess to which extent the links are identical. The values $Xs(1,2)$ and $Xs(1,3)$ represent the sum of delays of links 1 and 2, and links 1 and 3, respectively. In the following text we will call them summary delays.

Records of delays $Xs(1,2)$, $Xs(1,3)$, and $Xd(3,2)$ taken over 336 days (from MJD 49381 to MJD 49717) are shown in Fig.3. To better distinguish between the graphs, summary delays have been given arbitrary constants. One can see that the summary delays exhibit seasonal phase variations of about 10 ns with respect to the mean. In frequency they result in maximum deviation of 2 parts in 10^{15} for a two-month averaging interval. As for the differential delay, its seasonal phase variations are so small that they cannot be observed because of the quantization noise of the counters.

For the averaging interval $\tau = 10$ days (i.e. the shortest interval reported to the BIPM), all three records in Fig.3 have about the same Allan standard deviation of frequency which reads 2.5 parts in 10^{15}.

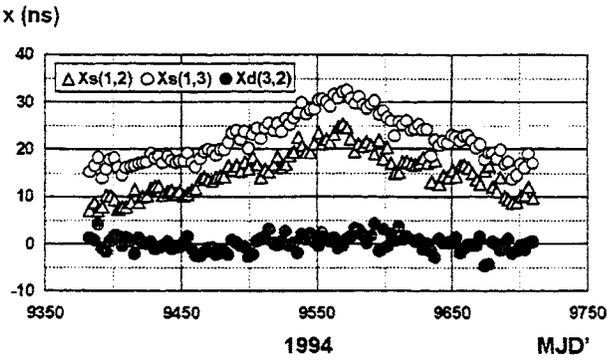


Fig.3. Records of the summary and differential delays

The summary delays have been gradually correlated with the outdoor temperature (daily mean in Prague), with the temperature in the rooms at ITE and IREE where the link electronics is located, and with atmospheric pressure and humidity. Significant correlation has been found with the outdoor temperature when shifted by an amount of 15 days. This can be explained only by the temperature sensitivity of the fiber optic cable buried under ground. There is also correlation with temperature at the IREE room where the link transmitter and the link receivers are placed..

Records of the outdoor temperature as well as the temperature at the critical room at IREE are shown in Fig. 4.

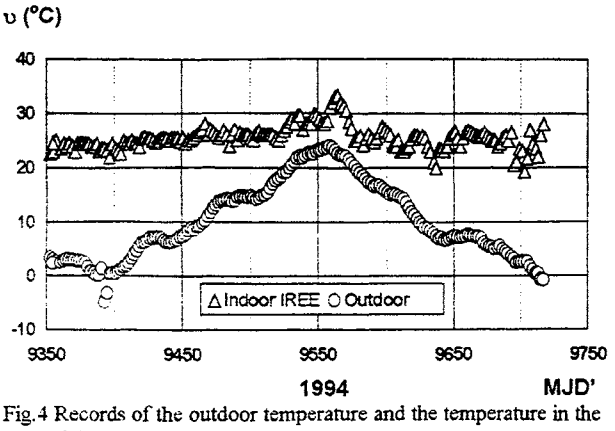


Fig.4 Records of the outdoor temperature and the temperature in the IREE room

The temperature dependence of the summary delay can be written in the form

$$Xs(1,i) = K_{1,i}\,\upsilon_1 + K_{2,i}\,\upsilon_2 \qquad (10)$$

where i = 2,3 and

υ_1 is the smoothed outdoor temperature (moving average over 31days as a rough approximation of the temperature under ground)

υ_2 is the indoor temperature in the room at IREE

$K_{1,i}, K_{2,i}$ are the corresponding temperature coefficients.

By applying the least-squares method to daily measurements for a set of N equations (j = 1 to N)

$$K_{1,i}\,\upsilon_{1j} + K_{2,i}\,\upsilon_{2j} - Xs(1,i) = v_j, \qquad (11)$$

we get the solved-for temperature coefficients $K_{1,i}$, $K_{2,i}$ in ns/°C for the summary delays as follows:

$K_{1,2}$ = 0.682, $K_{1,3}$ = 0.617
$K_{2,2}$ = −0.057, $K_{2,3}$ = −0.052

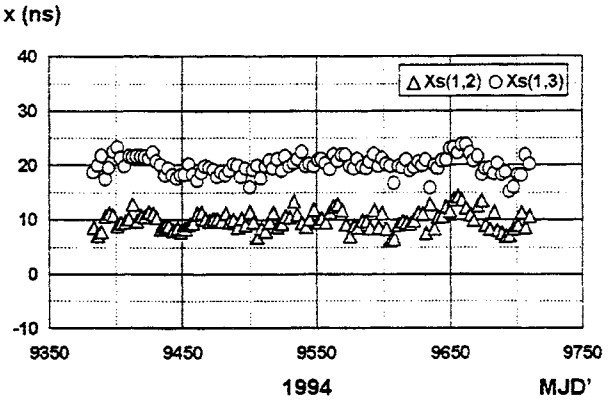


Fig.5. Summary delays after correction

Fig.5 shows the effect of the temperature correction on the summary delays $Xs(1,i)$ from Fig.3. In both cases the residual classical standard deviation is on the order of the quantization noise of the counters involved in the measurement (σ=1.6 ns).

The results indicate that the most temperature sensitive link element is the fiber optic cable itself while the sensitivity of the link electronics is smaller. This fact supported by the close correspondence of the above computed coefficients leads to the conclusion that the delays of the individual links may be considered equal within the measurement resolution limits.

3.2 Short-term Instability

In order to estimate the short-term instability of the link due to its phase noise (here „short-term" means averaging intervals 0.2 s to 1000 s), an identical 5 MHz signal was applied to links 2 and 3 at ITE and phase differences were measured between links 2 and 3 at the remote end. For this purpose, the measurement sensitivity was increased by using a frequency deviation multiplier operating at 5 MHz with the cut-off frequency 4.3 kHz. To measure only the „raw" noise, the phase differences were taken before the PLL clean-up.

Fig.6 shows the differential noise $Xd(3.2)$ in terms of Allan standard deviation of frequency on the log-log plot. The decreasing straight-line indicates prevailing phase modulation (with the slope -1) over the whole range of τ. To assess the degradation due to the links, Fig.6 also shows frequency stability of the ITE clocks as computed from the $AT(154)$-$AT(163)$ differences measured directly at ITE. One can see that if no PLL filtration is used at the remote end then $AT(154)^*$-$AT(163)^*$ begins to be degraded with respect to $AT(154)$-$AT(163)$ for τ smaller than about 100 s.

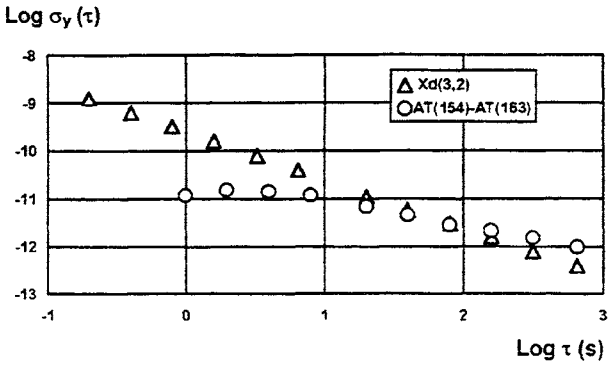


Fig.6. Differential noise $Xd(3.2)$ together with mutual frequency stability of ITE clocks

The link differential noise has been also computed by means of another stability measure [3]. TVAR or $\sigma_x^2(\tau)$, which is more appropriate for our case where phase noise is expected. It is defined as

$$\sigma_x^2(\tau) = \frac{1}{6}\left\langle \left(\Delta^2 \bar{x}\right)^2 \right\rangle \qquad (12)$$

where Δ^2 stands for the operator of the second difference .

For our computation we have used the common estimator of the true variance (12) which takes the form

$$\sigma_x^2(\tau) = \frac{1}{6n^2(N-3n+1)} \sum_{k=1}^{N-3n+1} \left\{ \sum_{i=k}^{n+k-1} \left(x_{i+2n} - 2x_{i+n} + x_i\right) \right\}^2 \qquad (13)$$

where $\tau = n\tau_0$ and τ_0 is the sampling interval.

For two different sampling intervals, we get $\sigma_x(\tau)$ as shown in Fig.7. Both straight-lines (computed for τ_{01} = 0.2 s and τ_{02} = 20 s, respectively) indicate the presence of white noise phase modulation (with the slope -1/2). The dash-and-dot line parallel to τ-axis

represents the classical standard deviation of x. The dashed-line prolongation for τ_{01} is justifiable only if the multiplier background noise can be neglected. In this case the noise has been verified only for $\tau = 0.2$ to 6.4 s, for $\tau = 6.4$ s giving $\sigma_x(\tau)$ about 0.5 ps.

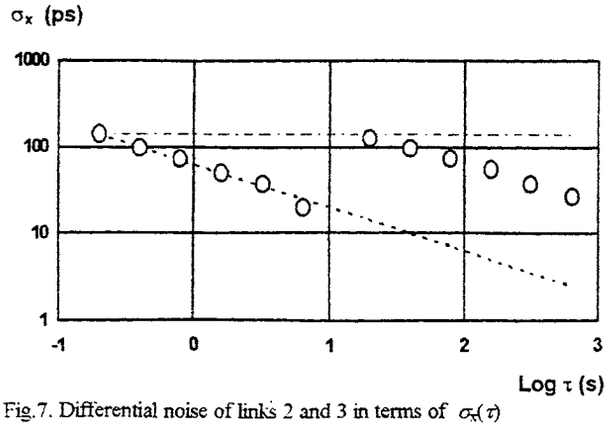


Fig.7. Differential noise of links 2 and 3 in terms of $\sigma_x(\tau)$

Assuming that the links behave identically and the phase noise in the links is uncorrelated, which may be the case for the above intervals, $\sigma_x(\tau)$ of an individual link can be estimated directly from the above differential noise.

4. External Traceability

The UTC(TP) time scale is traceable to the BIPM via the common-view GPS comparisons. The TTR-6 GPS receiver of IREE was calibrated against the GPS receiver of the Paris Observatory (OP) by means of the BIPM1 (SERCEL) portable receiver during a BIPM mission to Prague. As a result, the differential time correction of -12 ns (±1.5 ns) was added to the relation *UTC(TP)-UTC(OP)* [4]. Later on, W.Lewandowski of BIPM computed the differential coordinates with respect to the Technical University of Graz, Austria. Based on this computation, the original coordinates of the IREE antenna were corrected by adding
Dlat = -0.086", *Dlon* = -0.032", *DH* = +1.44 m.
The present coordinates read as follows
Latitude = 50°07'51.114"N,
Longitude = 14°27'10.968"E,
Height = 368.44 m.

IREE routinely checks its UTC(TP) against UTC(PTB) and and UTC(IEN) on a weekly basis. Thus, for example, in the last five weeks (MJD 49713 to 49747) the classical standard deviation computed from 27 daily tracks against PTB and from 24 daily tracks against IEN gives, after removing systematic drifts, 2.3 ns for *UTC(TP)-UTC(PTB)* and 1.9 ns for *UTC(TP)-UTC(IEN)*.

Using the same approach as in [5], i.e. treating the tracks as if they were equally spaced, we get the frequency stability for 24 track mid-points as depicted in Fig.8. The baselines corresponding to *TP-IEN* and *TP-PTB* are 750km and 350 km, respectively.

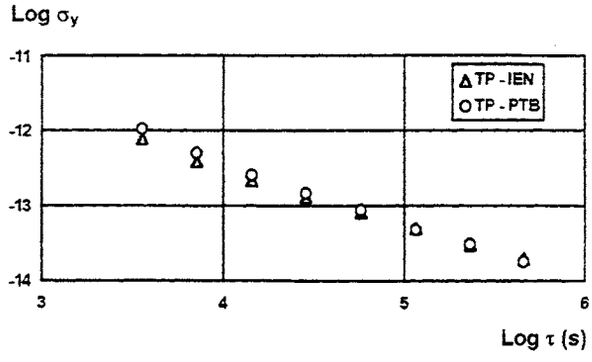


Fig.8 Standard Allan deviation of GPS common-view comparisons

Note that *TP-IEN* exhibits slightly less phase noise than *TP-PTB* though the baseline of *TP-IEN* is longer than *TP-PTB*.

A better picture of this case is provided again through $\sigma_x(\tau)$ which indicates the presence of white noise frequency modulation already for intervals larger than about four hours.

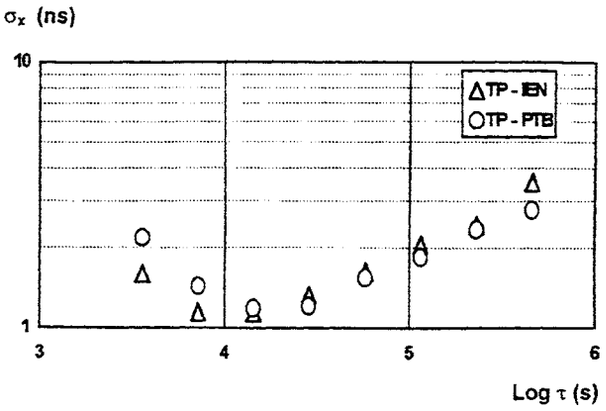


Fig.9. $\sigma_x(\tau)$ of the GPS common-view comparisons of UTC(TP)

5. Conclusions

This report has demonstrated the „state-of-the-art" of the Czech Time and Frequency Group Standard regarding its metrological parameters and both the internal and external traceability. While the external traceability is the standard case, i.e. it is given by the local performance of the GPS common-view time transfer, the internal traceability is determined by the fiber optic link that ensures the interconnection between the two sites, IREE and ITE, where the clocks are located.

Simultaneous phase-difference measurements at both sites have shown that the fiber optic link with the electronics, originally intended to meet less stringent demands of telecommunication than those of time and frequency metrology, exhibits group-delay variations which are small enough so as not to substantially degrade stability of the clocks reported to the BIPM. In addition, these variations predominantly due to the temperature sensitivity of the fiber itself, can be compensated for by using temperature coefficients.

Regarding short-term instability, the figures are not as good as one may wish and perhaps some of the link elements will have to be modified. The fact is that in the first stage the short-term stability has not been of primary importance. Also, no attempt has been made so far to transfer absolute time from IREE to ITE via the fiber optic link, at least at a level corresponding to the link potentials. This should be the subject of further work.

References

[1] Buzek O.,Čermák J., Šimša J.: Project of frequency reference system for synchronization of the SPT Prague Digital Communication Network (in Czech), IREE Report Z1923/A, Prague, Feb.1992

[2] Primas L., Lutes G., Sydnor R.: Fiber optic frequency transfer link, 42nd Annual Frequency Control Symposium, 1988, pp.478-484

[3] Allan D.W.: Time and frequency metrology: Current status and future considerations, 5th EFTF, Besançon, March 1991

[4] Lewandowski W. : Determination of Differential Time Correction between the GPS Time Receivers located at the Observatoire de Paris and the Ustav Radiotechniky a Elektroniky CSAV at Praha, BIPM Report WL/TM.004, Sept. 1992

[5] Thomas C., Wolf P., Tavella P.: Time scales, BIPM, Monographie 94/1, pp.21-23

TEST RESULTS OF KVARZ CH1-75 ACTIVE AND CH1-76 PASSIVE HYDROGEN MASERS

Edward D. Powers, Edward C. Jones, and Fredrick Danzy,
Naval Research Laboratory, Washington, D.C.20375-5320, USA

William M. Golding and Alick H. Frank
SFA, Inc.

ABSTRACT

The Institute of Electronic Measurements, KVARZ, in Russia has developed an active and passive maser for the market place. Ball Efratom of Irvine, CA is the distributor of these masers and the Naval Research Laboratory received one of each for testing.

The long-term drift characteristic of the CH1-75 active maser was measured in an environmental chamber whereas the CH-76 passive maser was measured in a room environment. Both masers were measured against the Naval Research Laboratory maser N1. The active maser was measured for 163 days and had a drift rate of less than 1 part in 10E-14/day with a temperature coefficient of less than 5 parts in 10E-15/C. The passive maser was measured continuously for 163 days (1.4x10E+7 sec.) and had a measured frequency drift of less than 3 parts in 10E-16/day; and a temperature coefficient was determined from room temperature changes to be less than 2 parts in 10E-14. Phase noise and Allan Deviation will be presented.

INTRODUCTION

The Naval Research Laboratory (NRL) has developed a test facility for the performance and environmental evaluation of atomic standards for spacecraft and ground applications[1,2]. The primary references for the dual mixer clock comparison systems are two Sigma Tau active hydrogen masers, designated N1 and N2. These reference standards have been evaluated against the master reference standards at the U.S. Naval Observatory to better than 3 parts in 10E-16 per day, and the dual mixer measurements systems has a measurement resolution well below that required for intercomparing these units.

The evaluation of the CH1-75 and CH1-76 hydrogen masers was conducted with the cooperation of Ball Efratom in both providing these units for an extended period of time and in testing support.

TEST CONFIGURATION

The basic block diagram of the measurement setup is shown in Figure 1. The isolation distribution amplifier is a multichannel device that has 125db isolation between channels. Each channel of the isolation amplifier has 4 outputs; and each output is also isolated by 125db from each other. The outputs of the isolation distribution amplifier go into the multichannel short-term and long-term dual mixers systems which provide high precision phase measurement. The phase information is then processed to determine frequency drift, environmental effects, such as temperature, stability, and the Allan Deviation. Least squared function fits to the data to determine and remove systematic behaviors, and the calculation of frequency stability, Allan Deviation, was performed by using software called stable[3].

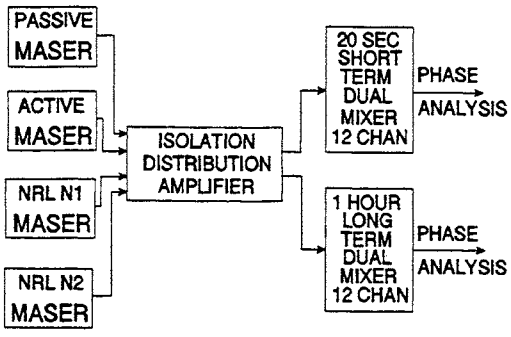


Figure 1. Block configuration of the NRL measurement system.

RESULTS OF THE CH1-75 ACTIVE HYDROGEN MASER

Data from the CH1-75 active maser, with the auto tune mode turned off, were taken for 162 days with measurements being made once an hour as shown in Figure 2. During the entire measurement period of 163 days the CH1-75 was in a constant temperature environment. Variations were measured also at hourly intervals. A least squares quadratic fit was made to the phase data. The minimum time between measurements on the long term system are the hourly data points which were used in this effort. The coefficients for the data fits and others are contained in Table 1. Figure 3 is a plot of the residuals after the quadratic fit has been subtracted. The frequency derived from the phase data of the active maser is shown in Figure 4, and a linear least squares fit is made to the data. The amount of drift is reasonably constant so that a significant improvement can be made in removing the drift. As before, the least squares fit is subtracted from Figure 4 and the residuals plotted in figure 5. A quadratic fit was made to the frequency in Figure 4 (not shown) and the slope of the fit is plotted on Figure 6. This shows the rate of change of the frequency.

Figure 7 is a plot of the Allan Deviation with and without the drift removed over the period of 162 days. Removing drift improved Allan Deviation about one order of magnitude at 162 days. Over the entire time the maser was observed, it was well-behaved and exhibited no out put problems.

The temperature of the active maser's environment was systematically changed for a period of about a week during the test period. The temperature response measured are shown in Figure 8. The results are not a smoothly-varying function so that a function fit is not feasible. Also note that there is an hystersis as well as plus and minus temperature coefficients in the ascending and descending temperature data. From the data, a temperature coefficient in the order of +/- 1 part in 10E-14/°C be expected.

RESULTS OF THE CH1-76 PASSIVE MASER

Data was taken continuously for 162 days test period on the passive maser in a normal room environment. The temperature variations were as much as 5 °C. Figure 9 shows the phase data with a quadratic fit, and figure 10 shows the phase data after the quadratic fit ha been subtracted, Figure 11 is the resulting frequency data with a least squares linear fit. The noise fluctuations in the data are about 10 times greater than those for the active maser frequency data. However, the frequency drift is much less than the active maser which is influenced with the fact that the active maser was not operated in the auto tune mode. Some of the changes in frequency were be correlated with changes in room temperature giving an approximate temperature coefficient of less than 2 parts in 10E-14/°C. The Allan Deviation shown in Figure 12, gives a value ao about 10E-14 over the entire measurement period. The unusual S shape of the curve is believed to be do to weekly temperature variations of the laboratory environment.

TABLE 1. Coefficients of Least Squares Fits

Figures	a	b	c
2	-8.799e-08	4.983e-10	7.387e-13
4	1.267e-13	4.099e-16	
#4	6.2953e-14	5.0795e-16	-2.511e-20
*6	5.0795e-16	-5.0220e-20	
9	7.076e-06	-9.617e-10	8.640e-15
11	-2.710e-13	5.889e-18	

Quadratic fit to the frequency in Figure 4 (not plotted).
* Slope of the quadratic fit to the frequency in Figure 4.

REFERENCES

1. R.L. Beard, J.A. Murray, and J.D. White, "GPS Clock Technology and Navy PTTI Programs at the Naval Research Laboratory', Proceedings of the Eighteenth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 1986

2. R.L. Beard, F. Danzy, E. Powers, and J.D. White, "Test and Evaluation Methods at the NRL Clock Test Facility", Proceedings of the 43RD Annual Symposium on Frequency Control, 31 May - 2 June 1989

3. Stable, Frequency Stability Analysis, Version 2.73, Hamilton Technical Services,
S. Hamilton, MA 01982 USA.

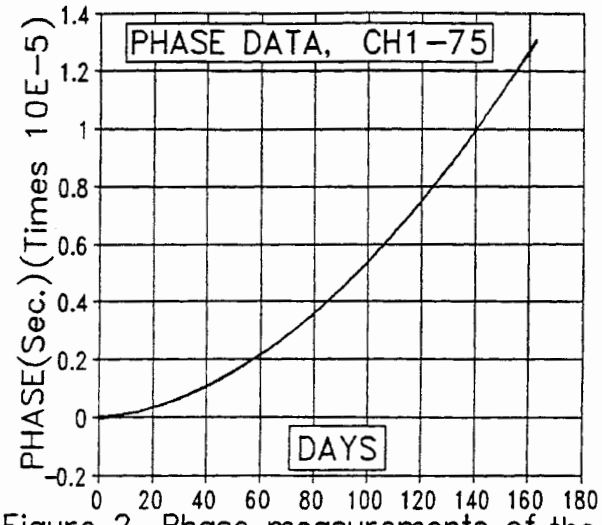


Figure 2. Phase measurements of the active maser with a quadratic fit.

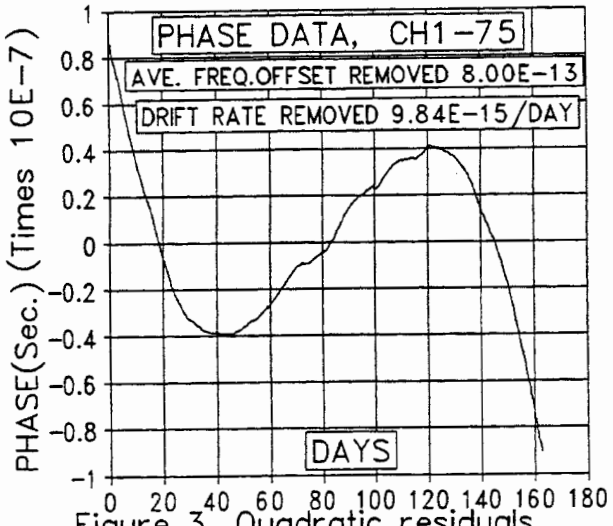


Figure 3. Quadratic residuals phase measurements.

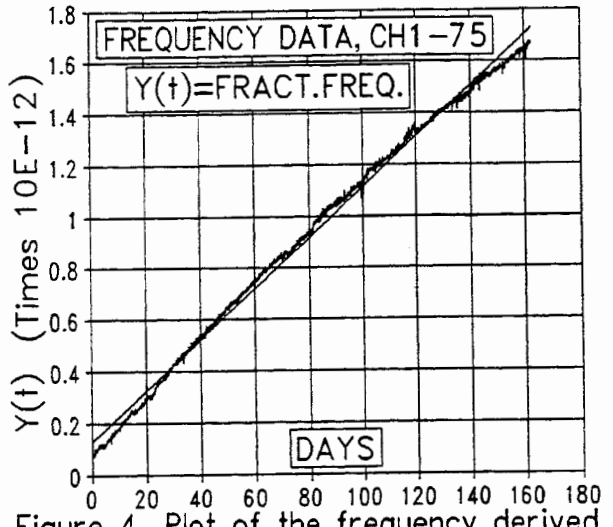


Figure 4. Plot of the frequency derived from the phase data with a linear fit.

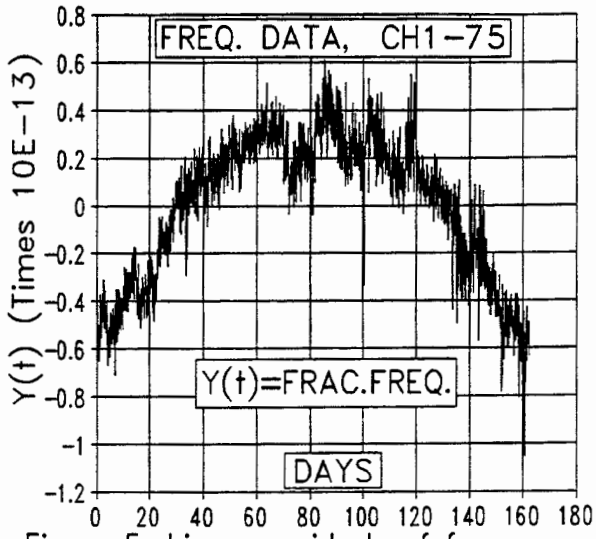


Figure 5. Linear residuals of frequency.

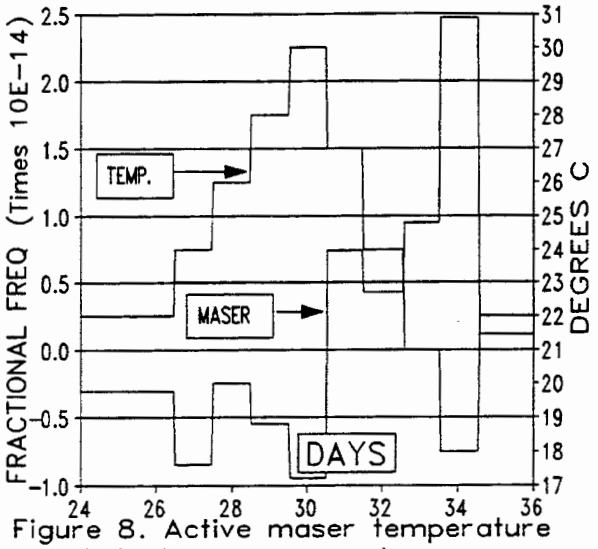


Figure 8. Active maser temperature coefficient measurements.

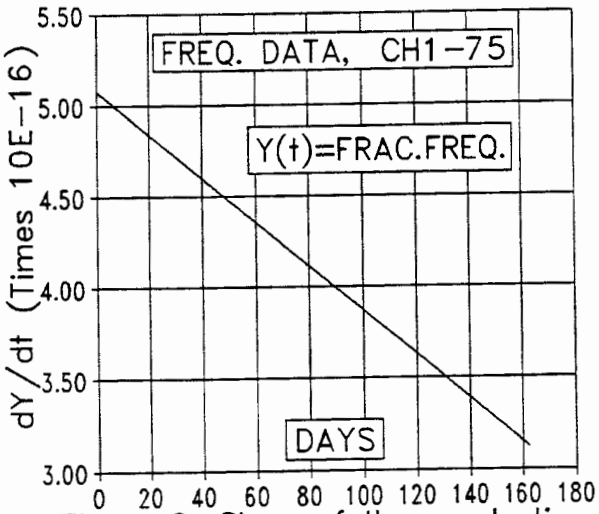


Figure 6. Slope of the quadratic fit to the frequency data.

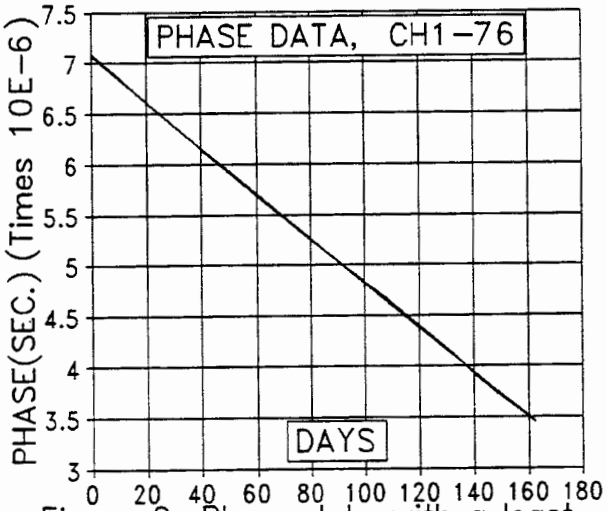


Figure 9. Phase data with a least squares quadratic fit to the data.

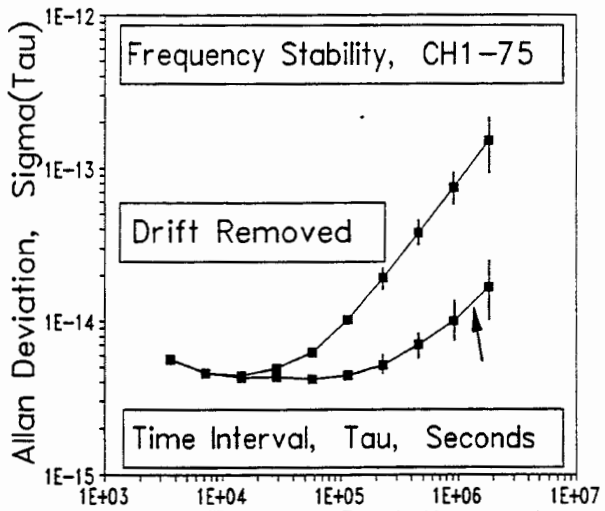


Figure 7. Allan Deviation with and without drift removed.

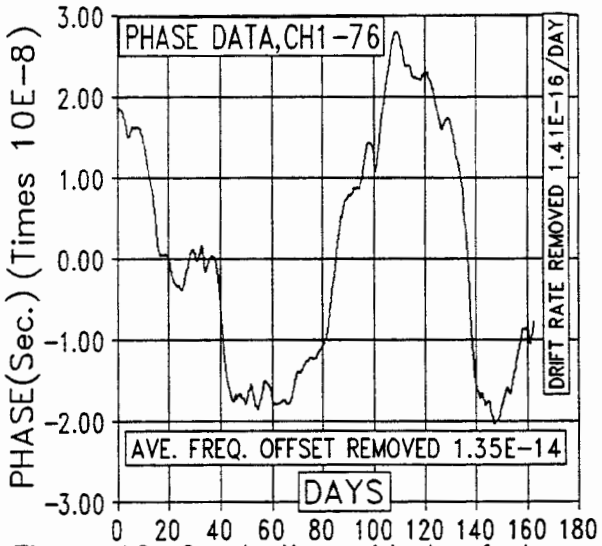


Figure 10. Quadratic residuals of phase.

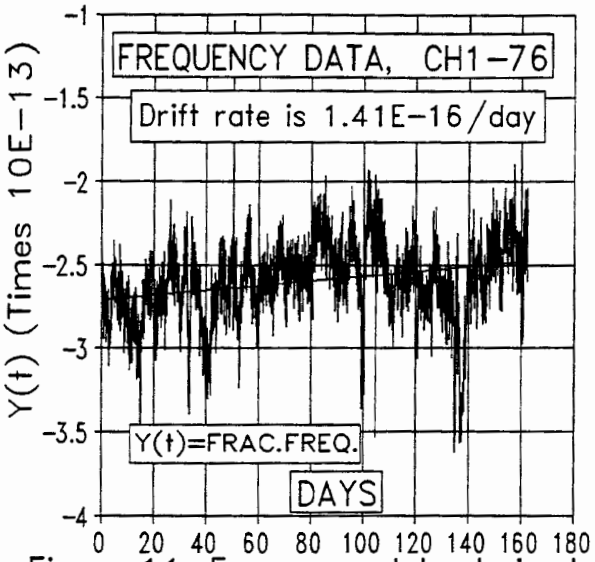


Figure 11. Frequency data derived from the phase data with a linear fit.

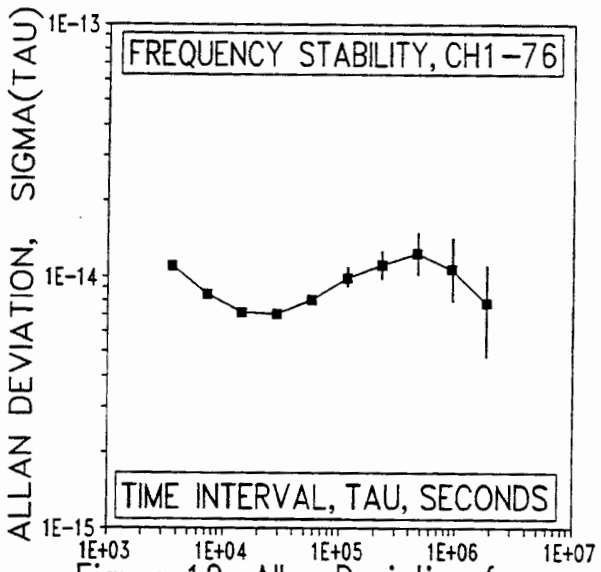


Figure 12. Allan Deviation for the passive maser.

PHASE COHERENT FREQUENCY MEASUREMENT CHAIN TO THE VISIBLE

H. Schnatz, B. Lipphardt, J.Helmcke, G. Kramer, F. Riehle

Physikalisch - Technische Bundesanstalt (PTB) **D-38116 Braunschweig**

ABSTRACT

Recently the Comite' International des Poids et Mesures (CIPM) has recommended the frequency of the Ca- intercombination line for the realization of the metre to be ν= 455 986 240,5 MHz with a standard uncertainty of 200 kHz ($\Delta\nu/\nu = 4{,}5 \cdot 10^{-10}$) [1]. Based on the existing frequency chain at PTB to the infrared [2], we extended the phase coherent frequency measurement chain to a Ca stabilized dye laser. The statistical uncertainty of the first frequency measurement is less than 100 Hz. The preliminary measurement of the frequency of the unperturbed intercombination line of Ca is ν=(455 986 240 492 ± 5) kHz.

I. INTRODUCTION

The Ca intercombination line $^3P_1$-$^1S_0$ was recommended in 1992 by the CIPM for the realization of the metre . Up to now this frequency has been determined by interferometric wavelength intercomparisons [3,4]. We have developed a dye laser stabilized to this transition [5]. To take full advantage of its potential as a precise optical frequency / wavelength standard, this frequency has to be measured with respect to the frequency of the primary standard of time and frequency.

The frequency multiplication steps from the Cs clock to the CO_2- laser radiation at 30 THz have been described earlier [6]. The scheme of the frequency measurement from the P(14)- line of a $^{12}C^{16}O_2$ - laser to the Ca intercombination line is depicted in figure 1. A $^{12}C^{18}O_2$ laser beam at ν= 28 534.3 GHz ($P^*(20)$) is superimposed by means of a beam splitter to the radiation of known frequency of a $^{12}C^{16}O_2$ laser at ν= 28 464.7 GHz (P(14)), which is phase locked to the Cs clock. One output of the beamsplitter is used to mix the difference frequency of the CO_2 lasers and the third harmonic of a dielectric resonator oscillator (DRO) at 23.2 GHz on a W- Ni point- contact diode (MIM).

At the other output the second harmonic of the sum frequency of both CO_2 lasers is compared with the radiation of a color- centre laser (CCL), operating at λ= 2.63 µm, by means of a second MIM- diode.

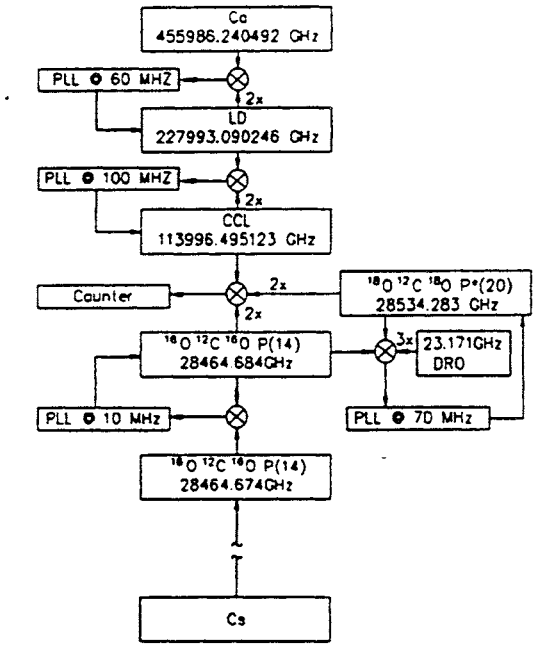


Fig. 1: Extension of PTB's frequency chain to the Ca- intercombination line

A nonlinear crystal ($AgGaS_2$) doubles the frequency of the CCL, and the corresponding second harmonic is compared with the radiation of a laser diode operating at λ= 1.315 µm. The frequency of this laser radiation is again doubled by means of a temperature tuned LiB_3O_5 crystal to reach the frequency of the Ca stabilized dye laser.

II. EXPERIMENTAL RESULTS

A. CO_2 LASERS

The CO_2 lasers used in the experiment are commercially available models, sealed off and grating tuned, with output powers of a few watts. Typically 20 mW from each laser is sufficient to saturate the MIM diodes.

B. MICROWAVE OSCILLATOR

The frequency difference of the two CO_2 lasers is bridged by the third harmonic of a DRO at 23.2 GHz, which itself has is phase locked to the Cs standard.

C. COLOR CENTRE LASER

The CCL is pumped by 3W cw radiation of an intensity stabilized Kr^+ ion laser at λ= 647 nm. The CCL is frequency stabilized by means of an intracavity electro- optical modulator (EOM).

The second harmonic of the CCL is generated in an angle tuned $AgGaS_2$ crystal. A fundamental power of 10 mW produces a second harmonic power of 1 nW in single path.

D. DIODE LASER AT λ= 1.315 µm

In the next step a DFB laser diode (LD) operating at λ= 1.315 µm is employed as a phase- locked local oscillator. The pre-stabilization technique of the LD and the cavity- enhanced generation of the second harmonic radiation in a LiB_3O_5 crystal have been described previously [7,8].

A fundamental power of 1.5 mW yields up to 70 nW of second harmonic light, which is sufficient for our applications.

E. CALCIUM FREQUENCY STANDARD

A high resolution dye laser spectrometer [9] has been stabilized to an effusive thermal Ca atomic beam using optical Ramsey resonances in a four running- wave configuration [5]. For stabilizing the laser frequency to the central fringe of either recoil component, the laser frequency was square wave modulated (0.5 Hz) and a first harmonic detection scheme was used.

F. DETECTION OF THE BEAT FREQUENCIES AND PHASE LOCKING

The beat note between the radiation of the Ca stabilized dye laser and the second harmonic radiation of the laser diode is detected by means of a Si avalanche diode. The beat signal at 60 MHz is tracked and its frequency is phase locked to a stable reference by a conventional phase- locking technique which uses the fast current control input of the laser diode, thus locking the LD to the Ca stabilized dye laser.

The radiation of the LD is superimposed collinearly on the second harmonic light of the CCL and focused on a small area InGaAs diode. Again, the beat signal at 100 MHz is tracked and locked to a reference. The control signal is fed to the input of a high voltage amplifier driving the intracavity EOM of the CCL.

The beat note between the CO_2 lasers and the third harmonic of the DRO has been detected with a signal to noise ratio of 35 dB in a bandwidth of 100 kHz, which is sufficient to assure phase locking of the $P(20)^*$-laser to the P(14)- laser.

The part of the frequency chain from the Cs clock to the P(14)- CO_2 laser is located in a different building which is about 60 m away from the laboratories of the Ca frequency standard and the high frequency part of the chain. Due to this we had to employ an additional CO_2 laser operating at the P(14) line. In an underground tunnel, optical systems with servo controlled mirrors and telescopes guide the radiation of the CO_2 laser to the other building. The phase coherent optical linkup between the two parts of the chain is performed with a signal to noise ratio of more than 50 dB.

In the final step of the frequency measurement experiment the second harmonic of the sum frequency of the two CO_2 lasers is compared with the radiation of a color centre laser (CCL), operating at $\lambda = 2.63$ µm, by means of a MIM-diode.

With powers of approximately 20 mW (P(20)*), 15 mW (P(14)) and 2 mW (CCL) the beat note at ν=1438 MHz could be obtained with a S/N- ratio of 15 dB in a bandwidth of 100 kHz. This beat signal was filtered, down converted to 100 MHz by means of a frequency synthesizer, tracked with a phase-locked electronic tracking oscillator and finally counted in a totalizing counter.

G. PRELIMINARY RESULTS OF THE FREQUENCY MEASUREMENT

All intermediate frequency steps from 30 THz to 455 THz have been set up and we have performed a preliminary frequency measurement of the Ca intercombination line. During the experiment a dye laser was successively stabilized to the two recoil components of the Ca intercombination line by means of Ramsey excitation of a thermal Ca atomic beam.

Figure 2a shows the frequency data of the measurement of the low frequency recoil component during a periode of approximately 500 seconds. All data coincide within the range of ± 1kHz. Figure 2b gives the distribution of the data in Fig.2a . The data are in good agreement with the fitted gaussian probability distribution.

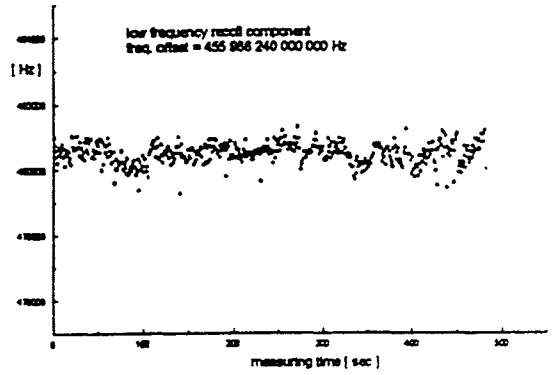


Fig. 2a: Frequency data of low frequency recoil component

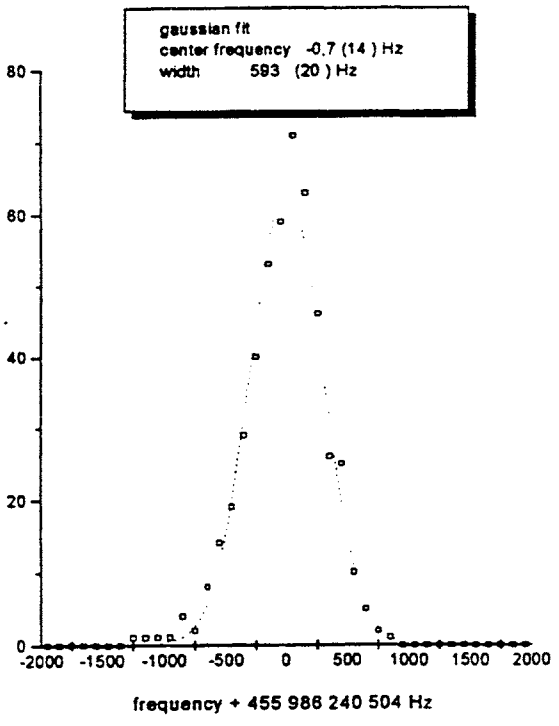


Fig. 2b: Frequency distribution of data of fig.2a and gaussian fit.

The statistical uncertainty (1σ) of the frequency measurement is less than 100 Hz. The preliminary overall uncertainty of the frequency of the

unperturbed Ca in this experiment was limited by residual first order Doppler effect of the thermal atomic beam to 5 kHz and the obtained frequency value of the unperturbed Ca transition is ν= 455 986 240 492 (5) kHz. The frequency measurement chain will be used to measure the transition frequency of Ca atoms stored in a magneto optical trap.

References

[1] T. Quinn
"Mise en Pratique of the Definition of the Metre (1992)", Metrologia, 30, 523-541, (1994)

[2] G. Kramer, B. Lipphardt, C. O. Weiss
"Coherent frequency synthesis in the infrared", Proceeding of the 1992 IEEE Frequency Control Symposium,pp. 39-43,Hershey, IEEE Catalog No. 92CH3083-3

[3] G. Bönsch, A. Nicolaus, U.Brand
"Wellenlängenbestimmung der Ca- Interkombinationslinie mit dem Michelson- Interferometer der PTB"
PTB Mitteilungen, 99, 329-334, (1989)

[4] J. Ishikawa, N. Ito, A. Morinaga
"Accurate Wavelength Measurement of the Intercombination Line of Calcium"
J. Appl. Phys. B, 33, 1652-1654, (1994)

[5] A. Morinaga, F. Riehle, J. Ishikawa, and J. Helmcke
"A Ca Optical Frequency Standard: Frequency Stabilization by Means of Nonlinear Ramsey Resonances ", Appl. Phys. B, 48, 165-171, (1989)

[6] C. O. Weiss
"Frequency measurement chain to 30 Thz using FIR Schottky diodes and a submillimeter backward wave oscillator", Appl. Phys. B, 34, 63-67, (1984)

[7] S. Ohshima and H. Schnatz
"Optimization of injection current and feedback phase of an optically self- locked laser diode ", J. Appl. Phys. 71, 3114- 3117, (1992)

[8] H.Schnatz, S. Ohshima, H.R. Telle, F. Riehle
"Extension of the PTB frequency chain towards the Ca intercombination line at 657 nm ", IEEE Trans. Instr. Meas. IM-42, 273-275, (1993)

[9] J. Helmcke, A. Morinaga, J. Ishikawa, F. Riehle
"Optical Frequency Standards ", IEEE Trans. Instr. Meas.38, 524 - 532, (1989)

MATERIALS

HYDROXYL DEFECTS IN GERMANIUM-DOPED QUARTZ: DEFECT DYNAMICS AND RADIATION EFFECTS

Harish Bahadur

National Physical Laboratory
Hilliside Road, New Delhi-110012, INDIA.

ABSTRACT

Irradiation effects on Ge, Al and other H-related point defects in quartz were investigated. Irradiation at 77 K in prior 300 K-irradiated Ge-doped samples showed a novel defect center absorbing at 3400 cm^{-1} representing a possible configuration of $[GeO_4^{4+}e^-Li^+e^-H^+]^0$ center.

INTRODUCTION

Hydrogen is an ubiquitous impurity in quartz and affects nearly every application of it in numerous precision aerospace precision equipments. Additionally, in all quartz, natural and most of the commercially available cultured material, Al and Ge are always present in varying proportions. While Al^{3+} in quartz acts as an acceptor atom and needs charge compensation, Ge^{4+} exists as a neutral atom. When exposed to ionizing radiation, the electrons and holes produced move through the crystal until they recombine or are trapped by these impurities. While the electrons are trapped at Ge sites, the holes are captured by Al atoms forming Al-hole centers. A variety of Al centers exist in quartz and have been extensively studied[1-3]. On the other hand, the Ge-related point defects in quartz relatively less studied. Figure 1 illustrates three of the commonly known Ge defect models. The upper diagram represents the neutral substitutional Ge atom and is labeled as the $[GeO_4]^0$ center. The defect center in the middle diagram has an extra electron trapped at a substitutional Ge and is labeled as $[GeO_4]^-$. The lower diagram has both an alkali interstial ion,-either Li^+ or Na^+, and the extra electron trapped at the substitutional Ge siteand is labeled as $[GeO_4/M^+]^0$ center. Mackey[4] obtained the fundamental $[GeO_4]^-$, alias Ge(I) by low-temperature x-irradiation of Ge-doped quartz. Isoya, Weil and Claridge[5] found that in addition to Ge(I) centers, another thermally excited-state configuration Ge(II) becomes observable upon warming the sample. Weil[6] has given a complete discussion of paramagnetic Ge and Al related centers in quartz.

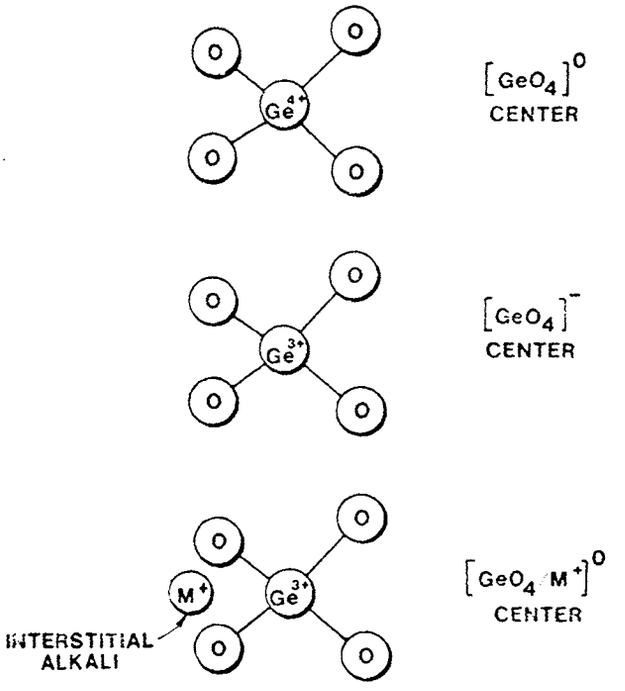


Figure 1: Models of germanium centers in crystalline quartz.

While the contribution of Al-centers on the radiation response on quartz crystals has been extensively studied[3], the Ge effects are relatively less studied. Euler and Kahan[7] reported the production of a radiation-induced anelastic relaxation at 246 K in Ge doped cultured quartz. Martin et.al[8] observed this peak in two different lightly irradiated natural quartz and not in oscillator crystals made from conventionally grown quartz. Euler and Kahan[7] have shown that Ge(C)-Li center is responsible for the 246 K peak.

A variety of hydroxyl defects are known to occur in both natural as well as cultured quartz[2,9]. The OH^- mplecules are trapped at numerous unidentified sites in quartz lattice. Most of the defects have their stretching frequencies in the near infrared region and thus are easily detectable by infrared absorption measurements. Much of the earlier work on OH^- -related defect centers in crystalline quartz has been summarized by Bahadur[9]. These hydroxyl defects serve as a source of protons to compensate the electron-excess defects of aluminum centers during irradiation. Of these numerous OH^- defect centers in quartz, a model is known only for the 3367 and 3306 cm^{-1} bands representing the long- and short-bond Al-OH^- centers[2,9,10]. Conventionally grown cultured quartz always has a set of four growth-defect bands positioned at 3350, 3400, 3440 and 3581 cm^{-1}. Martin et al[8] have shown that these growth-defect centers can also trap alkalis and produce anelastic relaxation peaks at 305 and 335 K found to be Li and Na related.

Bahadur[11] has recently reported that Ge-doping in quartz does not affect the production of the conventional growth defect bands. Peaks at 3481 and 3656 cm^{-1} were related to the presence of Ge incorporated in quartz material[11,12]. Limited information was presented on irradiation effects at 77 K and 300 K. Near infrared bands at 3506 and 3525 cm^{-1} were noticed to get produced at 77 K- and 300K-irradiations respectively[12].

In this paper we present some new results on Ge-doped cultured quartz crystals irradiated at 77 K before and after 300 K-irradiation and compare he results with the conventionally grown Sawyer Premium Q material. It is expected that these studies will be directly useful in evaluating the radiation response of Ge-containing quartz especially relating to hydroxyl defects and the associated frequency changes of quartz resonators. Samples were designated as GeQ-1 and GeQ-2. The GeQ-1 quartz was grown at Sawyer Research Products, Eastlake, OH. The intensities of the conventional growth defect bands of this sample were comparable to the Sawyer Premium Q quartz. Thus, it was a slowly grown material. Sample GeQ-2 was grown at Bell Laboratories, Murray Hill, NJ and was a faster grown material.

EXPERIMENTAL DETAIL

Near infrared absorption measurements were done on a Beckmann 4240 Beckman double beam spectrophotometer. Optical samples were prepared from the Z-growth regions and attached to the cold finger of a cryostat capable of reaching 77 K within 2 deg. All spectra were recorded for E ⊥c. Irradiation was done using a Van de Graaff electrostatic accelerator at Oklahoma State University. During irradiation, the samples were not demounted but remained in the Dewar itself. A detailed description of the measurement and irradiation techniques has been given earlier[9].

RESULTS AND DISCUSSION

Fig.2 shows the spectra of sample GeQ-2 in its as received state and after irradiation at 77 K and 300 K respectively. Comparison of the spectra of as received crystal to those of irradiated at 77 K and 300 K shows that the low temperature irradiation produces a band at 3506 cm^{-1} along with the suppression of other bands. The 300 K irradiation,[12] in addition to the 3367 and 3306 cm^{-1} bands which represent long- and short-bond Al-OH^- centers respectively, produced a band at 3525 cm^{-1} which represents a possible configuration of $[GeO_4\ e^-/Li^+]^0$.

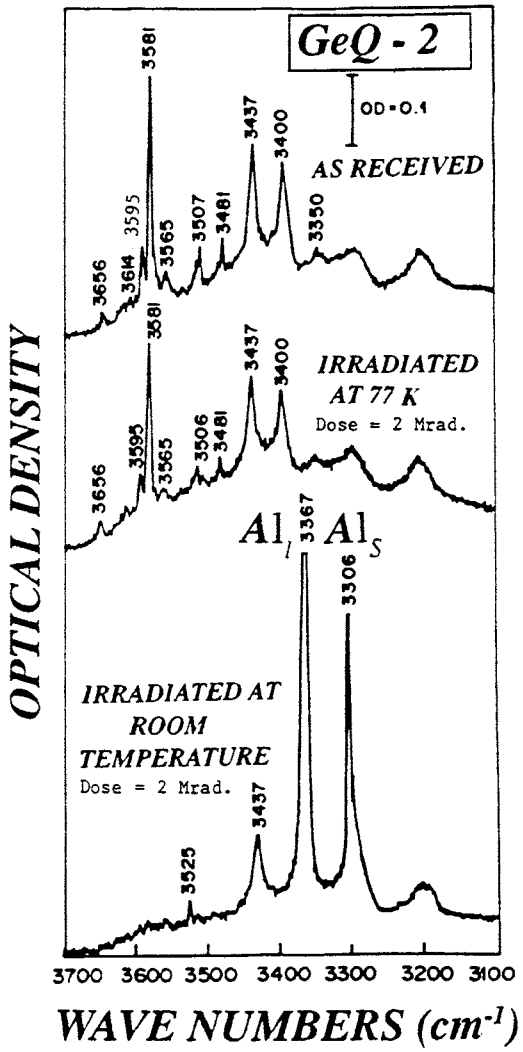


Figure 2: Near infrared absorption spectrum of a Ge-doped sample (GeQ-2) of crystalline quartz in as-received, and after irradiation at 77 K and 300 K.

Subsequent irradiation at 77 K, Fig.3, shows that a band at 3400 cm^{-1} was produced in both the crystals and the Al-OH^- bands behaved like those either in the conventionally grown cultured crystals[2,9] or natural crystals[9]. The new band at 3400 cm^{-1} represents a different species of defect center than one of the conventional growth defect bands absorbing at this frequency. Production curve of this band showed a linear response versus the accumulated doses at the final low-temperature irradiation. 300 K-irradiation was a necessary step to produce this band at the subsequent low-temperature irradiation. Both the samples GeQ-1 and GeQ-2 showed qualitatively similar results on the irradiation sequence adopted but the radiation induced 3400 cm^{-1} band was much prominent in the faster grown material, sample GeQ-2. It must be pointed out here that the Sawyer Premium Q quartz investigated did not produce any of the 3506, 3525 and the 3400 cm^{-1} bands. This shows that these bands are Ge-related hydroxyl defects and produced as a result of irradiation. Ge-doping in cultured quartz did not affect the irradiation characteristics of Al-related hydroxyl defects. These bands showed, as mentioned above, identical irradiation and also anneal characteristics as found for cultured [13] or natural quartz[8].

Isochronal anneal studies carried out on strongly irradiated such samples showed that the low-temperature irradiation induced 3400 cm^{-1} band decays during 180-200 K region. During the temperature region the conventional growth defect bands started a recovery. This shows an interdependence of Ge-related and the conventional growth defect bands.

Turning our attention back to the radiation induced 3400 cm^{-1} band , it is clear that the band is Ge-related and the 300 K-irradiation is a necessary step to produce this band. Therefore, this band is a result of irradiation cycle carried out on Ge-doped quartz. The first irradiation at 77 K would give rise to $[GeO_4]^-$ center because of trapping of electrons at Ge sites[4,6]. The 300 K-irradiation would make an interstitial alkali move to Ge site after getting released from the Al centers. The subsequent irradiation at 77 K will again be effective in having an electron trapped at Ge center. To preserve the charge neutrality a proton will also come nearby. This proton with neighboring oxygen would make an OH^- bond which would oscillate under the applied infrared irradiation. It appears that this is the most likely configuration of the defect responsible for the absorption band at 3400 cm^{-1} produced upon 77 K-irradiation in prior 300 K-irradiated samples From this discussion, a Ge-Li-H complex of the form $[GeO_4 e^- Li^+ e^- H^+]^0$ shown in Figure 4 appears to represent the radiation induced near infrared band at 3400 cm^{-1}.

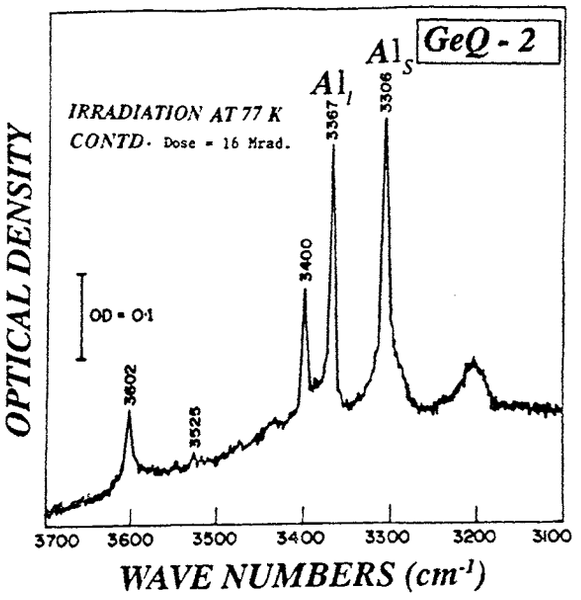


Figure 3: Near infrared absorption spectrum of the quartz sample GeQ-2 upon subsequent irradiation at 77 K after the crystal was prior irradiated at 300 K. Accumulated radiation dose at the final 77 K-irradiation was 16 Mrad.

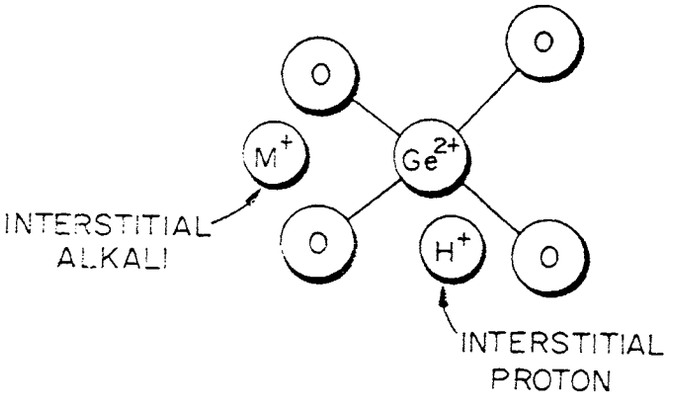


Figure 4: A proposed model for the radiation induced 3400 cm^{-1} band in Ge-doped quartz.

SUMMARY

The present work reports new results on irradiation characteristics of hydroxyl defects in Ge-doped quartz crystals. The analysis presented is expected to evaluate the quantitative radiation response for their application in satellite-borne frequency standards.

REFERENCES

1. J.A. Weil, Radiation Effects, 26, 261 (1975).

2. L.E. Halliburton, J.J. Martin and D.R. Koehler, in *Precison Frequency Control*, edited by E.A. Gerber and A.D. Ballato (Academic, Orlando, FL, 1985), 1, pp. 1-46.

3. J.J. Martin, J. Appl. Phys. 68, 5095 (1990).

4. J.H. Mackey, Jr, Chem. Phys. 39, 74 (1963).

5. J. Isoya, J.A. Weil and R.F.C. Claridge, J. Chem. Phys. 69,4876 (1978)

6. J.A. Weil, Phys. Chem. Miner. 10, 149, (1984).

7. F. Euler and A. Kahan, Phys. Rev.B 35, 4351 (1987).

8. J.J. Martin, H.B. Hwang, H. Bahadur and G.A. Berman, J. Appl. Phys. 65, 4666 (1989).

9. Harish Bahadur, J. Appl. Phys. 66 4973 (1989).

10. Harish Bahadur, J. Appl. Phys. 73, 7790 (1993)

11. Harish Bahadur, J. Appl. Phys. 75, 4741 (1994).

12. Harish Bahadur, Phys. Rev.B 50, 717 (1994).

13. W.A. Sibley, J.J. Martin, M.C. Wintersgill and J.D. Brown, J.Appl. Phys. 50, 5449 (1979).

A CONTINUUM THEORY OF PIEZOELECTRIC AND FERROELECTRIC DIELECTRICS INCLUDING SINGULAR SURFACES AND INTERFACIAL PROPERTIES

N. Daher

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

Abstract

Various phenomenological theories for the interaction of the electromagnetic fields with deformable bodies are presently available. However, at a certain level of complexity, particularly when interfacial properties are taken into account, these theories cease to apply. In recent works, the author placed in evidence different ambiguities and difficulties encountered in the classical methods (degeneracy of the vectorial approach in dealing with interfacial electromagnetism, difficulty associated with distribution products when energy methods are considered). A new energy formulation has been proposed and applied to simple situations in relativistic and nonrelativistic frameworks. This formulation heavily lies on the use of invariance principles (gauge invariance for electromagnetism, rotational invariance for mechanics and scale-change invariance for the interaction of the bulk with the interface). In the present work, after a brief recall of the basic ideas relative to the new procedure, a gauge and rotationally invariant, thermodynamically admissible nonlinear theory of deformable piezoelectric and ferroelectric dielectrics including singular surfaces and interfacial properties is presented. Such a theory is necessary for a consistent study of volume and surface acoustical, optical and acousto-optical wave modes which can propagate in such materials. A special attention is devoted to surface equations and particularly those which could not be obtained by previous formulations.

Introduction

The necessity for the introduction of interfacial physical properties in mechanics [1, 2] electromagnetism [3] thermodynamics [1, 2] and their different couplings [4] is now a well established fact. However, although this idea, that showed to be efficient in a variety of local circumstances, is conceptually simple, its mathematical development is not an easy task. How to introduce surface magnetic field and induction through the postulational well-known vectorial approach ? How to avoid multiplying two distributions or discontinuous functions which is not a well defined mathematical operation when dealing with an energy like formulation ? These mathematical difficulties led to the use of a number of tricks and procedures obtaining only partial solutions [2, 4, 5]. Recently, the author proposed a global solution applying to mechanics, electromagnetism, thermodynamics and their couplings, through the use of an energy formulation [6]. The latter is governed by three basic invariance principles, two of which are local and related to mechanics and electromagnetism (rotational invariance also called objectivity requirement and gauge invariance). The third is global and accounts for the presence of discontinuities and interfaces (scale-change invariance).

This formulation follows the same line of thought as the one given by Germain [7] in the volume and in a purely mechanical context and by Maugin and co-workers [4, 8] for electromagnetomechanical couplings. In this framework, the different interactions present in the energy formulation called principle of virtual power or d'Alembert principle (not to confuse with the Lagrangian approach) are governed by the principle of objectivity or rotational invariance.

The extended formulation based on three invariance requirements , two of which are new (gauge and scale-change invariance principles) allows the inclusion of Maxwell equations as well as general mechanical, electromagnetic and coupled interfacial properties in a unified mode of expression. Although some of the basic ideas take their origin in the works of Germain, Maugin and co-workers, innovations are performed in two directions : interfaces and electromagnetism. The one associated with interfacial properties and governed by the scale-change invariance principle may be regarded as a natural generalization of the above mentioned previous works. Besides, let us note that this problem was evoked by Germain about twenty years ago in Ref. [7] and a partial solution can be found in different papers published about ten years ago byMaugin and the present author [2, 4, 5]. As to the second direction which is related to the inclusion of Maxwell equations, it was somehow imposed on the author by the difficulties encountered in the derivation of interfacial electromagnetism through the vectorial approach [6]. These two different extensions led to an aesthetically pleasing and efficient method that deals with mechanics and electromagnetism on an equal footing and that remains sufficiently general for the study of nonintegrable systems. The last requirement is essential for the treatment of dissipative phenomena and irreversible processes which play a major role in technical applications. Because of the inclusion of electromagnetism, it seemed to the author more adequate to speak of a duality between geometrical and dynamical features rather than between kinematical and dynamical ones as usual. Moreover, in order to distinguish this formulation from the Langrangian approach (integrable systems) and from d'Alembert's principle which is of a purely mechanical origin, the extended formulation was called : "Leibniz - d'Alembert formulation". Indeed, Leibniz has much contributed to the development of energy conservation and invariance principles (see Refs. [6, 9] for details).

A - Coupled continua including interfacial properties

In the forthcoming developments, a special attention is devoted to the construction of two types of coupling effects. These have been evoked previously [6] without sufficient details and in a rather particular context. The first of these is relative to the objectivity invariance principle and concerns the interactions between different continua in the bulk as well as at singular surfaces and lines. The second is associated with the scale-change invariance principle and concerns the passage from one scale to another (volume → surface, surface → line) independently of any particular consideration as to the type of physical interactions.

Volume-volume, surface-surface, line-line couplings

The procedure to follow extends to interfaces the method given in Refs [7, 8] through the so-called first gradient theory. According to the problem in consideration, one

starts by defining the geometrical features (kinematical in pure mechanics) denoted by $g_i^{K\alpha}$. The indices i, K and α denote respectively, the vector components $i = (1, 2, 3)$, the space dimension $K = (3, 2, 1) = (V, S, L) = (volume, surface, line)$ and the nature of the phenomenon $\alpha = (1, 2, 3, ...) = (ferroelectricity, ferromagnetism, nematic state, ...)$. In the framework of a first gradient theory, a power may be defined in a compact manner as follows

$$P = \sum_K \sum_\alpha \int_{S^K} \left(D_i^{K\alpha} G_i^{K\alpha} + D_{ij}^{K\alpha} \nabla_j^K G_i^{K\alpha} \right) d^K x \quad (1)$$

Finite fields $G_i^{K\alpha} = d^K g_i^{K\alpha} / dt$ replace here the infinetisimal fields given in Refs. [6, 9] ($\delta^K g_i^{K\alpha}$)

K-space and time derivatives (∇_i^K, d^K/dt) extend the volume derivatives to account for surface and line inhomogeneous and nonstationnary physical processes. They are defined as follows

$$\frac{d^K}{dt} = \frac{\partial}{\partial t} + v^K . \nabla \; ; \; \nabla_i^K = \delta_{ij}^K \nabla_j \quad (2)$$

with

$$\delta_{ij}^3 = \delta_{ij} \; , \; \delta_{ij}^2 = \delta_{ij} - n_i n_j \; , \; \delta_{ij}^1 = \delta_{ij} - n_i n_j - b_i b_j \quad (3)$$

and where δ_{ij}, $\boldsymbol{n}$, $\boldsymbol{b}$ and $\boldsymbol{v}^K$ denote respectively the Kronecker tensor, the normal to the singular surface, its binormal and the velocity of "particles" relative to the different spaces (volume, surface and line velocities).

The dynamical fields $D_i^{K\alpha}$ and $D_{ij}^{K\alpha}$ are introduced by duality to obtain a scalar whose dimension coincides with a Power. The infinitesimal parameters $d^K x$ correspond to elements of volume, surface and line ($K = 3, 2, 1$). In continuum physics one distinguishes between given vectorial and tensorial fields such as the force of gravity or the Maxwell tensor and interaction fields such as the stress tensor (matter-matter interaction). The latter dynamical fields are to be related to kinematical or geometrical ones through constitutive relations. In an energy like approach as the one presented here, the particular structure of the interaction fields is obtained by use of invariance principles. For a good understanding of the subject matter concerning interacting fields in the bulk through the invariance under a rigid body motion (absence of interfacial properties and gauge invariance) one may refer to Maugin's remarkable paper [8]. To fix ideas, let us consider the case of deformable bodies with electromagnetic interactions and including ordering phenomena such as those encountered in the study of ferroelectricity [10], ferromagnetism [8] and nematic states [11]. The basic volume, surface and line elements are the mechanical velocity v_i^K and the nonmechanical contributions associated with the material time derivatives of polarization, magnetization and director denoted by

$$\frac{d^K r_i^{K\beta}}{dt} = R_i^{K\beta} \Leftrightarrow \frac{d^K \pi_i^K}{dt} \; , \; \frac{d^K \mu_i^K}{dt} \; , \; \frac{d^K r_i^K}{dt}$$

In a first K-gradient theory ($K = 3, 2, 1$) associated with 3, 2 and 1-dimensional spaces one may distinguish between the mechanical and nonmechanical features and writes locally

$$p_i^K = \left(f_i^K v_i^K + t_{ij}^K \nabla_j^K v_i^K \right) + \sum_\beta \left(D_i^{K\beta} R_i^{K\beta} + D_{ij}^{K\beta} \nabla_j^K R_i^{K\beta} \right) \quad (4)$$

The interactions between the different continua are subject to the objectivity requirement so that the geometrical admissible fields are the deformation tensor

$$\nabla_{(j} v_{i)}^K = \frac{1}{2} \left(\nabla_j v_i^K + \nabla_i v_j^K \right)$$

and the Jaumann derivative $R_i^{K\beta} - \Omega^K_{i\ell} r_\ell^{K\beta}$ and its gradient where $\Omega^K_{i\ell}$ is the rotation tensor subject to $\Omega^K_{i\ell} = - \Omega^K_{\ell i}$ and $\nabla_k \Omega^K_{i\ell} = 0$ (see Ref. [8] for details). After introducing dual dynamical fields with respect to the above mentioned objective fields, an identification procedure with eq. (4) leads to the following restrictions on the different dynamical fields

$$f_i^K = 0 \; , \; \sigma_{[ip]}^K \equiv t_{[ij}^K \delta_{jp]}^K + \sum_\beta \left(D_{[i}^{K\beta} r_{p]}^{K\beta} + D_{[i\ell}^{K\beta} \nabla_\ell^K r_{p]}^{K\beta} \right) = 0 \quad (5)$$

where we have set

$$A_{[ij]} = \frac{1}{2} \left(A_{ij} - A_{ji} \right) \quad (6)$$

In the particular case of deformable ferroelectric dielectrics, the basic elements reduce to the mechanical and to the polarization continua. Thus eq. (6) may be written as

$$\sigma_{[ip]}^K \equiv t_{[ij}^K \delta_{jp]}^K + {}^L E_{[ij}^K \delta_{j\ell}^K \nabla_\ell n_{p]}^K - \rho^K {}^L E_{[i}^K n_{p]}^K = 0 \quad (7)$$

This equation valid for volumes, surfaces and lines ($K = 3, 2, 1$) generalizes eq. (57) of Ref. [10].

The parameters t^K_{ij}, ${}^L E_i^K$ and ${}^L E^K_{ij}$, for which constitutive equations must be given, are respectively a stress tensor and local vectorial and tensorial fields. The first of these is nonsymmetric because of the presence of the fields ${}^L E_i^K$ and ${}^L E^K_{ij}$ which account for the polarization and its first gradient. By comparison or analogy with results obtained from the lattice theory of alkali halides, the tensor ${}^L E_{ij}^K$ is refered to as the shell-shell interaction tensor. Before ending this paragraph, let us note that eq. (7) generalizes the one given in Ref. [4] by accounting simultaneously for

a first gradient theory with respect to polarizations and for the presence of nonmechanical interfacial properties (${}^{L}E^{K}_{ij}$, ${}^{L}E_{i}^{K}$) for which constitutive relations should be constructed. In addition, there is also a difference with previous works [2, 4] as to the introduction of interfacial parameters. This may be seen even in a purely mechanical context. On letting ${}^{L}E_{i}^{K}$ and ${}^{L}E^{K}_{ij}$ vanish only the first term of eq. (7) remains. Unlike previous works [2], eq. (7) exhibits the influence of interfacial anisotropy through δ^{K}_{jp} so that Cauchy tensor t^{K}_{ij} is symmetric only in the bulk equation where t^{3}_{ij} coincides with the usual stress tensor of elasticity theory.

Volume-Surface and Surface-Line couplings

In the previous section we dealt with dynamical interactions that occur in the bulk, at singular surfaces and singular lines. Here, the attention is focused on geometrical interactions which appear only if one considers spaces of different dimensions. The simplest example is that of a free boundary where physical properties are defined only in the bulk. In this case, the boundary contribution associated to a term like $\int_{s3} D_{ij}\, \nabla_j\, G_i\, d^3x$ is written as $\int_{s2} R_i\, G_i\, d^2x$. After the application of the usual procedure associated with d'Alembert's principle [7] one deduces $R_i = D_{ij}\, n_j$.

The situation is more complicated when singular surfaces and lines are embedded into the volume and carry their own physical properties. In this case, in addition to the usual external boundary there is an internal "boundary" that may be expressed as the combination of three contributions as follows

$$P_{sc} = -\sum_{K}\sum_{\alpha}\int_{S^{K-1}}\left\{ R_i^{K\alpha+}\, G_i^{K\alpha+} - R_i^{K\alpha-}\, G_i^{K\alpha-} + R_i^{(K-1)\alpha}\, G_i^{(K-1)\alpha}\right\} d^{K-1}x \tag{8}$$

The first two account for the energies that occur at either side of the interface and the third contribution includes purely interfacial properties. One should remark that eq. (8) is an additional interaction power or energy variation which must vanish in the absence of singular surfaces and interfaces. If one uses the language of invariance principles, then it is sufficient to require that eq. (8) remains invariant under the following scale-change transformation

$$\overline{G}_i = G_i + C_i \tag{9}$$

where C_i is any continuous vector $[\![C_i]\!] = C_i^{+} - C_i^{-} = 0$ across the interface, and G_i is one of the three vectors $G_i^{K\alpha+}$, $G_i^{K\alpha-}$ or $G_i^{(K-1)\alpha}$. This requirement leads to the following restriction

$$R_i^{K\alpha+} - R_i^{K\alpha-} + R_i^{(K-1)\alpha} = 0 \tag{10}$$

so that eq. (8) may be written as

$$P_{sc} = -\sum_{K}\int_{S^{K-1}} P^{K}_{sc}\, d^{K-1}x \tag{11}$$

where we have set

$$P^{K}_{sc} = \sum_{\alpha}\left[\!\!\left[R_i^{K\alpha}\left(G_i^{K\alpha} - G_i^{(K-1)\alpha}\right)\right]\!\!\right] \tag{12}$$

This procedure allows the reduction of the three vectorial degrees of freedom where $R_i^{(K-1)\alpha}$ is expressed in terms of $R_i^{K\alpha+}$ and $R_i^{K\alpha-}$. As to the remaining two vectors, they are determined by the application of the energy formulation that we do not recall here.

If one considers the example of ferroelectric dielectrics derived in the previous section, then one gets

$$P^{K}_{sc} = \left[\!\!\left[L_i^{K}\,(v_i^{K} - v_i^{K-1}) + R_i^{K}\left(\frac{d^{K}\pi_i^{K}}{dt} - \frac{d^{K-1}\pi_i^{K-1}}{dt}\right)\right]\!\!\right] \tag{13}$$

where v_i^{R} and π_i^{R} are respectively the mechanical velocity and polarization density in an R-dimensional space. $R = \{K, K-1\}$.

Now that we have explained with sufficient details the extended procedure distinguishing between rotational invariance and scale-change invariance principles, we can apply this method with no ambiguity. In previous works [5] although the problem of interfaces has been dealt with in various contexts (mechanics, simple electromagnetic materials and deformable semiconductors), and consistent results were obtained, the given solutions were partial and could not be extended to account for electromagnetic ordering (ferroelectricity, ferromagnetism, etc.) and other nonmechanical situations. This is due to the fact that the situations studied previously were thought of in mechanical terms. The study of diffusion processes in semiconductors, for example, was done by treating the diffusion current per unit charge, which has the dimension of a velocity, by analogy with the mechanical velocity concept. Let us note that when $G_i^{K\alpha}$ is a velocity eq. (12) includes a relative velocity which is objective or rotationally invariant. Here, one understands why the scale-change invariance principle did not appear as a necessity in a purely mechanistic interpretation. Indeed, rotational invariance appears to be sufficient for the determination of volume, surface and volume-surface interactions simultaneoulsy, but this is a mere coincidence. One may say that although the previous method led to correct results it was based on an accident where two distinct features coincide for certain particular applications. The distinction between two elements that coincide for particular situations is a well-known fact in the process of generalizing physical theories. One of the most famous examples is the passage from newtonian to einsteinian mechanics where the distinction between proper and improper times becomes a necessity if one wishes to treat a wider range of phenomena including high velocities.

B - Application to deformable ferroelectric dielectrics

Thermoelectromechanical equations

The thermoelectromechanical balance laws may be deduced in an elegant manner from three general principles written in global form for the material volume D swept out by the singular surface Σ. These are the principle of virtual power and the first and second principles of thermodynamics.

We call ${}^tP_a$, ${}^tP_i$, ${}^tP_{sc}$, ${}^tP_v$, ${}^tP_c$, ${}^tP_e$, K, $\hat{K}$, E, $\hat{E}$, N, $\hat{N}$, U^{em}, $\dot{Q}_h$, and $\mathscr{N}^e$, respectively, the total power of inertia forces, the total power of internal forces, the total power of internal contact forces (scale change), the total power of "volume" forces, the total power of external contact forces, the total power of prescribed forces (of any type, in the bulk, on surfaces or on lines), the total kinetic energy of the volume (regular region), the total kinetic energy of material points belonging to the singular thermodynamical surface, the total internal energy of the regular material region, the total internal energy of the singular thermodynamical surface, the total surface entropy of the singular thermodynamical surface, the total electromagnetic fields, the total rate of supply of heat, and the total rate of supply of entropy. In the sequel a superscript asterisk will indicate a virtual field or the value of an expression in such a field (which is not, in general, a solution of the actual problem).

Principle of virtual power

In a Galilean frame and for an absolute Newtonian chronology the total virtual power of inertial forces of the system balances the sum total of the powers of internal and external forces impressed on the system for any virtual velocity fields. With the above notation, this reads

$$ {}^tP_a^* = {}^tP_i^* + {}^tP_{sc}^* + {}^tP_v^* + {}^tP_c^* \tag{14} $$

First principle of thermodynamics

The time rate of change of the total energy contained in $D-\Sigma$ and on Σ is equal to the sum of the power developed by prescribed forces, the energy supply by radiation in $D-\Sigma$ and on Σ, and the flux of energy through the boundaries $\partial D-\Sigma$ and $\partial\Sigma$. This may be stated as

$$ \frac{d}{dt}(K+E) + \frac{\hat{d}}{dt}(\hat{K}+\hat{E}) + \frac{d}{dt}U^{em} = {}^tP_e + \dot{Q}_h \tag{15} $$

Second principle of thermodynamics

For any thermodynamical process the time rate of change of the total entropy of the system is never less than the sum of the total entropy supply and the total flux of entropy through the boundaries. Mathematically, this reads

$$ \frac{d}{dt}N + \frac{\hat{d}}{dt}\hat{N} \geq \mathscr{N}^e \tag{16} $$

For lack of space we do not reproduce all the needed expressions. We simply give the contributions that have to replace those given in Ref. [4]. Consequently for an electrically polarized, heat conducting deformable ferroelectric dielectric, the basic expressions to be carried out in eqs. (14) - (16) and which differ from those given in Ref. [4] are given or constructed as follows

$$ {}^tP_{(i)}^* = -\int_{D-\Sigma} P_i^*\, dv - \int_\Sigma \hat{P}_i^*\, da \tag{17} $$

$$ {}^tP_{sc}^* = -\int_\Sigma {}^{\Sigma}P_{sc}^*\, da \tag{18} $$

with

$$ P_i = F_i\, v_i + t_{ij}\nabla_j\, v_i - \rho\, {}^LE_i\,(\frac{d\pi_i}{dt}) + {}^LE_{ij}\nabla_j\,(\frac{d\pi_i}{dt}) \tag{19} $$

$$ \hat{P}_i = \hat{F}_i\, \hat{v}_i + \hat{t}_{ij}\hat{\nabla}_j\, \hat{v}_i - \hat{\rho}\, {}^L\hat{E}_i\,(\frac{\hat{d}\hat{\pi}_i}{dt}) + {}^L\hat{E}_{ij}\hat{\nabla}_j\,(\frac{\hat{d}\hat{\pi}_i}{dt}) \tag{20} $$

and

$$ {}^{\Sigma}P_{sc} = Z_i^+\, v_i^+ - Z_i^-\, v_i^- + \hat{Z}_i\, \hat{v}_i + K_i^+\left(\frac{d\pi_i}{dt}\right)^+ - K_i^-\left(\frac{d\pi_i}{dt}\right)^- + \hat{K}_i\left(\frac{\hat{d}\hat{\pi}_i}{dt}\right) \tag{21} $$

The application of the rotational invariance principle or objectivity requirement in the volume and at the interface lead to the following restrictions

$$ \sigma_{[ij]} = 0 \quad , \quad \hat{\sigma}_{[ij]} = 0 \quad , \quad F_i = 0 \quad \hat{F}_i = 0 \tag{22} $$

where the volume and surface symmetric tensors σ_{ij} and $\hat{\sigma}_{ij}$ show how the volume and surface mechanical continua are coupled with the polarizable ones.

$$ \sigma_{ij} \equiv t_{ij} - \rho\, {}^LE_i\, \pi_j + {}^LE_{ik}\, \nabla_k\, \pi_j \tag{23} $$

$$ \hat{\sigma}_{ij} \equiv \hat{t}_{ik}\left(\delta_{kj} - n_k\, n_j\right) - \hat{\rho}\, {}^L\hat{E}_i\, \hat{\pi}_j + {}^L\hat{E}_{ik}\left(\delta_{k\ell} - n_k\, n_\ell\right)\nabla_\ell\, \hat{\pi}_j \tag{24} $$

As to the application of the scale-change invariance principle, it leads to the following restrictions

$$ \hat{Z}_i = -\left(Z_i^+ - Z_i^-\right) = -[\![Z_i]\!] \tag{25} $$

and

$$ \hat{K}_i = -\left(K_i^+ - K_i^-\right) = -[\![K_i]\!] \tag{26} $$

The present formulation generalizes the procedures given in Refs. [4, 10] by taking into account volume and surface polarization gradients. Moreover, unlike eq. (4.5) in Ref. [10] the expression of ${}^tP_i^*$ governed by the objectivity requirement is split into two different contributions noted by ${}^tP_i^*$ and ${}^tP_{sc}^*$ and given by Eqs. (17) and (18). Only the first of these is subject to the objectivity requirement. As to the second, it is governed by the so-called scale-change invariance principle.

As pointed out earlier, the boundary contributions associated with $\int K_{ij} \nabla_j A_i^* \, dv$ and $\int \hat{K}_{ij} \hat{\nabla}_j \hat{A}_i^* \, da$ are $\int L_i A_i^* \, da$ and $\int \hat{L}_i \hat{A}_i^* \, d\ell$. Consequently, in the framework of a first gradient theory with respect to polarization one is led to the following contributions

$$\int_{\partial D - \Sigma} Q_i \left(\frac{d\pi_i}{dt} \right)^* da + \int_{\partial \Sigma} \hat{Q}_i \left(\frac{d\hat{\pi}_i}{dt} \right)^* d\ell \quad (27)$$

that must be added to the expression of ${}^tP_c$ in eq. (4.7) of Ref. [4]. The contributions Q_i and $\hat{Q}_i$ are to the polarizable continuum what the tractions T_i and $\hat{T}_i$ are to the lattice continuum. Although the resulting local volume and surface differential equations are not given here, their derivation is a simple question of calculation which is now a well established procedure that follows a purely deductive scheme. The basic elements included in the above mentioned global formulation will be developed elsewhere leading to local volume and surface differential equations needed for the study of wave propagation problems in deformable ferroelectric dielectrics

Conclusion

Let us conclude by recalling that in Part A we have considered a general and systematic formulation that applies to different nonmechanical continua (ferroelectric, ferromagnetic and nematic states) and to different scale changes (volume-surface, surface-line, line-point). In part B, a particular case relative to deformable ferroelectric dielectrics with interfacial properties has been given. Although the latter case includes only one nonmechanical continuum (ferroelectricity) and one scale change (volume-surface), it generalizes theories that could not be extended through the systematic procedure given previously in Refs. [2,4].

References

[1] R. Gatignol and P. Seppecher, Journal de Mécanique Théorique et Appliquée, numéro spécial, 1986, p. 225 à 247.

[2] N. Daher and G.A. Maugin, "The method of virtual power in continuum mechanics. Application to media presenting singular surfaces and interfaces, Acta Mech. **60**, 217-240 (1986).

[3] A.M. Albano, D. Bedeaux, J. Vlieger, "On the description of interfacial electromagnetic properties using singular fields, charge density and currents at a dividing surface", Physica **102A**, 105-119 (1980).

[4] N. Daher and G.A Maugin, "Virtual power and thermodynamics for electromagnetic continua with interfaces", J. Math. Phys. **27**, 3022-3035 (1986).

[5] N. Daher, "Principe des puissances virtuelles étendu aux discontinuités et interfaces, application à l'acousto-électronique". Doctoral Thesis, Besançon (1987).

[6] N. Daher, "Energy formulation for electronic, optical and acoustical applications including interfacial properties and irreversible processes", Synthetic Metals **67** (1-3), 287-291 (1994).

[7] P. Germain, "La méthode des puissances virtuelles en mécanique des milieux continus : théorie du second gradient", J. Mécanique **12**, 235-274 (1974).

[8] G.A. Maugin, "The method of virtual power in continuum mechanics. Application to coupled fields", Acta Mech. **35**, 1-70 (1980).

[9] N. Daher, "Leibniz-d'Alembert energy formulation, application to piezoelectric media with interfaces", EFTF 94, Technical University Munich, Weihenstephan, Germany, vol. II, pp. 749-760 (1994).

[10] G.A. Maugin, J. Pouget, "Electroacoustic equations in elastic ferroelectrics", J. Acoust. Soc. Amer. **68**, 575-582 (1980).

[11] S. Chandrasekhar, "Liquid crystals", second edition, Cambridge University Press, 1992.

9 TH EUROPEAN FREQUENCY AND TIME FORUM - BESANÇON - 8 - 9 - 10 MARCH 1995

Session POSTER Number

STUDIES OF CRYSTALLINE QUALITY AND VIBRATION MODES IN LANGASITE SINGLE CRYSTALS.

A. ZARKA*, B. CAPELLE*, J. DETAINT, D. COCHET-MUCHY***.**

*Laboratoire de Minéralogie-Cristallographie,Université Paris VI et Paris VII,Associé au CNRS 75005 Paris France.

Centre National d'Etudes des Télécommunications,dept. PAB/BAG/MCT/CMM 92220 Bagneux France.* Crismatec 2, Rue des Essarts 38610 Gieres France.

Langasite ($La_3Ga_5SiO_{14}$ or L.G.S.) is a new piezoelectric material with promising properties.Observations using X-ray transmission topography have permitted to verify the quality of samples of different origins. Defects more generally encountered are dislocations, growth bands and stacking faults. The dislocations are oriented perpendicularly to the growth striations. In some cases, they are grouped in more or less compact bundles. On the other hand, small stacking faults were also observed in several crystals. In some samples, the density of the dislocations is extremely reduced (several lines per cm^2) and frequently the growth bands are very few.

Generally speaking, the crystalline quality of the samples can be estimated as excellent. However, studies concerning the growth of langasite must be continued to enhance their quality even further (to reduce the number of plane, linear and point defects). Vibration modes of different types of devices made with this material have been studied by X-ray transmission topography either in the classical or stroboscopic mode. This study was concerned with bars vibrating in elongation, plane or plane-convex resonators. For the resonators vibrating in thickness shear mode a lateral anisotropy variation of the overtone range predicted by a theoretical model was observed.

SWEEPING OF VERY HIGH PURITY SYNTHETIC QUARTZ UNDER VACUUM AND AIR ATMOSPHERE

Jean-Pierre BACHHEIMER

Université Joseph Fourier (Grenoble I)
Laboratoire de Spectrométrie Physique (associé au CNRS)
B.P. 87 - 38402 Saint-Martin-d'Hères Cedex (France)

ABSTRACT

Comparative air and vacuum sweeping at ~ 500°C have been performed on high purity (Aℓ < 0,1 ppm) synthetic quartz with very low (10 /cm²) and high (200/cm²) dislocation densities. Results of measurements of sweeping currents, activation energies, ohmicity tests and infrared spectra are very different under air or vacuum but only minor differences appear between low and high dislocation density samples. Comparison is also given with higher impurities concentration sample.

INTRODUCTION

"Sweeping" [1] of quartz (SiO_2) is an industrial high temperature electrolysis treatment which is used to improve the radiation hardness of piezoelectric devices. This occurs because sweeping gives preventively similar effects , but more stable, on the Al-M^+ ($M^+ = Na^+, Li^+$) and OH impurity centers which are introduced during hydrothermal growth of cultured quartz. Schematically sweeping in an air (wet) atmosphere replaces the alkali swept towards the cathode by hydrogen to form Aℓ-OH centers, while sweeping in a vacuum atmosphere should form at the completion Al-hole centers after the depletion of the as grown OH. [In fact depending on the relative Aℓ/OH content [2] an intermediate Aℓ-X center [3] can be implicated between the Aℓ-OH and Aℓ hole states - and it is then necessary to irradiate the sample to get finally only Aℓ-hole centers].

Monitoring of the sweeping process can be performed by the evolution of the infrared (I.R.) and visible spectra before and after the process because : i) as grown OH displays peaks S_4 (the largest at 3585 cm^{-1} or 2789 nm), S_3, S_2, S_1 on a large background absorption (usually measured at 3500 cm^{-1} and 3410 cm^{-1}) ii) Aℓ-OH displays two I.R. peaks at 3380 cm^{-1} ("e_2" large) and 3306 cm^{-1} (e_1 weak) iii) Aℓ-hole is involved in the A-bands smoky coloration which absorption peaks at ~ 460 nm. But during the process practically the only in situ control consists in measuring the electrical current variations as a function of elapsed time. Roughly the initial current is due to alkali displacement and the final current is due to H^+ displacement [1,4]. However there is a lack of data to correlate these currents with the impurities contents and modifications as well as the structural defects concentration. In this aim we present the results (sweeping currents, ohmicity tests, activation energies, infrared spectra) obtained at 500°C under air or vacuum atmosphere from very high purity synthetic quartz with low and high dislocation densities.

EXPERIMENTAL

The samples (~ 1 x 1 x 1 cm^3) were successively cut from two bars of pur Z synthetic quartz (GEC) of very high purity (Aℓ < 0,1 ppm). One with a low (< 10/cm^2) dislocations density, the other with a high (~ 200/cm^2) dislocations density. Their I.R. spectra, obtained with a λ/9 Perkin Elmer double beam (Φ3) spectrometer (nearly polarized) is shown fig. 1.

Table 1 gives the averaged infrared absorption coefficients (cm^{-1} at room temperature) before sweeping and [x,xxx] after vacuum sweeping (see further) in the OH free part of the samples.

Table 1.

Samples	αS4	α 3500	α3410
"L" low(10/cm^2) dislocation density	0,0075 ± 2 [0,000]	0,0260 ± 5 [0,023]	0,075 ± 1 [0,070]
"H"high (200/cm^2) dislocation density	0,006 ± 5 [0,000]	0,026 [0,023]	0,075 ± 1 [0,069]

Electrodes are Pt (annealed and washed with HCl-NOH 50 % solution) foils pressed (~ 0,4 kg/cm^2) on the Z faces which are optically polished or finely (7 μm) ground. The electric field is applied slowly (1 to 3') when the sweeping temperature (500°C) is reached. Ohmicity tests are obtained by step to step decrease of the electric field. Electrolysis in air was carried out in an open furnace while vacuum electrolysis was performed in a furnace operating at ~ 3.10^{-6} mm Hg*. Because the small Aℓ concentration of the samples Aℓ-hole absorptions cannot be detected.

RESULTS

. Comparison between L and H samples (500° - 2000 V/cm)

- air : fig 2* shows the sweeping current as a function of time up to 25 days : they are practically identical and decrease from 1,3 μA/cm^2 at 3' to ~ 0,033 μA/cm^2 where we still observe a slow decay. (The quasi stabilization after 12 days can be due to a long term effect of an Ω test although it lasted itself only ~ 2 hours). Comparison between H samples with polished and non polished Z faces show a slight difference which do not affect the overall behaviour. H and L samples also exhibit the same Ω test (fig. 3) at the end of a short transient state of 1 to 2' (with a tendency to increase for the low voltages and when the sweeping time increases). Activation energies are about 1,95 eV [500 to 300°C] at the end of the sweeping.

* passing from 10^{-6} to 10^{-3} mm Hg induces only negligeable sweeping current variations after a transient state.

Fig. 4-5 show the evolution along the Z axis of the S4 peaks before and after sweeping .The effects are qualitatively similar : both samples display a depletion of the OH (S4) content over about half the sample. We note however an additional depletion on the cathode side on the L sample. The "front-lag" on the depolished H sample can be perhaps attributed to a shorter sweeping time (21 instead of 25 days). Normally, we should find e_2 (Aℓ-OH) peaks where S4 is not totally depleted but probably because of the low Aℓ content we did not succeed to detect e_2 with a low (- 194°C) temperature Z scan [except two doubtfull points corresponding -according to our calibration-to a~ 0,02 ppm (w) Aℓ content : that could be a confirmation of the maker's data - X irradiation of another sample from the same L bar show e_2 peaks corresponding to 0,06 ± 0,03 ppm (w) Aℓ concentration-]

- vacuum : fig. 6 : as for the air case sweeping currents appear also similar for the H and L samples but here the currents are about an order lower. A fact that we have already noted [5] for higher impurities concentrations. Another difference with the air curve is a steepest decrease of the currents between ~30' and 10h in vacuum atmosphere. Ω test (fig. 7) show further non linear effects and the transient state to reach an equilibrium value after each step can be very long (10-30'). Activation energy is stronger (2,3 eV for L) than for air sweeping. I.R. spectra show fig. (8-9) a general S4 depletion which becomes total on about half the sample for both qualities.

. Comparison with another quartz of higher impurities concentrations. Fig. 10 shows the air sweeping currents obtained at 1000 V/cm - 500°C with a sample of L quality and another sample with 6 , 5 ppm (weight) Aℓ, Na concentrations and α S4 ~ 0,05 cm^{-1}. Both samples had Z faces finely grounded. We see that the current from this last sample is about 10-5 times higher than from the L quality.
These two samples show at the beginning of the electrolysis a remarquable diffusive ($\sim t^{-1/2}$) currents variations. (Generally such~ $t^{--1/2}$ current -not always at the beginning- variations are often observed on certain parts of the i = f(t) curves as well under air as under vacuum atmosphere).

DISCUSSION

These quartz of very low impurities concentrations give very low sweeping currents as expected [1, 4]. The small difference between L and H samples means that the diffusion coefficients of the impurities implicated in the sweeping process are practically little affected by these levels of dislocations densities. The initial sweeping current depends on the alkali concentration * which is itself roughly proportionnal to the Aℓ concentrations for electric charge compensation reasons so, if we admit-as for natural quartz [4] - that the final H^+ current depends on the Aℓ influence, the whole curves of sweeping current are roughly translated, all things being equal, as a function of Aℓ concentrations. Considering the incertitude on Aℓ concentrations determinations and also that they can vary inside the sample, it seems that we could keep in mind as an order of magnitude a $[A\ell]^{1/2}$ dependence between the overall sweepings currents (within the same experimental conditions). Under vacuum experiments, due to the lack of H^+ furniture at the anode, we get sweeping currents which are about an order of magnitude smaller than those obtain under air as for samples of higher impurities contents [5]. For the same reason vacuum ohmicity tests give non linear phenomena in contrary to air test which gives only small non linearity showing that space change formation and electrode processes are probably much less important for this last case. Otherwise space charge formation could explain long term decay of the sweeping currents.

CONCLUSION

As expected from very low impurities concentrations, the sweeping currents are found to be very low in the very high purity synthetic quartz investigated. In a first approximation dislocation density from < 10 to 200/cm^2 show only a minor effect as well on the electrical response of the samples as on the final result of the sweeping on the OH content. We have also found (contrary to other results [7]) that there is no relation between the dislocation density and the [OH] concentration of these samples of very high purity.

ACKNOWLEDGEMENTS :

The author wishes to thank D. Bellet (L.S.P and E.S.R.F. Grenoble) who performed X-ray topographies to determine dislocation densities ; G. Dolino (L.S.P) for suggestions ;D.F. Croxall from GEC Research-England-where were grown the very high purity quartz investigated in this study.

REFERENCES

(1) see the review of J.J. Martin, IEEE Trans. on Ultrasonics Ferroelectrics and Frequ. Control 35 , 288-296 (1988)
(2) J.P. Bachheimer , Proceeding of the 7th EFTF Neuchâtel 671-675 (1993)
(3) J.P. Bachheimer , Phys. Chem. Solids. 54, 1501-1507 (1993)
(4) A.K. Kronenberg, S. H. Kirby , American Mineralogist 72, 739-747 (1987)
(5) J.P. Bachheimer , G. Dolino , 43rd Ann. Freq. Control Symp. IEEE, 517-520 (1989)
(6) J. G Gualtieri and D.W Eckart , Proc. 40 th An Freq. Cont. Symp. IEEE, 115-120 (1986)
(7) J.C. Brice , Rev. Modern. Phys. 57, 105-146 (1985)

* We can determine the transported charge [6] before and above the stabilized currents for the two qualities of the fig. 10 (i.e. t < 30' for L and t<10 h for the other quartz) : we find $0,3.10^{-3}$ and 20.10^{-3} cb/cm^2 which can be compared with the ~ 100(with a possible 2 factor) ratio of the impurities contents of these samples

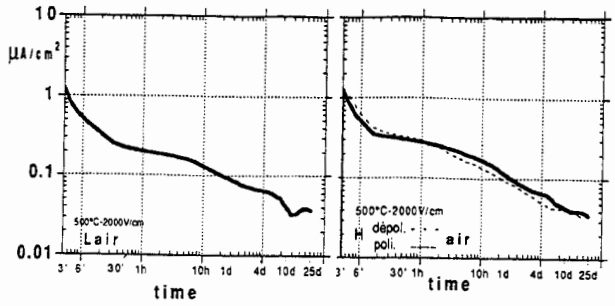


Fig. 2 : sweeping currents versus time (air) for the L and H quartz

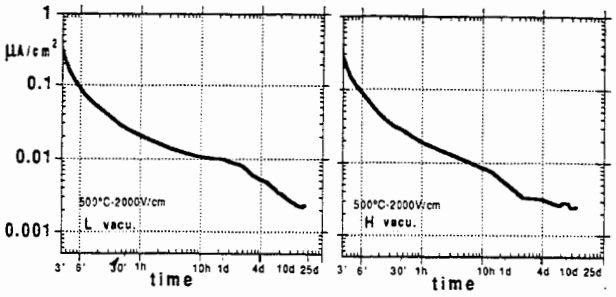


Fig. 6 : sweeping currents versus time (vacuum) for the L and H quartz

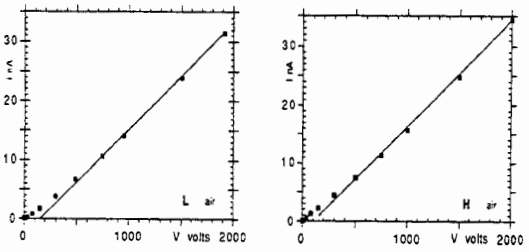


Fig. 3 : Ohmicity test (air)

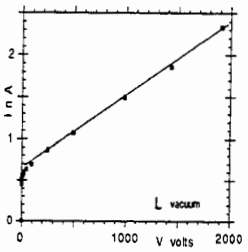


Fig. 7 : Ohmicity test (vacuum)

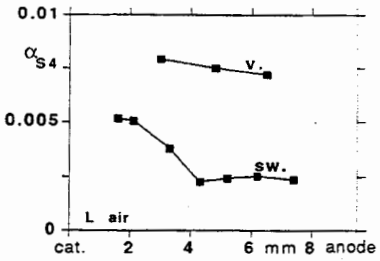


Fig. 4 : S4 (OH) before and after the sweeping of the Fig. 1 (L-quartz)

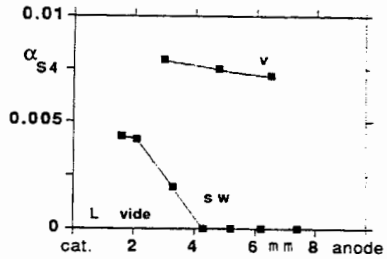


Fig. 8 : S4 (OH) before and after the sweeping of the Fig. 6 (L-quartz)

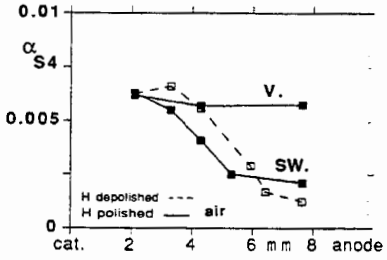


Fig. 5 : S4 (OH) before and after sweeping of the Fig. 1 (H-quartz)

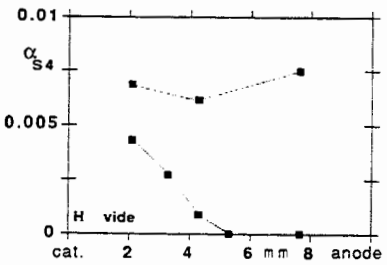


Fig. 9 : S4 (OH) before and after the sweeping of the fig. 6 (H-quartz)

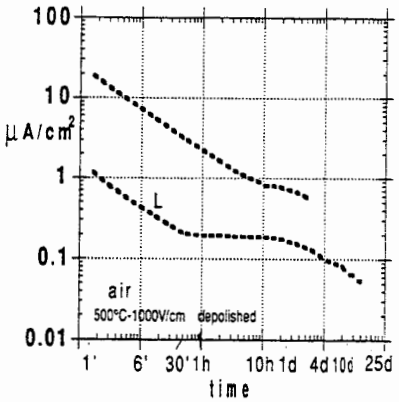


Fig. 10 : Sweeping currents comparison between a L quartz (Aℓ < 0,1ppm) and another quartz of higher impurities concentrations (Aℓ ~ 5 ppm w)

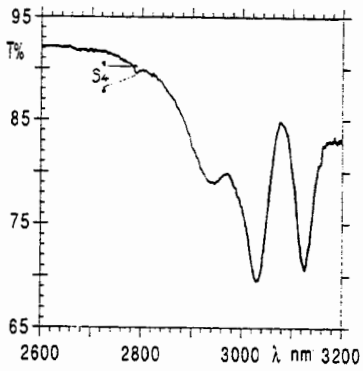


Fig. 1 : typical I.R. spectrum (r.t) of the L and H samples showing the S4 (OH) peak superimposed on the overtone vibrations (+ negligeable large background OH) of the SiO_2 lattice

MICRO-FABRICATED SILICON STRUCTURES : EXPERIMENTAL AND THEORETICAL INVESTIGATIONS OF 3-D ETCHED SHAPES

C. TELLIER AND S. DURAND

Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
F - 25030 Besançon Cedex, France

ABSTRACT : The anisotropic etching behaviour of doubly-rotated silicon plates was studied. For a 35% KOH solution 3D structures etched on {hhl} planes are bounded by curved regions and limiting planar facets whose directions of alignment and extents vary with the angles of cut. Tentatives are made to predict dissolution shapes for etched structures using a tensorial model of the anisotropic dissolution process. Experimental shapes are compared with theoretical shapes as derived from a 3D numerical simulation.

1. INTRODUCTION

Usually the sensing elements of micro-resonators are based on quartz crystals. But silicon crystal constitutes an alternative material. Actually surface acoustical waves devices are composed of silicon resonant membranes. Some attempts are also made to realize bulk waves multi-layered structures such as metal-ZnO-Si microstructures. In these devices resonant or sensing elements are obtained by anisotropic wet etching of (100) silicon wafers. But it is also interesting to work with singly (hk0) or doubly (hhl) rotated silicon plates in order to minimize temperature effects or to enhance the sensor sensitivity specially for piezoresitive sensors. With this object the theoretical prediction of etched shapes constitutes a necessary stage for a computer-aided design of silicon sensors.

2 EXPERIMENTAL RESULTS

2.1 Experimental

In the experiments p-type silicon wafers (- ∅ = 13 mm) with different {hhl} orientation were used. For passivation the wafers were thermally oxided or/and a CVD silicon nitride layer was deposited. The samples were etched in a 35% KOH solution maintained at a constant temperature of 55°C. Standard photolithographic techniques were used to obtain masks of circular, square or star-like patterns. In addition some special masks were designed for the fabrication of mechanical devices such as mesa in suspension or cantilever. Qualitative and quantitative evaluations of the anisotropy were made by Scanning Electron Microscopy (SEM) and additionally by profilometry measurements (evaluation of the etch rate from the decrement in thickness).

2.2 Results

Figures 1 and 2 give two structures etched respectively on a (100) and on a (112) silicon wafer. These figures give evidence for an increasing anisotropy of the chemical etching when we deviate from a {100} or from a {110} plate. In particular one can observe that the (112) cantilever-like structure is composed of a mesa which seems, in the + X' direction (Figure 2), to be bordered by a curved region formed by successive planes slightly disoriented with respect to the reference surface.

Let us recall that {hhl} planes are doubly-rotated plates whose orientation is determined by means of two angles of cut φ and θ [1]. It is clear that here φ = - 45° and that θ varies with the Miller indexes h and l. Thus in order to follow the changes in the anisotropy with the angle of cut, θ, we have etched eleven masked and differently-oriented plates corresponding to (hhl) planes ranging from (112) to (118) and from (221) to (661). Some results related to (661), (112) and (221) silicon wafers are respectively illustrated in Figures 3a, 4a and 5a. We have chosen to present etched "holes" and mesa fabricated with masks of initially circular shape. In this case all the (hkl) planes are potentially present in the etching sequence and can "a-priori" participate to an etched structure. In this condition we can really appreciate the anisotropy effect which leads to the disappearance of some (hkl) planes [2-6].

Figures 3a to 5a show that etched holes and mesa are bounded either by curved regions (CR) or by quasi-planar facets (f). Moreover in some specific directions bordering curved and/or planar facets form sharp edges (9) which penetrate within the crystal under the mask. In general the top-view etched shapes differ markedly from the initially circular shape. Another interesting result which emerges from these figures is that the etched shapes of all structures satisfy the symmetry properties relative to the <110> axis.

Let us now consider Figure 4a, one may observe that the (661) etched shape for mesa does not deviate very much from the well-known shape related a (110) mesa [6]. In particular we can crudely recognize some symmetries specific to {110} planes even if the bordering inclined shoulder is more extented in direction - X' than in direction + X'. Turning now our attention to (112) and (221) etched structures we remark the marked convex undercutting of mesa specially for the (221) mesa (Figure

5a). Moreover as expected the slightly inclined shoulder S_1 lies in the + X' direction for the (112) structure and in the - X' direction for the 221 structure. This may be attributed to the fact that we pass from a (112) plane to a (221) plane by rotating the (111) plane from respectively - 16° and ≈ + 19° about the [110] axis. Finally the reader can also appreciate the influence of orientation on the surface texture of un-masked regions.

In conclusion all structures show marked anisotropic effects which firstly cause important changes in the initial "circular" shape and secondly, induce the formation of planar limiting facets or curved facets under the mask. It is clear that firstly, dissolution profiles characterized by the development of sharp edges under the mask produce a mechanical weakness of etched structures and that, secondly, the formation of large curved regions or of large lateral underetchings is undesirable in many cases because of limited dimensions of structures. Consequently a computer-aided design of etched mechanical devices passes through a theoretical simulation of three-dimensional etched shapes of structures.

3. 3D SIMULATION OF ETCHED SHAPES

3.1 The tensorial model for anisotropic chemical etching

In the tensorial three-dimensional model [5] we associate to each surface element of orientation (φ, θ) a dissolution slowness vector, $\vec{L}$, which characterizes the anisotropic wet etching process. The normal etch rate, R, of a surface is simply connected to the magnitude, L, of the slowness vector by means of R = 1/L. Tellier and co-workers [8-10] have shown that it is possible to calculate the displacement, $\vec{P}$, of a surface element moving within the crystal during the chemical attack starting from the analytical equation for the dissolution slowness, L(φ, θ) i.e. from the representative surface for the dissolution slowness vector $\vec{L}$. For silicon crystal which belongs to class m3m we can adopt the following general form [9] :

$$L(\varphi,\theta) = \sum_{k=0} a_k^* (\cos\theta \sin\theta)^{2k} + \sum_{l=1}\sum_{m=0} b_{l,m}^* \left(\sin\varphi \cos\varphi \cos^2\theta\right)^{2l} (\cos\theta)^{2m}$$

The coefficients of series are expressed in terms of dissolution constants which may be evaluated from experiments. In reality we estimate only constants related to tensors of higher rank. In practice constants for tensors of rank 12 or 18 where recently determined by a systematic analysis of bi-dimensional etched shapes [4,5]. Here attemps are made to derive 3D dissolution shapes using the previously proposed set of dissolution constants.

3.2 The 3D simulation

The simulation program [11] works with doubly-rotated surface elements whose orientations are now defined by four angles, then calculates the displacements $\vec{P}$ of surface elements potentially present at the starting structure and finally, eliminates elements which disappear with further etching. In the final stage constant-level contours of etched mesa and hole as well as cross-sectional dissolution profiles can be viewed. Dissolution profiles correspond to cross-sectional planes which intersect the reference surface along various lines L_i (see Figure 5c for example).

Figures 3d, 4d and 5b give some typical results for holes. From a general point of view we observe that, for (112) and (221) silicon planes, predicted shapes exhibit contours which differ markedly from the starting circular shape. This departure is more marked for the (221) hole than for the (112) hole in good agreement with experiments. The anisotropy effect for the (661) hole seems less pronounced for the theoretical structure (Figure 4b) than for the experimental one (Figure 4a). Nevertheless in Figure 4b we recognize the facets f_1 and f_2 as well as the slightly curved region (CR) present on the SEM micrograph even if the simulation leads to a too accentuated extent of the curved region as compared with experimental data. In contrast the predicted (221) and (112) holes are composed of a curved region (CR) and of facets (f_1, f_2) whose directions of alignment coïncide quite accurately with the theoretical intersecting lines L_1 and L_2 (Figures 3d and 5c). Moreover the theoretical relative extents of facets with respect to that of the curved region appear to be in close agreement with the relative extents illustrated by SEM images.

Turning our attention to dissolution shapes for the (112) mesa a crude agreement between predicted and experimental shapes can at the first sight, be depicted in Figures (3a and 3b). But a careful examination of the SEM micrograph reveals some differences with in particular the development of a fragmented cross-sectional profile in region R_1. The experimental dissolution profile is certainly composed of two limiting facets forming a sharp edge whereas in theory only a facet develops which penetrates deeply under the mask.

4. CONCLUSION

The few examples presented in the above section allow us to draw a conclusion. The agreement between theory and data concerns in particular reference surfaces (hhl) with lower Miller indexes. This agreement covers the formation of sharp edges, curved regions and limited planar facets in some specific regions. For planes corresponding to higher Miller indexes deviations are observed in the relative extent of successive curved or planar facets limiting the structure. These departures may be attributed to a lack of precision in the estimation of the relative amplitude of the successive maxima in $\vec{L}$. Thus we are now working on further adjustments of the dissolution constants in order to derive 3D shapes which coïncide closely with etched shapes whatever the orientation of the starting reference surface.

REFERENCES

1. I.E.E.E. "Standard on Piezoelectricity (IEEE, New-York, 1978).
2. K.E. PETERSEN, Proc. IEEE, 70 (1982) 420.
3. H. SEIDEL AND L. CSEPREGI, Sensors and Actuators, 4 (1983) 455.

4. C.R. TELLIER AND A. BRAHIM-BOUNAB, J. Mater. Sci., 29 (1994) 5953.
5. C.R. TELLIER AND A. BRAHIM-BOUNAB, J. Mater. Sci., 29 (1994) 6354.
6. C.R. TELLIER AND T.G. LEBLOIS, J. Phys. III, 2 (1992) 1259.
7. C.R. TELLIER, D. BENMESSAOUDA, A. BRAHIM-BOUNAB AND S. DURAND 2nd Japan-France Congress on Mechatronics, Takamatsu (Japan), November 1-3, 1994.
8. C.R. TELLIER AND J.L. VATERKOWSKI, J. Mater. Sci., 24 (1989) 1077.
9 A. BRAHIM-BOUNAB, J.Y. AMAUDRUT AND C.R. TELLIER, J. Mater. Sci., 26 (1991) 5585.
10. C.R. TELLIER, J. Crystal Growth, 96 (1989) 450.
11. C.R. TELLIER AND D. BENMESSAOUDA, 8th European Frequency and Time Forum, Munich (Germany), March 8-9, 1994.

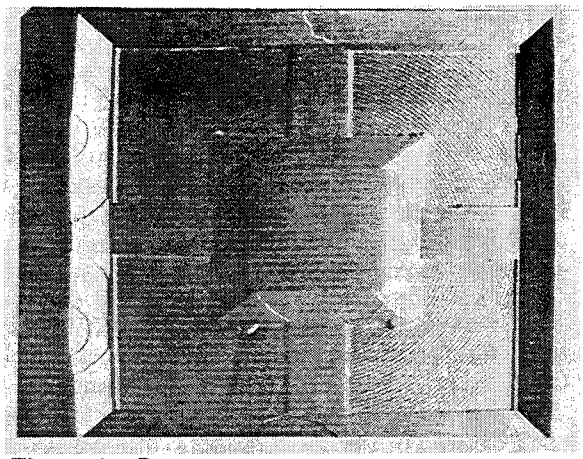


Figure 1 : Resonant structure composed of a mesa etched on a (100) plane.

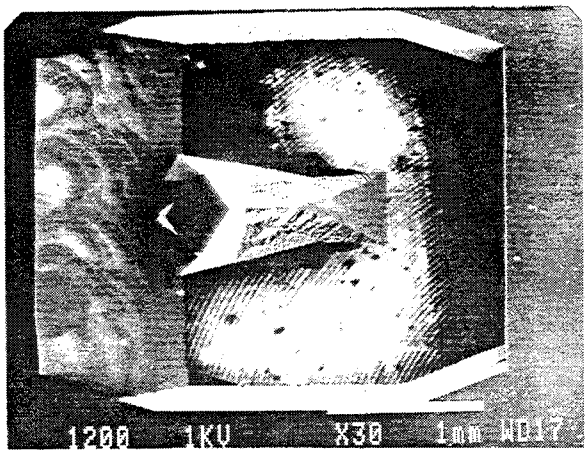
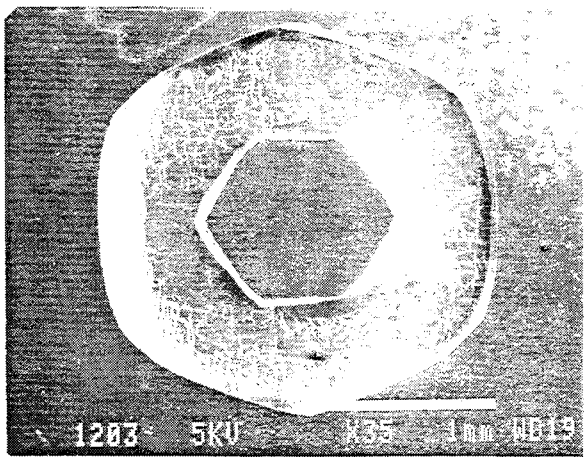
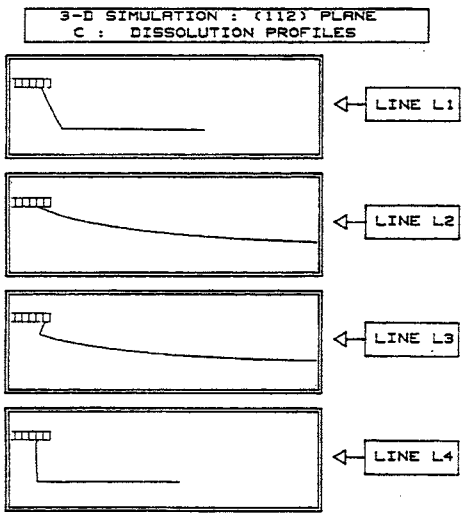
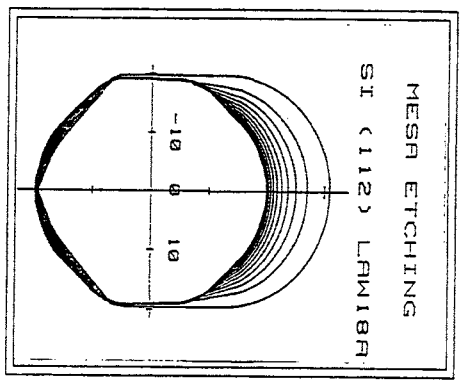


Figure 2 : A cantilever-like structure etched on a (112) silicon plane.







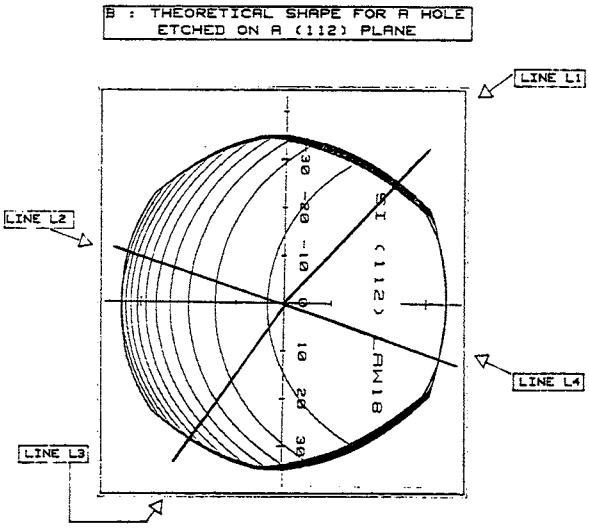
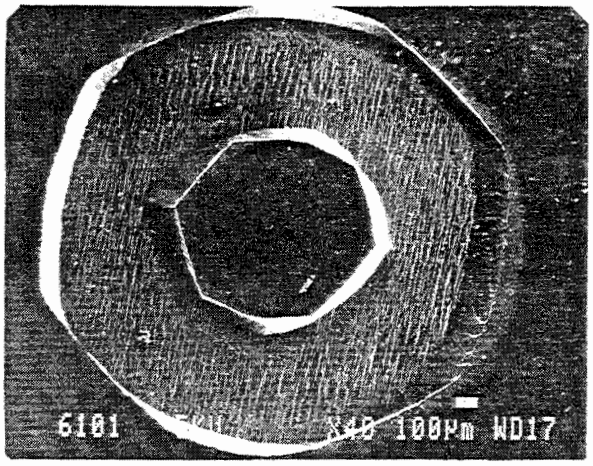


Figure 3 : Experimental and theoretical shapes for mesa and diaphragms etched on a (112) plane :
(a) Experimental structure
(b) Theoretical constant level contours for a hole
(c) Cross-sectional profiles corresponding to various intersecting lines.



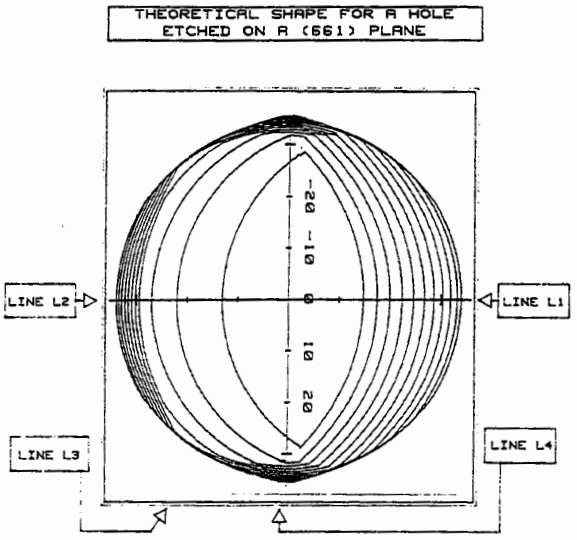
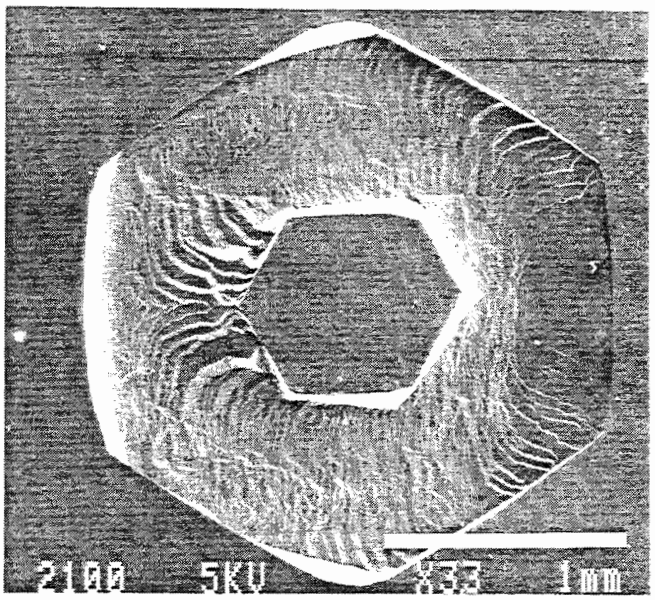
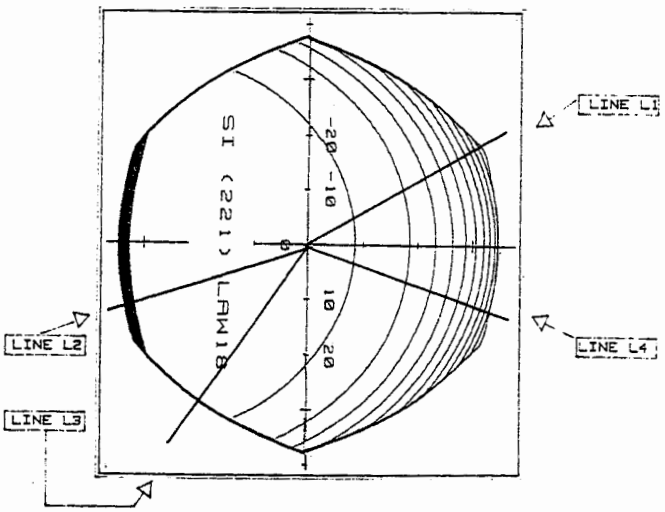


Figure 4 : Experimental and theoretical shapes for mesa or hole etched on a (661) plane :

(a) : etched structure, **(b)** : theoretical shape for a hole



B : THEORETICAL SHAPE FOR A HOLE
ETCHED ON A (221) PLANE



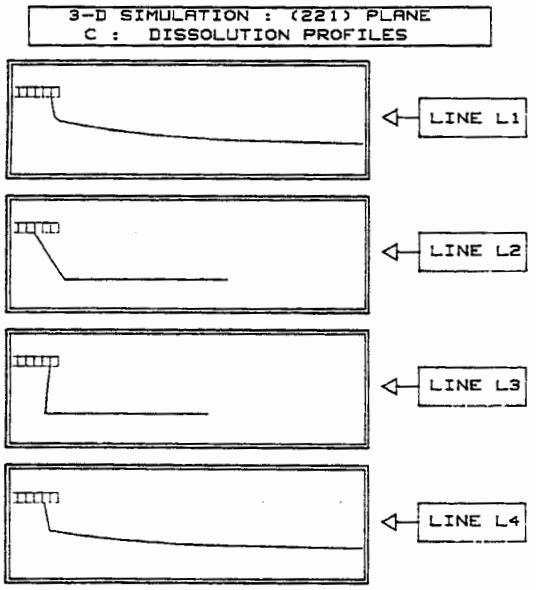


Figure 5 : Experimental and theoretical shape for mesa or/and hole etched on a (221) silicon plane

(a) S.E.M. micrographs of etched structure
(b) Theoretical constant-level contours for a (221) hole
(c) Cross-sectional profiles corresponding to various intersecting lines.

ACKNOWLEDGMENT : The authors wish to acknowledge the Direction of Recherches et Etudes Techniques (D.G.A.) for its financial support.

RESONATORS

$(Ba_{5-x}Sr_x)Nb_4O_{15}$: A NEW CERAMIC DIELECTRIC RESONATOR FOR APPLICATION IN COMMUNICATION SYSTEMS

H. SREEMOOLANADHAN, M.T. SEBASTIAN and P. MOHANAN*

Regional Research Laboratory, Thiruvananthapuram - 695 019, India.
*Dept. of Electronics, Cochin Univ. Sci. Tech., Kochi - 682 022, India.

Abstract

A new dielectric resonator system of materials has been identified with formula $Ba_{5-x}Sr_xNb_4O_{15}$. The synthesis and characterisation of these ceramics for different values of x by the solid state method produced suitable properties such as high ε_r (>40), high quality factor (>10,000) and low τ_f. The structure of these ceramics is hexagonal.

Introduction

Single crystals of Quartz possess excellent quality factor and thermal stability. Hence their use for tele-communication purposes is well-known; but is confined in the MHz region. At microwave frequencies, ceramics with high relative permittivity, called 'Dielectric Resonators' (DR) are used [1]. Dielectric Resonators possess high quality factor and low coefficient of thermal variation of resonant frequency. For practical applications, one must have ceramics with ε_r in the range 25-100, unloaded quality factor (Q_u) > 2,000 and low τ_f, preferably within $\pm$ 20 ppm/°C. By virtue of these properties, DRs permit miniaturisation of Microwave Integrated Circuits (MICs). The versatility and adaptability of DRs have made them indispensable components for application such as filters, oscillators etc. [2] in Cellular Telephones and Satellite communication [3]. To suit specific applications, the dielectric properties of ceramics are adjusted or optimised with the help of various dopants and additives [4,5]. Recently, Sreemoolanadhan *et al.* reported [6] the possibility of using $Ba_5Nb_4O_{15}$ based ceramics for application as dielectric resonators in communication systems. The present work is an investigation into the preparation and microwave characterisation of newly found ceramics in the BaO-SrO-Nb_2O_5 system. The effect of substitution of SrO for BaO has been studied and it improves the dielectric properties at microwave frequencies.

Experimental procedure

The ceramics are prepared by the conventional solid state method. The starting materials are $BaCO_3$ (99.9%) $SrCO_3$ (99.9%) and Nb_2O_5 (99.9%). Six compositions, corresponding to x=0,1,2,3,4 and 5 are separately mixed using distilled water in an agate mortar for 30 min. and dried. The dried powders are again ground and calcined in Alumina (Al_2O_3) crucibles at temperatures in the range 1200°C-1350°C for 4 hours in air. The calcination temperature increases as the value of x increases. Calcined powders are thoroughly ground for about 1½ hours and pressed into cylindrical compacts of size 10 mm diameter and 6-7 mm thickness, under a pressure of 4.5 tonnes. The compacts are then sintered at temperatures 1380°C-1440°C for 4 hours. The sintering temperature is slightly lowered with increase in x. However, the ceramic with x = 5, i.e., $Sr_5Nb_4O_{15}$ did not densify even at 1600°C. So it is found that an addition 1 wt. % of La_2O_3 to the calcined powder enhanced the sinterability and the $Sr_5Nb_4O_{15}$ was sintered at 1450°C for 4 hours. Sintered compacts are used for density measurement, X-ray powder diffraction and microwave dielectric property measurements.

For the microwave dielectric constant and quality factor measurements, HP 8510B Network Analyser coupled to HP 8514 B Test set and HP 8341 B Sweep oscillator are used. The ε_r is measured by the method of Hakki and Coleman [7] which was later modified by Courtney [8]. A cylindrical DR of 10 mm diameter and 6-7 mm length is resonated in the end- shorted condition. The ε_r is calculated from the frequency of TE_{011} mode and the dimensions of the DR with the help of HP 9000, 300 series Instrumentation Computer. The unloaded quality factor is measured using a stripline method proposed by Khanna and Garault [9]. The DR coupled to a 50 Ω microstripline, enclosed in a brass cavity of 5 x 5 x 3 cm^3, acts as a band rejection filter. From

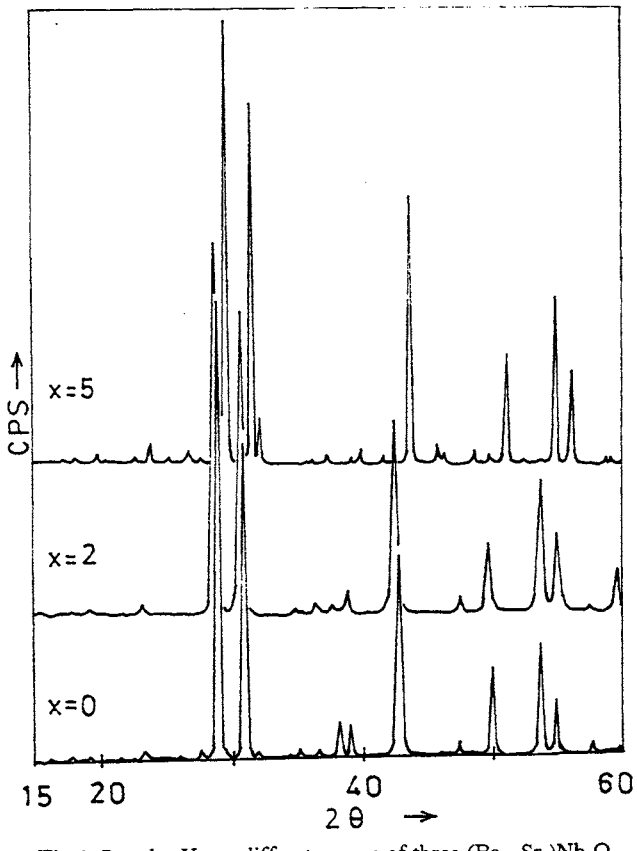


Fig.1. Powder X-ray diffractograms of three $(Ba_{5-x}Sr_x)Nb_4O_{15}$ ceramics.

the different modes, Q_u is obtained by the computer and the maximum Q_u is taken. The temperature coefficient is measured by keeping the DR in an aluminium cavity of 4.5 cm diameter and 7 cm height keeping the top open. The range of temperature is 25°C-80°C.

Results and Discussion

Fig.1 shows the powder X-ray diffractograms obtained using CuKα radiation for $(Ba_{5-x}Sr_x)Nb_4O_{15}$ ceramics. The patterns show the isostructural nature of these ceramics as well as the phase purity. The ceramics have hexagonal structure with one molecular formula weight. The pattern of $Ba_5Nb_4O_{15}$ (x = 0) is in agreement with the earlier reports (JCPDS file No. 14-28). The microwave dielectric properties such as ε_r, Q_u and τ_f are shown in Table I. The unloaded

Table I. Microwave dielectric properties of $(Ba_{5-x}Sr_x)Nb_4O_{15}$ ceramics

Material	Resonant freq. (f_o) (GHz)	ε_r	Q_u	$Q_u \times f$ (GHz)	τ_f ppm/°C
$Ba_5Nb_4O_{15}$	5.07	38.4	2,500	12,750	-9
$Ba_4SrNb_4O_{15}$	4.70	48.4	2,100	9,870	+11
$Ba_3Sr_2Nb_4O_{15}$	4.71	49.5	3,500	16,400	+7
$Ba_2Sr_3Nb_4O_{15}$	4.61	50.6	4,600	21,200	+85
$BaSr_4Nb_4O_{15}$	5.00	44.9	8,800	44,500	+7
$Sr_5Nb_4O_{15}$	-	40.5	-	-	-

quality factor (Q_u) has been normalised using the relation that $Q \times f$ is a constant in the microwave region. The variation of resonant frequency with temperature is shown in Fig.2. Since the ceramic, $Sr_5Nb_4O_{15}$ (x = 5) did not show resonance sharp enough for measurement, the values of Q_u and τ_f are not

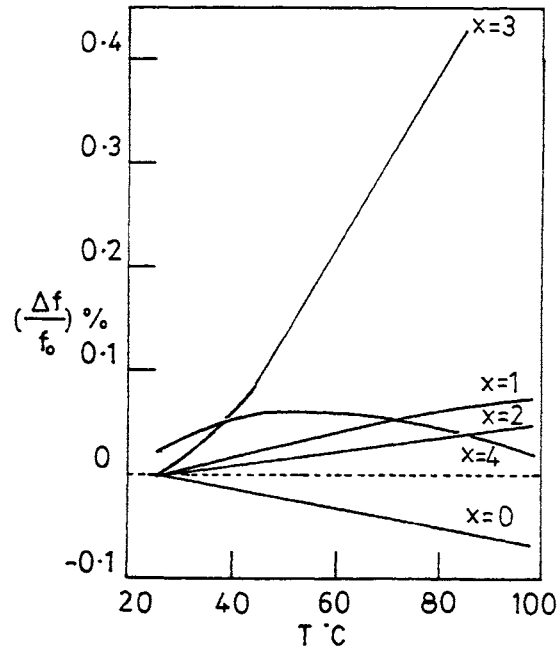


Fig.2. Variation of f_0 with temperature.

obtained. From Table I, it is clear that ε_r and Q_u increase with strontium substitution showing maximum at x = 3. Table II shows the preparation temperature and duration and

Table II. Characteristics of $Ba_{5-x}Sr_xNb_4O_{15}$ ceramics

Material	Calcination		Sintering		Bulk density g/cc
	Temp.	Time	Temp.	Time	
$Ba_5Nb_4O_{15}$	1200	4 hrs	1420	2 hrs	6.07
$Ba_4SrNb_4O_{15}$	1250	4 hrs	1420	4 hrs	5.64
$Ba_3Sr_2Nb_4O_{15}$	1250	4 hrs	1420	4 hrs	5.44
$Ba_2Sr_3Nb_4O_{15}$	1320	4 hrs	1380	4hrs	5.62
$BaSr_4Nb_4O_{15}$	1320	4 hrs	1380	4 hrs	5.46
$Sr_5Nb_4O_{15}$	1350	4 hrs	1450	4 hrs	-

the bulk density. The typical microstructure is shown in Fig.3. The $Ba_5Nb_4O_{15}$ ceramic was thermally etched and the surface was analysed using Scanning Electron Microscope (SEM). It has been reported earlier [6] that the $Ba_5Nb_4O_{15}$ has

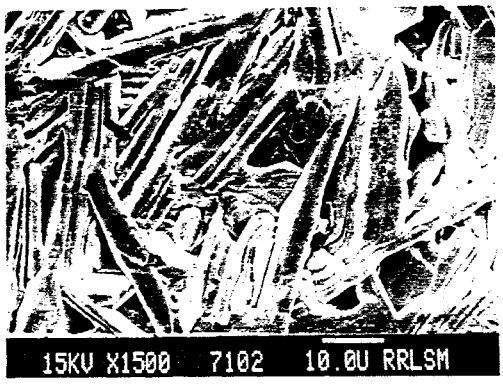


Fig.3. SEM photograph of $Ba_5Nb_4O_{15}$ ceramic.

a τ_f of +120 ppm/°C. The measurement of τ_f was done in a cavity which is 10 times larger than the sample. Moreover, the ceramic had a low quality factor and hence it was strongly coupled to the microwave probe to carry out the τ_f measurement. Hence the τ_f now obtained is much less than that and is more accurate since the ceramic is loosely coupled. For all microwave dielectric property measurements weak coupling is given [1,2,10].

Acknowledgements

The authors are thankful to Dr. A.D. Damodaran and Prof. K.G. Nair for their constant help rendered in this programme. H. Sreemoolanadhan is thankful to CSIR, Government of India for the Research Fellowship.

References

1. D. Kajfez and P. Guillon, 'Dielectric Resonators', Artech House, Massachusetts, (1986).

2. L.A. Trinogga, G. Kaizhou and I.C. Hunter, 'Practical Microstrip Circuit Design', Ellis Horwood, (1991).
3. J.K. Plourde and C.L. Ren, IEEE Trans. Microwave Theory Tech., **MTT-29(8)**, 754, (1981).
4. K. Wakino, K. Minai and H. Tamura, J. Am. Ceram. Soc., **67(4)**, 278, (1984).
5. H. Sreemoolanadhan, Johny Isaac, M.T. Sebastian, K.A. Jose and P. Mohanan (communicated to Ceram. Int.).
6. H. Sreemoolanadhan, Johny Isaac, Sam Solomon, M.T. Sebastian, K.A. Jose and P. Mohanan, Phys. Stat. Sol., **143(1)**, K45, (1994).
7. B.W. Hakki and P.D. Coleman, IRE Trans. Microwave Theory Tech., **MTT-8**, 402, (1960).
8. W.E. Courtney, IEEE Trans. Microwave Theory Tech., **MTT-18(8)**, 476, (1970).
9. Aps Khanna and Y. Garault,IEEE Trans. Microwave Theory Tech., **MTT-31(3)**, 261, (1983).
10. Y. Kobayashi and M. Katoh, IEEE Trans. Microwave Theory Tech., **MTT-33(7)**, 586, (1985).

MICROMINIATURE CRYSTAL UNITS AND TEMPERATURE SENSORS AT 150 - 1000 kHz

M.E. Bogoush, V.B. Grouzinenko

Scientific Research Institute "Phonon"
44 Krasnobogatyrskaya str., Moscow 105023, Russian Federation

Summary

Crystal units with bar piezoelements vibrating on torsion mode were the subject of wide theoretical and experimental investigations allowed to design and put into production microminiature crystal units 150-1000 kHz having high electrophysical parameters and characteristics: quality factor Q up to 10^5, spurious resonance attenuation down to 30 dB, stability to external mechanical effects, low dynamical inductance L_1 down to 10 H/mm, low dynamical resistance down to 1 kOm, low capacitance ratio down to 450, low aging, low input and matirial consumption, mass $\approx$ 4 g, volume $\approx$ 0,15 cm^3. Calculation results showed that quartz crystal units may have both high temperature stability and linear frequency-temperature characteristics by changing the cut angle and across dimensions relations of single rotated piezoelements vibrating on torsion mode.

The article being presented contains the results of theoretical and experimental investigations of microminiature crystal units vibrating in torsion mode at 150-1000 kHz. The torsion theory was devised by B.Saint-Venant and W.Foigt [1,2]; the overview of initial investigations of torsion mode crystal units was done by W.Cady [3]; the first crystal units for mass production were designed by P.G. Pozdnyakov and J.G. Vassin [4]; the first microcrystals were designed by V.B. Grouzinenko and A.P. Fedorkov in the Scientific Research Institute "Phonon" [5]. For microcrystals there are used bar piezoelements having the axis directed along the X-axis of quartz (cut angle yxl/ϑ°). Such orientation excludes elastic bounds of torsion vibrations with longitudinal and flexural ones. The article presents the results of calculation and analysis of temperature coefficients frequency of the 1st, 2nd, and 3rd orders (TCF), frequency-temperature characteristics (FTC), dynamical inductance dependence on the cut angle and the across dimensions relation. For designing crystal units of torsion mode it is significant to take into account piezoelectric modules, dielectric susceptibility and temperature coefficients of elastic compliances, densities and unit segments along the crystallographic axes [6].

Frequency coefficient

Calculation of the frequency coefficients of torsion vibrations requires the solution of Dirichlet problem for the stress function [7] with vanishes boundary conditions

$$s_{55}u_{yy} - 2s_{56}u_{yz} + s_{66}u_{zz} = -2,$$

where s_{55}, s_{56}, s_{66} - elastic compliances. Hence the frequency coefficient N is calculated

$$N = f \cdot \ell = \frac{1}{2}\sqrt{\frac{C}{\rho J}},$$

where f- frequency,
l - bar length,
J - polar moment of the cross-section inertia,

$$J = \frac{s b^3}{12}\left(s^2 + b^2\right),$$

s - bar dimension along Y,
b - bar dimension along Z,
C - torsional stiffness,
$C = 2\oint u dS$, where the integration is fulfilled over the area of the bar cross-section.

The more the piezoelement length is and the less the influence of the leads and solder doze is, the better results of calculating frequency coefficient [8] correspond to experimental data.

Frequency-temperature characteristics

Temperature dependence of crystal unit frequency is expressed by using the power series.

$$FTC(T) = \frac{f(T) - f(T_1)}{f(T_1)} = \sum_{n=1}^{3} TCF^{(n)} \cdot (T - T_1)^n,$$

where f(T) - frequency at the temperature T,
TCF(n) - temperature coefficients of the frequency of "n" -orders.

Temperature coefficients of the frequency are calculated on the basis of three points of the frequency-temperature characteristic being T, 2 T, 3 T apart from the temperature T1:

FTC1 = FTC (T1 + Δ T),
FTC2 = FTC (T1 + 2$\cdot\Delta$ T),
FTC3 = FTC (T1 + 3$\cdot\Delta$ T).

The solution of the linear equation system gives formulas for temperature coefficients:

$$TCF^{(1)}=\left(3\cdot FTC1-\frac{3}{2}\cdot FTC2+\frac{1}{3}\cdot FTC3\right)/\Delta T,$$

$$TCF^{(2)}=\left(2\cdot FTC2-\frac{5}{2}\cdot FTC1-\frac{1}{2}\cdot FTC3\right)/\Delta T^{2},$$

$$TCF^{(3)}=\left(FTC1-\frac{1}{2}\cdot FTC2+\frac{1}{6}\cdot FTC3\right)/\Delta T^{3}.$$

Extremum temperatures defined by the condition $\frac{d\,FTC(T)}{dT}=0$ are calculated from the formula

$$T_e = T1 - \frac{TCF^{(2)} \mp \sqrt{(TCF^{(2)})^2 - 3\cdot TCF^{(1)}\cdot TCF^{(3)}}}{3\cdot TCF^{(3)}}$$

Fig. 1 presents theoretical and experimental data relating to the dependence of the FTC extremum upon the across dimensions relation of piezoelements of yxl/45° cut angle. Experimental data are from [9]. It may be seen that the more the bar length and the less the influence of the leads and the solder dots are, the less the calculation error is.

The dependence of TCF the 1st, 2nd, and 3rd orders upon the cut angle and the across dimensions relation of quartz bars are given in Fig.2,3,4. Analogous calculation was done in [10,11] without taking into account b_{eff}, which makes it actual only for two cut angles with b_{eff}. FTC of bars having the the cross-section close to rectangular one ($a/b=0,9-1,0$) are practically quadratic parabolas. Their TCF depend upon the cut angle insignificantly,

$TCF^{(1)} \approx TCF^{(3)} \approx 0$ and $TCF^{(2)} \approx (-0,1 \div 0,3)\cdot 10^{-8}\,°C^{-2}$.

The $TCF^{(1)}$, $TCF^{(3)}$ may not have simultaneously zero meanings at any combination of the cut angle and across dimensions relation. So, quartz piezoelements of yxl/ cut angle with a rectangular cross-section have no FTC of the cubic parabola configuration over the room temperature range.

Calculation results showed that it is impossible to improve the temperature stability by changing the cut angle and across dimensions relation of single-rotating piezoelements in comparison with AT cut angle.

The calculation points to a possibility of designing a thermosensor with a quartz crystal unit having a linear TFC. But such characteristic has, like that for the AT angle $\approx 35°$ and the relation $a/b \approx 1$, if $TCF^{(2)}$ and $TCF^{(3)}$ of such a bar simultaneously come to zero and $TCF^{(1)} \approx 30\times10^{-6}\,°C^{-1}$.

At present several quartz cuts are known, harrying a linear TFC: LC-cut and its modification FY (for crystal units with thickness modes) and AT-cut for contour mode crystal units [12-14]. The quartz bar presented is notable for it belongs to single-rotated cut group, known as being more convenient in production.

Dynamical parameters.

The inductance L_1 of torsion mode crystal units is in direct proportion to the piezoelement length $L_1 = K_L \cdot l$, where K_L - inductance coefficient depending upon the cut angle, across dimension relation and the electrode parameters.

The calculation of K_L, for piezoelements with electrodes covering practically in whole the general surfaces (division of the electrodes is along the edges) and the comparision with the experimental data are given in [15]. The good accordance of the experimental data to the calculated ones allowes us to analyze dynamical inductance by using Fig.5. The minimum meaning of the dynamical inductance coefficient reaches 11 H/mm for bars of a rectangular cross section of the cut angle $\approx 45°$. For cut angles close to 0° and 90° it is reasonable to use the electrode division along the main surfaces of the piezoelement [16].

The quality factor Q of microcrystals as, rough values equal 10^5 and strongly depends upon the strictness of centering the leads, upon the solder dose, the placement of soldering the attenuation needles, the gas pressure inside the crystal unit, the quality of processing the surfaces of crystal elements. Experiments showed a high temperature stability of R_1 and Q.

It follows from the formulas:

$$R_1 = \frac{\omega L_1}{Q},\quad \omega = 2\pi f = \frac{2\pi N}{l},\quad L_1 = K_L \cdot l$$

that $$R_1 = \frac{2\pi N K_L}{Q}.$$

Hence, variation of the resistance over the frequency range are due to the quality factor value scatter. From approximate relations for bars of a cross-section close to a quadratic one ($N \approx 1,0 \cdot 10^3$ kHz mm, $K_L \approx 10$ H/mm) we get the evaluation of the resistance $R_1 \approx 10^7/Q$ kOhm.

Dynamical resistance of torsion mode microcrystals reaches 1 kOhm over the whole frequency range.

Spectral characteristics of the crystal units don't contain any spurious resonances except fundamental frequency overtones. The calculation of the relative attenuation A of the spurious mode was carried out by using the formula from [17]:

$$A = 20\,lg\,\frac{\omega_n L_n}{\omega_1 L_1},$$

where ω_n, L_n, ω_1, L_1 - angular frequencies and inductances of the spurious and fundamental (n=1) resonances accordingly, from which and from the formula of [15]

$$L_n = \frac{\rho l s b (s^2+b^2)}{6\left[SD\,dS_e\right]^2 \sin^2\frac{n\pi l_e}{2l}},$$

one may derive a formula for calculating the relative attenuation of torsion vibration overtones:

$$A = 20\,lg\,\frac{n\sin^2\frac{\pi l_e}{2l}}{\sin^2\frac{n\pi l_e}{2l}},$$

where $l_э$ - electrode length.

The attenuation of the fundamental torsion resonance of the partially plated piezoelement relative to the wholly plated one may be calculated by using the expression:

$$A = -40 \lg \sin\frac{\pi l_э}{2l}.$$

In Fig.6 the overtone attenuation for torsion mode piezoelements versus the relative electrode dimension $l_э/l$ is given. Partial electrode with $l_э/l \approx 0{,}78$ allows to attenuate all the overtones by 20 dB. Experiments showered stronger attenuation values (by $\approx$ 30 dB) for spurious resonances $A=20\lg R_n/R_1$ owing to higher quality factor of the fundamental resonance.

The resonance spacing of torsion mode microcrystals have values over the range 0,08 - 0,1%, its maximum meaning is possible at $l_э/l=0.75$ and its dependence on the electrode length is defined by the multiplier:

$$\frac{\sin^2\frac{\pi l_э}{2l}}{\frac{\pi l_э}{2l}} \quad [11].$$

The parallel capacitance of microcrystals lies in the range 3-10 pF.

The aging of torsion mode microcrystals is less as compared with that of analogous crustal units using microtuning-fork piezoelements since their masses are more.

Such characteristics as low labour input and material consumption, mass $\approx$ 4g, volume $\approx$0,15 cm^3, stability to external mechanical effects make torsion-mode crystal units much more preferred than contour-shear crystals over the frequency range 150-700 kHz and in some cases than thickness-shear crystals over the range 500-1000 kHz. Torsion mode crystal units have significantly higher temperature stability and mechanical strength than crystal units vibrating in longitudial modes.

Perspective possibilities appearing in connection with the application of torsion-mode crystal units and only now discovered, condition the actuality of developing computerized methods of designing crystal unit types in question.

References

1. B.Saint-Venant. Memoire sur la torsion des prismes. Sciences math. et phys. 14, Paris, 1856, 233-560.

2. W.Voigt. Lehrbuch der Kristallphysik. Ten , Leipzig, 1910.

3. W.Cady. Piezoelectricity Dover, N.Y.-L., 1946.

4. P.G.Pozdnyakov, J.G.Vassin. Quartz crystal units vibrating in torsional modes. Reports of Academia of Sciences USSR, 1969, v.185, N4, p.809-812.

5. V.B.Grouzinenko, M.I.Yaroslavskyi. Quartz microcrystals. "Electronic Industry", 1982, v.4(110), p.6-7.

6. R.Bechmann, A.D.Ballato, I.I.Lukaczek. Higher order temperature coefficients of the elastic stiffnesses and compliances of alpha-quartz. Proc. IRE, 1962, August, p.1812-1822.

7. S.G.Lekhnitzkyi. Torsion of anisotropic and unhomogeneous bars. "Science", Moscow, 1971.

8. M.E.Bogoush, V.S.Mitroshkina, V.F. Malygin, V.N.Bankov, P.G.Pozdnyakov. Temperature-frequency characteristics of crystal units with torsional modes. "Electronic technics", ser.5, "Radiodetails and radiocomponents", 1980, v.1(38), p.45-50.

9. Reference book on quartz crystal units, edited by P.G. Pozdnyakov, "Communication", 1978, p.288.

10. S.S.Chernykh, V.N.Bankov, P.G. Pozdnyakov. On the calculation of eigen-frequencies of torsional vibrations for quartz crystal elements. "Crystallography", 1971, v.16, iss.4, p. 792-795.

11. C.Herman. DT-cut torsional resonators. Proc. XXXI Ann. Symp. Freq. Contr. May 1977, p.55-61.

12. Patent 3423609 (USA). Quartz crystal temperature transducer. D.L. Hammond.

13. A.F.Fedorkov, M.I.Yaroslavskyi, I.K.-I. Fekar. Method of producing piezoelements of quartz crystal units. Patent 234054 (USSR).

14. M.I. Yaroslavskyi, A.F. Fedorkov. Low-frequency thermosensitive quartz crystal units. "Measurement techniss", 1978, N.1, p. 71.

15. M.E.Bogoush, M.I.Yaroslavskyi, N.V. Volkova, T.M. Sergeeva, A.F.Fedorkov. Calculation of dynamical inductance of quartz crystal units with piezoelements of VP-cut. "Electronic technics", ser. "Radiodetails and radiocomponents", v.1(46), 1982, p.52-55.

16. I.Zelenka. Piezoelectric Resonators and their Applications, Academia-Elsevier, 1986.

17. B.M.Bukshtam, M.E. Bogoush, P.G.Pozdnyako., S.S.Orlovskaya, E.V.Gerasimov. Influence of configuration and sizes of electrod plating upon the spectrum of unharmonic vibrations of lenz piezoelements. "Electronic technics", ser. 5, "Radiodetails and radiocomponents", v.1(26), 1978, p.10-19.

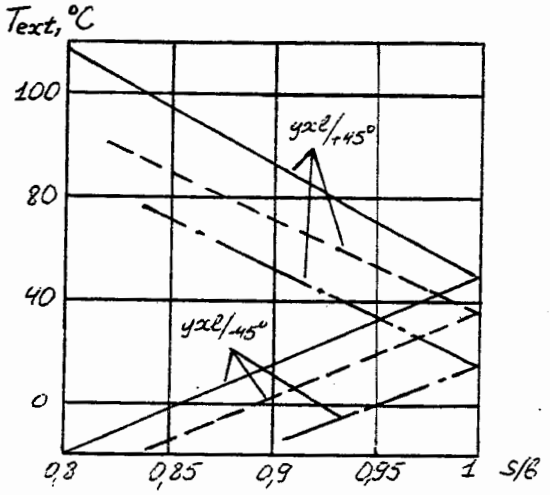


Fig.1. Temperature of extremum FTO dependence on the piezoelement across dimensions s/b
——— theory,
– – – – piezoelement b/l=0,11,
——·—— piezoelement b/l=0,22.

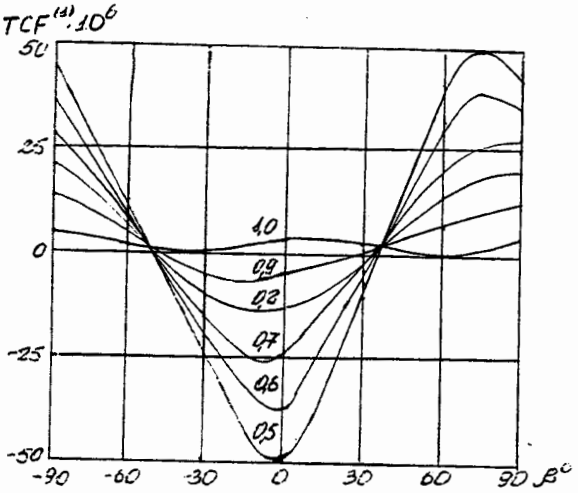


Fig.2. TCF(1) dependence on the cut angle of torsion piezoelements with different across dimensions s/b.

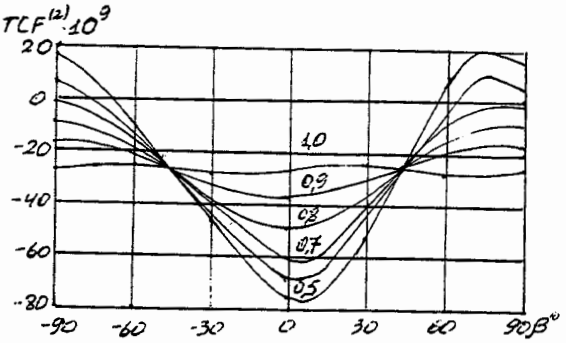


Fig.3. TCF(2) dependence on the cut angle of torsion piezoelements with different across dimensions s/b.

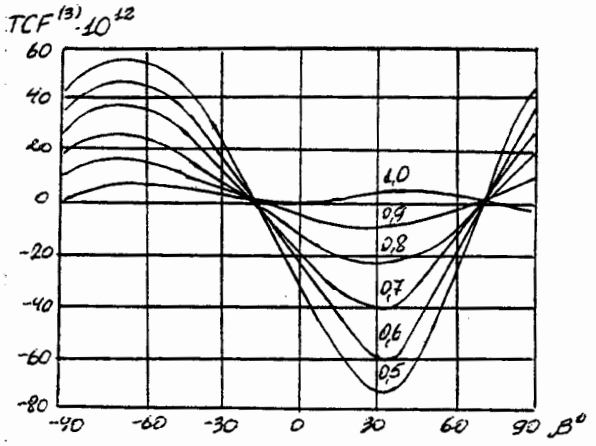
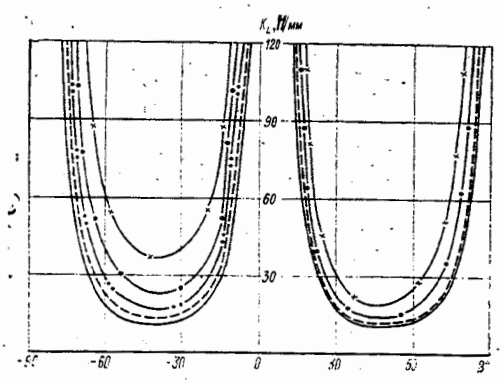


Fig.4. TCF(3) dependence on the cut angle of torsion piezoelements with different across dimensions s/b.



—x— 0,5 , —·— 0,6 , —··— 0,7,
– – – 0,8 , ——— 1

Fig.5. Dynamical inductance coefficient dependence on the cut angle of torsion piezoelements with different across dimensions s/b.

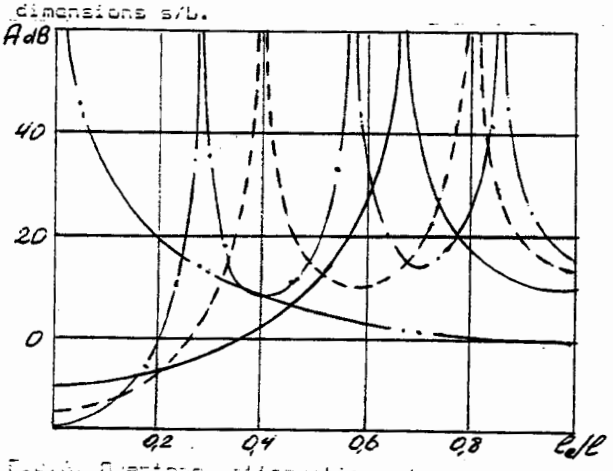


Fig.6. Overtone attenuation dependence on the electrode relative size l_0/l.
——— , – – – – , —·— attenuation overtones of the 3rd, 5th, 7th orders respectively 1st overtone, —··— the 1st overtone attenuation of partially plated piezoelement comparatively completely plated.

INFLUENCE OF CRYSTALLOGRAPHIC ORIENTATION ON FERROBIELASTIC TWINNING IN QUARTZ

P.L.GUZZO, J.J.BOY
Laboratoire de Chronométrie, Electronique et Piézoélectricité
Ecole Nationale Supérieure de Mécanique et des Microtechniques
25030 Besançon Cedex, France

ABSTRACT

The ferrobielastic twinning in quartz has been investigated in relation to the crystal orientation. A suitable mechanical stress is used to alternate electrical twin domains in X-cut and AT-cut samples of synthetic quartz. The coercive stress of each orientation was determined and are used to estimate the order of magnitude of the threshold energy for this state shift. It is shown that the optical discernment between domains is due to the piezo-optic effect. For X-cut samples, stripe-shaped patterns are observed with walls parallel to the optical axis but rather irregular domains expands in AT-cut samples.

Keywords : quartz crystal, ferrobielastic twinning, piezo-optic effect, X-cut, AT-cut.

I. INTRODUCTION

The harmful effect of electrical (Dauphiné) twins on the performances of quartz resonators was largely reported since the use of natural crystals as plate vibrators [1]. In the present time, synthetic quartz provided by various companies world-wilde are usually free of electrical twins. However, the application of quartz elements in pressure and force transducers is drastically limited by electrical twinning due to severe environmental conditions (high stress and temperature) [2]. Additionally, the nucleation of electrical twins under mechanical or thermal stresses can be observed in the manufacturing process of quartz plates [3]. These technological aspects have rised the interest in the generation of post-growth electrical twins.

Until now, theoretical considerations have been confined to the general thermodynamic theory of ferroic crystals [4, 5]. In this context, electrical twins are seen as domains of a ferrobielastic crystal. The formulation of a microscopic theory for the generation of twinned domains under appropriate stresses is still to be done. Additionally, there is no information whether electrical twinning is dependent of lattice imperfections such as dislocations and impurity-related point defects or not. Thus, more experimental works should help to provide information about this phenomenon.

The orientations of electrical twin domains are related by a 180° rotation about the trigonal axis. Consequently, the elastic compliance s_{14}, the piezo-electric coefficient d_{11} and the piezo-optic coefficient π_{14} change sign between twin states [6]. The ferrobielastic behaviour of α-quartz under compressive stress experiments have been investigated by several authors. The twin domains have been observed by using piezoelectric and/or piezo-optic effects [7, 8, 9].

In the present work, the ferrobielastic twinning is studied in relation to the crystal orientation. Stress-strain experiments were carried out on X and AT-cut samples and the force was applied in different directions. Stripe-shaped twin domains were stored by means of a CCD video camera. It is observed that the stress necessary to reverse twin states is dependent of the crystal orientation. For X-cut samples, these domains propagate first in the Z-direction and after in the Y-direction. For the AT-cut's domain patterns are more irregular. In both cases, the progress of domain patterns are often abrupt.

II. THERMODYNAMIC CONSIDERATIONS

Ferrobielasticity is a second order effect in which strain difference between orientation states is induced by applied stress [4, 5]. The switch from the original state (I) to the twinned state (II) may take place if the difference between the Gibbs-free-energy density of twin domains (ΔG) is negative. At constant temperature and zero electric field, ΔG can be denoted as :

$$\Delta G = G^{II} - G^{I} = -\frac{1}{2}\Delta s_{ijkl}\, T_{ij}\, T_{kl} \qquad [1]$$

where Δs_{ijkl} are the elastic compliance differences between the twin states, and T_{ij} the stress tensor elements. Taking into account the quartz symmetry and the symmetry operation which relates twin states, the above equation in matrix notation is given by :

$$\Delta G = 2 s_{14}\left(T_1 T_4 - T_2 T_4 + 2 T_5 T_6\right) \qquad [2]$$

where s_{14} is positive for the original state $\left(s_{14} = 4.5 \times 10^{-12}\ m^2/N\right)$. We assume that the compressive stress (P) is uniaxial and homogeneous in the faces of the crystal. We refer P to a coordinate system X" Y" Z" which is rotated by the angles φ, θ and ψ around Z, X and Y, respectively. XYZ is the Cartesian frame of the crystal. If P is applied in Z" direction, the energy ΔG as a function of the sample orientation described by φ, θ and ψ is denoted as :

$$\Delta G = 2 s_{14}\, P^2 \cos\psi \cos\theta \left(-\sin^3\psi \sin 3\varphi - \cos^3\psi \sin^3\theta \cos 3\varphi + 3\sin^2\psi \cos\psi \sin\theta \cos 3\varphi + 3\sin\psi \cos^2\psi \sin^2\theta \sin 3\varphi\right) \qquad [3]$$

In this equation, φ, θ and ψ are defined as the 1949 IRE Standard [10].

According to the principle of minimal energy, ferrobielastic twinning will occur when a threshold condition is achieved. Domain II will replace domain I only when ΔG becomes negative. As s_{14} is positive, the sign of ΔG is given by the transformation factor F(φ, θ, ψ) of the stress tensor T_{ij}^{*} to T_{ij} in the crystal axes. F(φ, θ, ψ) represents the stress combination required to produce ferrobielastic switching as a function of φ, θ, and ψ.

III. THE PIEZO-OPTIC EFFECT IN X-CUT AND AT-CUT SAMPLES

The crystal property of having two indices of refraction is known as birefringence or double refraction. This property usually results from the crystal anisotropy. However, birefringence can also be produced by the application of a suitable electric field or by suitable stresses on the crystal. The variation of refractive index due to electric field or mechanical stress is called electro-optic and piezo-optic (photoelastic) effect [11].

The refractive index of a crystal is determined by the components of the dielectric impermeability tensor $[B_{ij}]$ [11]. If we neglect higher-order terms than the first, the changes in the coefficients B_{ij} under mechanical stress is writted according to the relation :

$$B_{ij} = B_{ij}^{o} + \pi_{ijkl} T_{kl} \qquad [4]$$

where π_{ijkl} are the coefficients of the piezo-optic tensor, B_{ij}^{o} these of the impermeability tensor at zero stress and T_{kl} are the stress components. In the crystallographic frame XYZ, only the diagonal coefficients B_{ij}^{o} for α-quartz are different to zero and $[B_{ij}^{o}]$ is written as :

$$\left[B_{ij}^{o}\right] = \begin{pmatrix} 1/n_o^2 & 0 & 0 \\ 0 & 1/n_o^2 & 0 \\ 0 & 0 & 1/n_e^2 \end{pmatrix} \qquad [5]$$

where n_o = 1.544 and n_e = 1.553 are the ordinary and extraordinary refractive indices, respectively [12].

The piezo-optic tensor is symmetrical with respect to an interchange of indices i with j and k with l. However, there is no symmetry with respect to an interchange of ij with kl [11]. The independent elements of piezo-optic tensor for α-quartz are $\pi_{11}, \pi_{12}, \pi_{13}, \pi_{14}, \pi_{31}, \pi_{33}, \pi_{44}$ and π_{41}. The elements that reverse their sign on twinning are π_{14} and π_{41}. The values of π_{14} and π_{41} are 0.097 and 0.320 x 10^{-12} m^2/N respectively [13].

In the present work, we are interested by the piezo-optic effect when X and AT-cut crystals are submitted to the uniaxial stress P along Z" direction. Previously, Aizu [4] determined the angular difference between twin states in the YZ plane of an X-cut sample rotated of - 45°. The angular difference Δα as a function of ψ is given by :

$$\Delta\alpha \approx \frac{2(\sin\psi)^2 \pi_{41}}{\left(B_1^o - B_3^o\right)} P \qquad [6]$$

where π_{41}, B_1^o and B_3^o are coefficients corresponding to domain I. Following the sequence reported by Aizu [4] a similar relation showing the angular difference between twin domains for the AT-cut is denoted as :

$$\Delta\alpha \approx 3.42 \frac{\sin\psi \cos\psi \pi_{41}}{\left(B_1^o - B_3^o\right)} P \qquad [7]$$

The above relations have the same order of magnitude and can be used to determine the angular difffence between twin domains when polarized light is transmitted through the plane of the sample (X or AT).

IV. EXPERIMENTAL PROCEDURE

The device used to induce almost uniaxial compressive stress in quartz samples was reported in the previous work [14]. The sample displacement in the direction of the axial stress is now measured by using four (two pairs) of TESA inductive gauge heads. Thus, the strain is determined by the average of two differential displacements.

The twinned domains was observed by polarized ligth. Between the white light source and the polarizer, a 650 nm bandpass filter was inserted in order to minimize the light intensity. An IAC 500 CCD video camera with an 90 mm focal length objective and an extension tube of 25 mm were used to observe the development of twin domains. Several stages were stored by using the PC SCOPE treatment image software.

Synthetic right-handed quartz crystals grown by Sawyer Research Products were used. The samples were cut from non-swept Z-growth region free of inclusions twins and macroscopic defects. All the faces were ground and optically polished. The parallelism of faces is between ± 2.5 µm. Table 1 shows the dimensions of X-cut and AT-cut samples. One sample of natural quartz from Madagascar was also used. The stress was applied along Z" direction to switch from domain I to domain II, according to equation [3]. The domain patterns were observed in Y' faces. The experiments were carried out with a very slow strain-rate ranged between 1 and 2 x 10^{-5} mm/s.

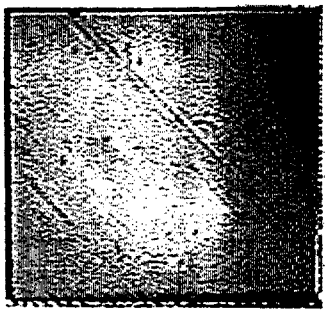
Table 1 : Orientations and dimensions of quartz samples.

Sample cut	Orientations			Dimensions (mm)		
	φ°	θ°	ψ°	X"	Y'	Z"
X-cut	90	0	-30	9.0	5.3	8.0
	90	0	-45	9.4	6.3	7.9
	90	0	-60	9.5	6.3	8.0
AT-cut	0	35	45	9.5	5.0	8.5

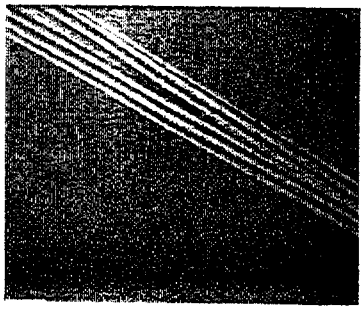
V. RESULTS

The domain patterns of X-cut samples rotated of ψ = -30°, -45° and -60° are shown in Fig. 1. It is observed that domain pattern is closely related to the crystal orientation. As it was previously reported [4,7,9], the domain patterns look like stripes oriented along the Z axis. The fractures that is observed in some pictures appeared either after an abrupt progress of domain patterns or immediately after the swich-over. Until now, we do not detect the cause of these fractures.

ψ = -30° P = 530 MPa



ψ = -45° P = 305 MPa



ψ = -60° P = 260 MPa

Fig. 1 : Optical observations of twin domains under mechanical stress for X-cut samples.

Table 2 shows the results obtained after several compressive experiments on X-cut and AT-cut samples. The domain pattern was firstly observed at a stress level P_o. This value does not necessarily correspond to the twin nucleation stress but only to the stress level in which twin domains were visualized for the first time during the first compressive cycle. It is observed that P_o varies in a large range even for samples of the same orientation. This fact seems to be associated to the stress distribution in the sample. Indeed, under normal experimental conditions, unhomogeneous uniaxial stress and sample geometrical defects play an important role in the stress level associated to twin nucleation.

Table 2 : The stresses evaluated to observe and to reverse twin domains in X- and AT-cut samples of synthetic quartz.

Sample		$F(\varphi, \theta, \psi)$	Stress (MPa)		
			N	P_o	P_r
X-cut	I.1	-0.108	2	280	650
ψ = -30°	I.2		3	430	630
X-cut	II.1	-0.250	1	505	565
	II.2		3	255	410
ψ = -45°	II.3		4	280	485
X-cut	III.1	-0.325	1	450	-
ψ = -60°	III.2		1	360	480
	III.3		1	235	485
	IV.1	-0.314	1	150	370
AT-cut	IV.2		1	185	435
ψ = 45°	IV.3		1	355	450
	IV.4		2	150	320

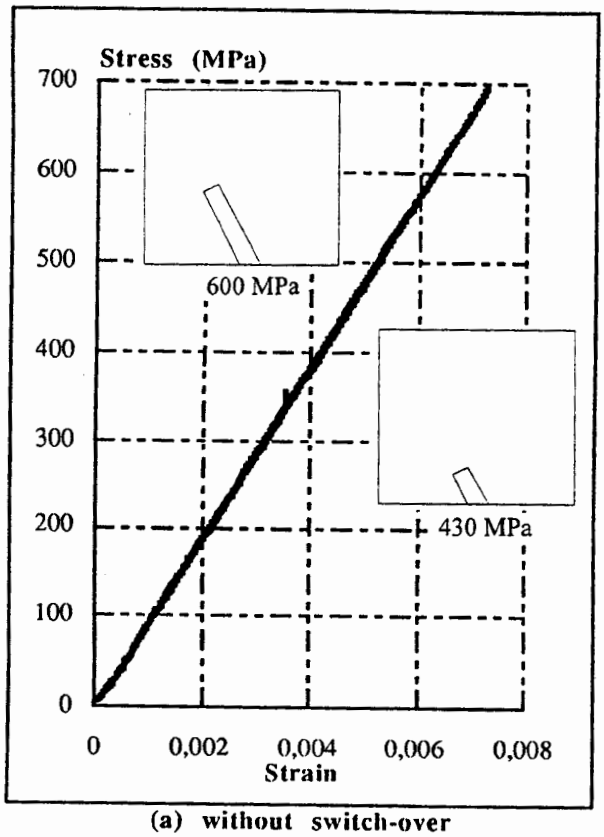
N : number of compressive cycles performed to obtain the switch-over.
P_o : stress level in which the first pattern is observed
P_r : coercive stress
- : not observed

For some samples, two or more compressive cycles were performed before attaining the switch-over. As it was previously reported [7], the stress evaluated to reverse twin domains (P_r) decreases with N. Although the samples were not submitted to the same tests, P_r seems to be dependent of crystal orientation. Except for II.1 and II.2 samples, P_r decreases as the following sequence :

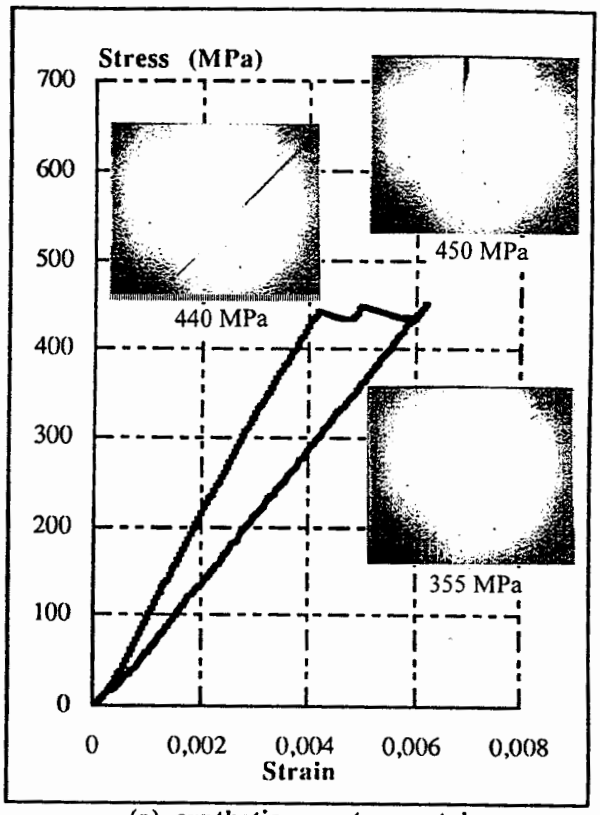
$$X(\psi=-30°) > X(\psi=-45°) > X(\psi=-60°) > (AT(\psi=45°))$$

This observation is in accordance with the previous results reported for X-cut samples [9]. As it could be expect from eq. [3], the level of P_r decreases as $F(\varphi, \theta, \psi)$ increases. Concerning the comparison between two different orientations, we think that it is necessary to take into account other factors as interatomic energy to better evaluate the coercive stress.

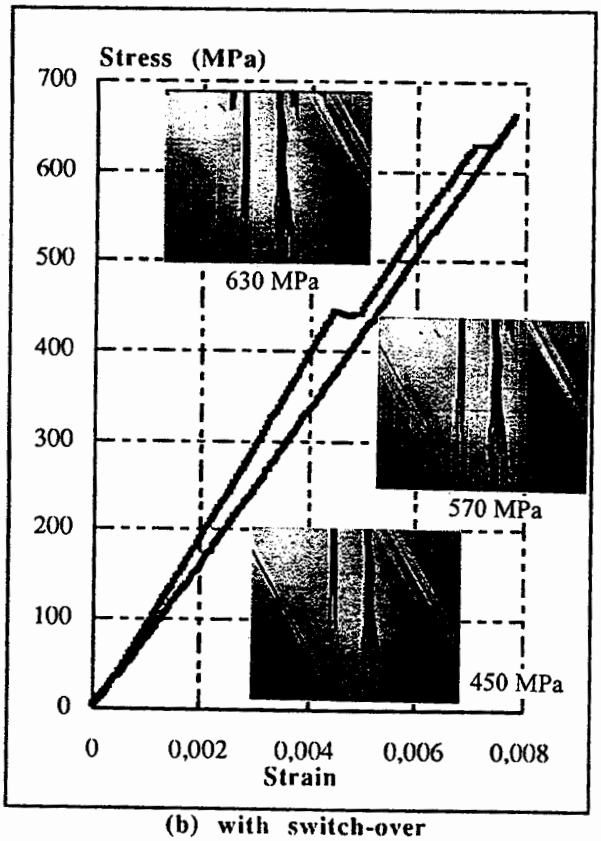
The stress-strain curves for a $X(\psi=-30°)$-cut sample are shown in Fig. 2. The two compressive cycles obtained before the switch-over did not modify the slope of the stress-strain curve. In the second cycle (Fig. 2a), one stripe was observed at about 430 MPa. This stripe did not progress until 600 MPa. In the third cycle (Fig. 2b), very clear stripe-shaped domains appeared at about 450 MPa accompanied by a discontinuity in the curve. By increasing the stress, it was observed that the advance of twin domains was abrupt and step-like, expanding in the direction of Y axis. Above 630 MPa no stripe-shaped domains was observed and the transition to the alternate domain was completed.



(a) without switch-over

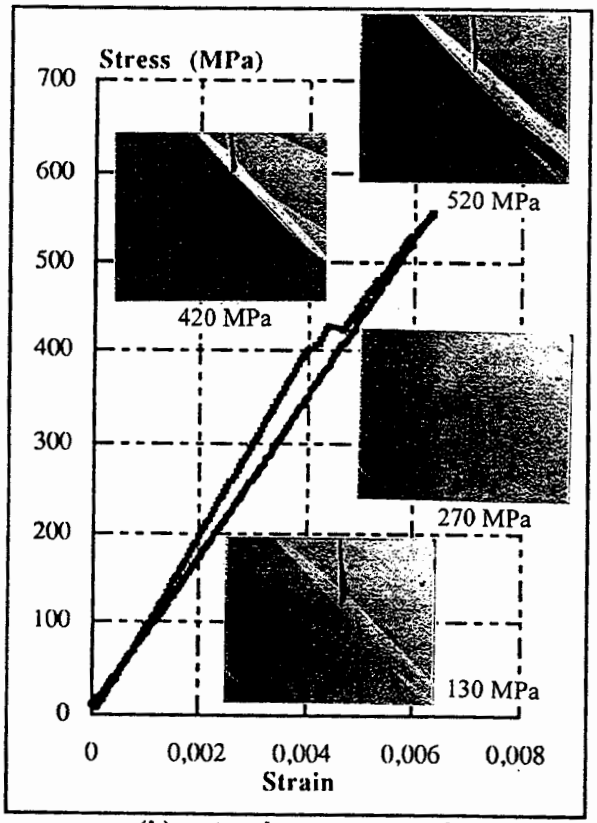


(a) synthetic quartz crystal



(b) with switch-over

Fig. 2 : Stress-strain curves and domain patterns for X-cut sample with ψ = -30°.



(b) natural quartz crystal

Fig. 3 : Stress-strain curves and domain patterns for AT-cut samples with ψ = 45°.

Fig. 3 shows the domain pattern obtained for several pressure stages on AT-cut samples. For the synthetic quartz (Fig. 3a), the domain pattern was firstly observed at 355 MPa on the right lower corner. By increasing the pressure with a slow rate, the domain pattern appeared abruptly in the central part of the sample at about 440 MPa. The stripe makes an angle of 45° in relation to the sample sides which corresponds to the direction of X-axis in the rotated AT plane. In the stress-strain curve, it is observed that an intermediate stage becomes stable between 440 and 450 MPa. The switching was completed and no more domain pattern was observed.

For the AT-cut natural quartz sample (Fig. 3b), the domain patterns were initiated at about 270 MPa. At 400 MPa stress level, the shape of domain patterns is modified and the first step of the stress-strain curve appeared. The optical contrast between domains increased with the pressure, as it was observed at 420 MPa. At this stage, the contrast in brightness near the upper right corner is similar to the brightness of the left part of the sample which was not reversed to the alternate state. The brightest region is oriented at about 45° in relation to the upper side. Above 440 MPa, the stress-strain curve seems to progress less abruptly than the synthetic sample. However, domain patterns continued to move as step-like. We started to remove the pressure before attain the coercive stress. Then, it was observed that the contrast between twin domains began to decrease. The brightest region was still observed at about 100 MPa without changing in position.

VI. DISCUSSION

The progress of twin domains under mechanical stress is not always continuous but often abrupt and step-like. It is confirmed that the ferrobielastic switching is also accompanied by an audible snap. This sound was usually heard at the instant of the swichover. However, it was also heard when the progress of domain patterns was very abrupt.

We estimated the magnitude of the threshold energy associated to the ferrobielastic twinning by using eq.[3] and the coersive stress showed in Table 2. The average value for ΔG is 5×10^5 N/m^2. It is important to emphasize that ΔG does not take into account neither the unhomogeneities of the uniaxial stress nor the imperfections of the crystal lattice. Additionally, it was previously reported that ferrobielastic twnning seems to be dependent of sample geometry [14] and the crystal past history [7]. These facts compel us to further investigate ferrobielastic twinning in relation to the stress distribution and quartz quality.

The angular difference between domains was evaluated at the strees level P_o. For both X-cut and AT-cut samples $\Delta\alpha$ is about 1.5° and is in accordance with the previous results [4]. The AT-cut twin domains are different from the well defined stripe-shaped domain patterns of X-cut samples. As it is observed in Fig. 3, the brightest region in rotated AT-cut is oriented along the X axis. Thus, similar as for X-cut samples, the brigthest zone is in the direction of one of the crystallographic axis.

ACKNOWLEDGEMENTS

One of the authors is grateful for scholarship provided by CNPq (Brazil).

REFERENCES

[1] R.A. HEISING, "Quartz crystals for electrical circuits", D. Van Nostrands. Co. Inc., New York, 563 p, 1946.

[2] E.P. EERNISSE, R.W. WARD, "Quartz resonator sensors in extreme environments", 45th Annual Symposium on Frequency Control, pp. 254-260, 1991.

[3] J.J. BOY, B. DULMET, P. CONRAUX, "Quartz crystal twinning under mechanical stress: Finite elements analysis", 6th European Frequency and Time Forum, pp. 323-327, 1992.

[4] K. AIZU, "Second order ferroic state shifts", J. Phys.Soc. Japan, 34 (1) pp. 121-128, 1973.

[5] R.E. NEWNHAM, "Domains in minerals", Am. Miner., 59, pp. 906-918, 1974.

[6] R.E. NEWNHAM, L.E. CROSS, "Symmetry of secondary ferroics II", Mat. Res. Bull., 9, pp. 1021-1032, 1974.

[7] T.L. ANDERSON, R.E. NEWNHAM, L.E. CROSS, "Coercive stress for ferrobielastic twinning in quartz", 31th Annual Symposium on Frequency Control, pp. 171-177, 1977.

[8] E. BERTAGNOLLI, E. KITTINGER, J. TICHY, "Ferrobielastic hysteresis in α-quartz", J. Appl. Phys., 50 (10), pp. 6267-6271, 1979.

[9] J. MARKGRAAFF, "Elastic behavior of quartz during stress induced Dauphiné twinning", Phys. Chem. Minerals, 13, pp. 102-112, 1986.

[10] "Standards on Piezoelectric Crystals", Proc. I.R.E., 37, pp. 1378-1395, 1949.

[11] J.F. NYE, "Physical Properties of Crystals", Oxford University Press, Oxford, 329 p., 1985.

[12] E.A. WOOD, "Crystals and Light", Dover Public. Inc., New York, 156 p., 1977.

[13] LANDOLT-BÖRNSTEIN, "Tabellen", New Serie Bd III/1, 134, Springer, Berlin, 1969.

[14] 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", 8 th European Frequency and Time Forum, pp. 802-810, 1994.

The effect of mass-loading on the frequency-temperature characteristics of fundamental AT-cut quartz resonators.

Irina Mateescu, Jiri Zelenka*, Ioan Mateescu

Institute of Physics and Technology of Materials MG-7 Bucharest- Magurele R-76900, Romania
*Technical University of Liberec Halkova 6, CZ-46117, Liberec, Czech Republic

Abstract.

The dependence of the frequency-temperature characteristics of AT-cut quartz resonators on mass-loading was studied.

The measurements were carried out for fundamental thickness-shear mode of vibration using various electrode materials of various thickness.

CrAu, Au, electrodes (100, 300, 600 nm thickness) have been deposited by vacuum evaporation on AT-cut quartz plates.

The measured frequency-temperature characteristics of the resonators were compared to the theoretical ones.

I. Introduction

For frequency control applications is necessary to have the weakest possible dependence of the resonator parameters on temperature.

The temperature dependence of the resonant frequency is one of the most important quartz resonators parameters.

The thickness shear vibration of AT-cut quartz resonators provides the minimum frequency-temperature variation in a large temperature range.

Due to the conditions required for this dependence it is very important to know the influence of various factors on this parameter.

The temperature dependence of the resonant frequency of the bulk piezoelectric resonators is a function of the material constants which are temperature dependents, on thermal expansion coefficients, on coupling between the vibration modes and on the nature and thickness of the electrodes [1]. The temperature characteristics of the material constants and the thermal expansion coefficients depend on the orientation of the quartz plate [2]. The coupling between the vibrations modes is conditioned by the shape and size of the plate. The electrodes cause the mass-loading of the plate, they define the electrical conditions of the vibrations and determine small changes in the elastic properties of the plate [3].

Some authors [4,5,6,7] revealed that the frequency-temperature characteristics are dependent on mass-loading caused by the metallic electrodes deposited on piezoelectric substrate.

The electrodes define the electrical conditions of the vibration and change a little the effective elastic properties of the plate by means of which is modified the variation of the frequency with the temperature.

In this paper are presented the results of the experimental investigation of the influence of the electrode material and thickness of the electrodes on the frequency-temperature characteristics of AT-cut quartz resonators, vibrated in a fundamental thickness-shear mode.

The measured frequency-temperature characteristics are compared with the calculated curves.

The new method for the computation of the temperature-dependence of the resonant frequency developed by J. Zelenka [1], assumes a small amplitude of the resonator vibration which is added to the large strain due to the temperature change.

The temperature change induces in quartz plate relatively large strains which have to be expressed by considering the nonlinear elastic effects too. For quartz plates covered by electrodes, the temperature change also induces strains in electrodes. In the strain formula we have considered only the first order term in the series expansion of the temperature coefficients.

II. Experimental

A set of seven AT-cut quartz resonators with the frequency 5.2 MHz working on the fundamental mode of the thickness-shear vibration was used for measurements.

The cultured Sawyer 40M plan-parallel quartz plates, with 14 mm diameter, 0.33 mm thickness, were lapped with 3 μm abrasive powder and beveled with a radius R=60 mm.

The orientation angle of the plates were in the range 35°20' - 35°23'.

The CrAu, Au electrodes with 7mm diameter have been deposited by thermal evaporation in vacuum up to a thickness of 100 nm, 300 nm and 600 nm .

The plates were bonded in the holder with silver epoxy after deposition of the first layer (100 nm) and remained fixed in holders for the next electrode deposition (300 nm and 600 nm).

The frequency-temperature dependence in the temperature range 25-65°C was measured after each stage of deposition of electrodes (Au, CrAu with the thickness 100, 300, 600 nm).

The results were fitted to the theoretical curves. An extrapolation of the fitted curve was carried out in the temperature range -15°C - +85°C.

III. Results and discussion

From the frequency - temperature characteristics of the resonators deposited with Au and CrAu electrodes (100, 300 and 600 nm thickness) it is concluded that not all the resonators shows a monotonous dependence between frequency deviation and electrode thickness.

The largest frequency change in the temperature range - 20°C to +80°C and the largest first order frequency temperature coefficients are shown by most samples for the thickness of electrode 300 nm.

The resonance frequency variation in the given temperature range and the first order temperature coefficients of the resonators with thicker or thinner electrodes are smaller.

This fact is opposite to the previous studies of the influence of the thickness of electrodes on the frequency - temperature characteristics of AT - cut quartz plates.

Many authors show that the first order frequency temperature coefficient is a monotonous function of the thickness of electrodes.

Figure 1 - 4 show the experimental frequency - temperature characteristics of the resonators A and C with Au and CrAu electrodes.

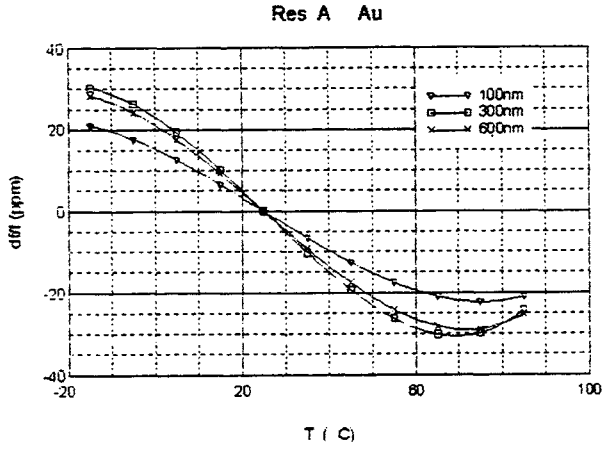


Fig.1. Resonator A with Au electrodes

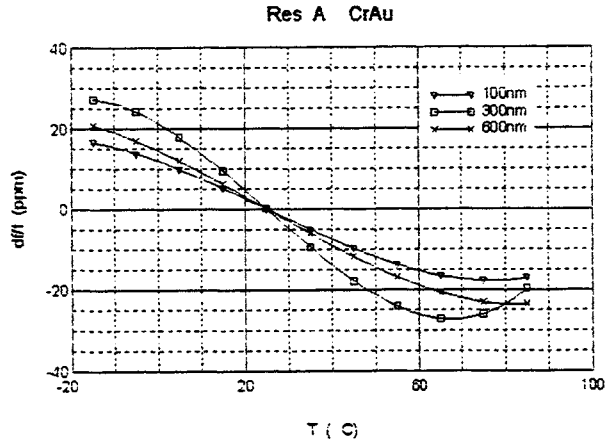


Fig. 2. Resonator A with CrAu electrodes

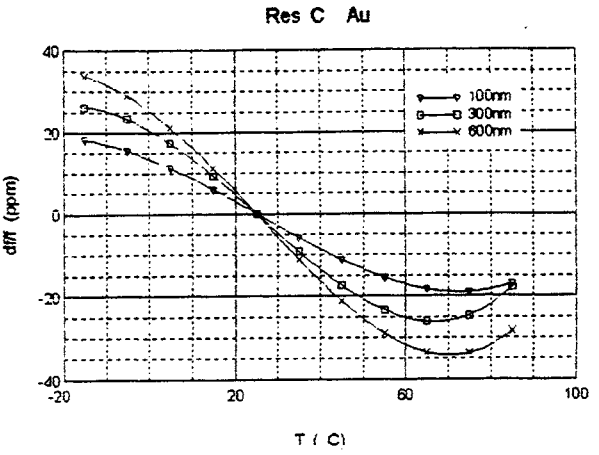


Fig. 3. Resonator C with Au electrodes

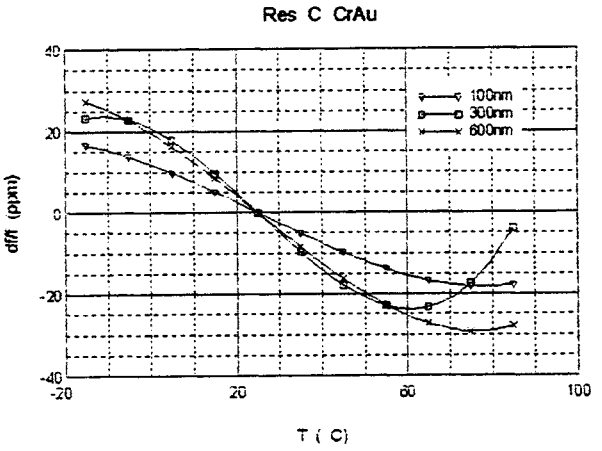


Fig. 4. Resonator C with CrAu electrodes

The dependence of the maximum frequency deviation on electrode thickness for A and C resonators is presented in Figure 5 and Figure 6.

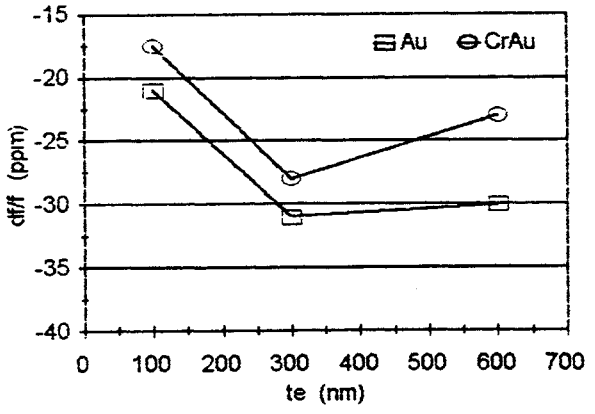


Fig. 5. Frequency deviation for resonator A with Au and CrAu electrodes

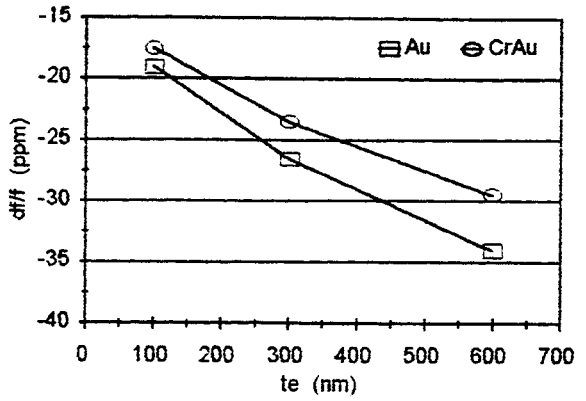


Fig. 6. Frequency deviation for resonator C with Au and CrAu electrodes

We observe that the first order frequency temperature coefficient is a monotonous function of the thickness of electrodes for C resonator and a non-monotonous one for A resonator.

To understand the various behaviour of these resonators we measured the uniformity of the electrode thickness using the microhardness method with a diamond prism indentor and the parallelism of the plates with an optimeter. The results of the measurements indicate a relation between the parallelism of the plates and the monotony of the frequency deviation curve. C resonator has a plate thickness non uniformity of 0.2 μm and A resonator has a 0.9 μm .

This observation led us to the idea of relating the frequency deviation to the appearance of a new mode of vibration[8], the most probable the flexure mode.

To explain the results of the measurements we analysed the influence of the coupling between the flexure and the thickness shear mode of vibration on the frequency - temperature characteristics of the AT - cut resonators.

The frequency spectrum of AT - cut plates for a strip of uniform thickness 2b and width 2a for the ratio a/b in the range 40 - 43 and the thickness of Au electrodes 100 nm, 300 nm and 600 nm is given in Fig.7.

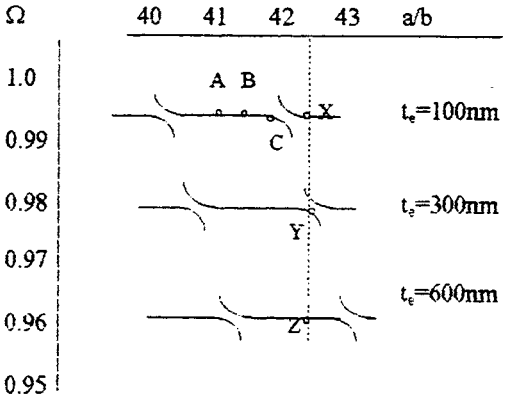


Fig.7. The frequency spectrum of AT-cut plates with Au electrodes 100nm, 300nm and 600nm thickness.

The resonant frequency - temperature characteristics computed for the points A, B and C are given in Fig. 8.

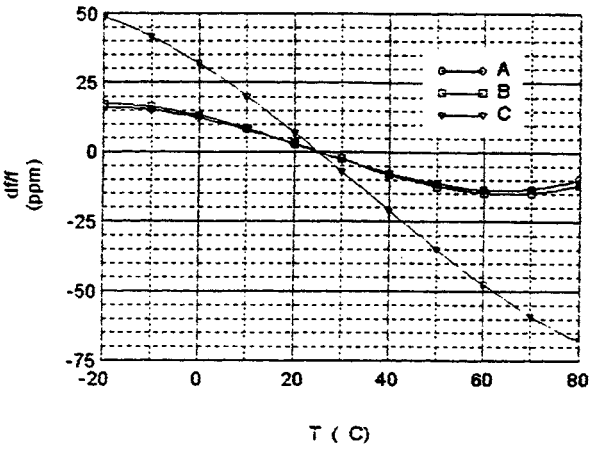


Fig.8. Computed frequency-temperature curves in points A, B, C.

Let us take AT - cut quartz strip resonators with the dimension ratio a/b = 41.15 as in our experiments.

The frequency - temperature characteristics for the points X, Y, Z are given in Fig. 9.

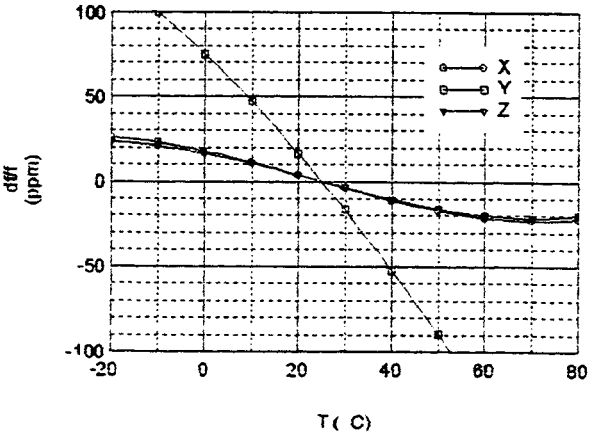


Fig.9. Computed frequency-temperature curves in points X, Y, Z..

The characteristic are similar to the frequency - temperature characteristic presented above.

A comparison between the frequency - temperature characteristics of Au electrodes resonators and CrAu electrodes shows that the maximum frequency deviation in the case of Au electrodes is higher than for CrAu electrodes. This fact is probably due to a stronger adherence of CrAu electrodes. For CrNi electrodes [9] deposited by vacuum and RF sputtering was also observed a similar behaviour of the frequency temperature dependence, which can be ascribed to the adherence effect.

For all electrode materials (Au, CrAu, CrNi) and for different mode of electrode deposition the maximum frequency deviation is a monotonous function on electrode

thickness for θ angles closed to $\theta=35^{0}15'$ and a non monotonous function for θ greater than $35^{0}22'$.

IV. Conclusions

The results obtained in this investigation and the theoretical calculations allowed to conclude that the quartz plate and electrode parameters (plate parallelism, diameter-thickness ratio) has a major influence on the frequency temperature characteristics of the resonators, especially for cut angles significantly higher than optimum AT-cut angle.

V. Acknowledgments

The authors are indebted to Dr. M. Popescu for fruitful discussions and advises. Thanks are due to D. Frăţiloiu, E. Jelea, P. Fărcăşanu and C. Mihail for experimental help. Special thanks are due to C. Bora for data processing.

References

1. J.Zelenka, The calculation of the resonant frequency temperature dependence of quartz bulk resonators, Proc. of 8th Piezoelectric Conference PIEZO - 94, Zakopane, 5 - 7 October 1994

2. R.Bechmann, Frequency - temperature angle characteristics of AT and BT - type quartz oscillators in an extended temperature range, Proc. IRE, No 48, 1960

3. J.Zelenka, Piezoelectric resonators and their applications, Academia Prague, 1986

4. A.J.Miller, Preparation of quartz crystal plates for monolithic crystal filters, Proc. 24th Annual Symposium of Control Frequency, 1970

5. J.Zelenka, The influence of electrode on the resonance frequency of AT-cut quartz plates, Int. J. Solids Structures, Vol.11, 1975

6. A.D.Ballato, T.Lukaczek, Higher order temperature coefficients of frequency of mass - loaded piezoelectric crystal plates, Proc. 29th Annual Symposium of Control Frequency, 1978

7. J.Pavlovec, J.Suchanek, J.Zelenka, The influence of the thickness of electrodes on the temperature - frequency dependence of AT - cut quartz resonators, 8th Piezoelectric Conference PIEZO - 94, Zakopane, 5 - 7 October 1994

8. I.I.Postnikov, Computer analysis of the characteristics of high - frequency quartz resonators, Proc. of IEEE Intern. Frequency Control Symposium, 1994

9. I. Mateescu, I.V.Mateescu, The effects of mass-loading on the frequency-temperature characteristics of AT-cut qurtz resonators. Contract report No.816C, Institute of Physics and Technology of Materials, Febr.1993

INTERMODULATION PRODUCTS OF QUARTZ RESONATORS AND THEIR MEASUREMENT

Jaroslav Nosek

Technical University in Liberec, Liberec, Czech Republic

Abstract

The relations used for expression of the intermodulation products of quartz resonators are described in this paper. The measuring set for the resonators vibrating at 10 MHz is described and the examples of experimental verifications are given briefly.

1 Introduction

Intermodulation products of AT cut quartz resonators were studied theoretically namely by Tiersten [1]. This paper is based on his work and defines the intermodulation current $\hat{I}_\Omega$ of frequency Ω in a circuit containing the resonator. It also derives its dependence on the resonator admittance $\hat{Y}_N$ and the value of control signals V_1 and V_2. A measuring station for intermodulation products measurement is described and received results are discussed.

2 Intermodulation of Piezoelectric Resonators

If two harmonic signal of levels V_1 and V_2 and frequencies ω_1 and ω_2 positioned uniformly with respect to resonant frequency ω_N acts simultaneously, the intermodulation signal of voltage $\hat{V}_\Omega$ and angular frequency $\Omega = 2\omega_1 - \omega_2$ is created due to the piezoelectric effect. Electrical circuit the intermodulation signal is acting in is given in Fig. 1. In the circuit R_G is the internal resistance of the generator. R_L is the load resistance the intermodulation current $\hat{I}_\Omega$ is flowing through.

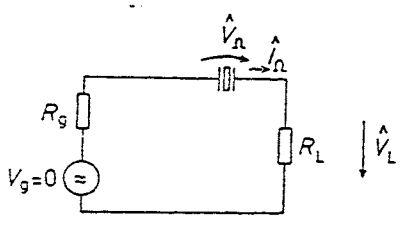


Figure 1: Circuit with intermodulation current

The complex value of the intermodulation voltage for the circuit in Fig. 1 can be given in the form according to the work [1]

$$\hat{V}_\Omega = \frac{-R^2 A_{1\Omega}^2 A_{23} A_4 + jRA_{1\Omega}A_4}{1 + R^2 A_{2\Omega}^1 A_{23}} \tag{1}$$

where $R = R_G + R_L$ and

$$A_{1\Omega} = S_e \Omega \frac{\epsilon_{22}}{2h} \tag{2}$$

$$A_{23} = 1 + k_{26}^{*2} + \frac{8k_{26}^{*2}}{N^2\pi^2}\frac{1}{\Delta_\Omega} = A_2 + \frac{A_3}{\Delta_\Omega} \tag{3}$$

$$A_4 = \frac{18\gamma k_{26}^2 \epsilon_{26}^2 V_1^2 V_2}{c_{66}^3 h^2 N^4 \pi^4 \Delta_{\Omega c}} = A_4' \frac{V_1^2 V_2}{\Delta_{\Omega c}} \tag{4}$$

$$\Delta_\Omega = \frac{\omega_N^2}{\Omega^2} - 1 \tag{5}$$

$$\Delta_{\Omega c} = \left(\frac{\Omega^2}{\omega_N^2} - 1\right)\left(\omega_N^2 \frac{\omega_1^2}{-} 1\right)\left(\omega_N^2 \frac{\omega_2^2}{-} 1\right) \tag{6}$$

The non-linear elastic module γ is given by formula

$$\gamma = \frac{1}{2}c_{22}^E + c_{266}^E + \frac{1}{6}c_{6666}^E \tag{7}$$

The circuit in Fig 1 with intermodulation current $\hat{I}_\Omega$ and frequency Ω that is flowing in it is described by equation

$$\hat{V}_\Omega + (R_G + R_L)\hat{I}_\Omega = 0 \tag{8}$$

If we calculate the current $\hat{I}_\Omega$, after simple arrangement we get an interesting relation

$$\hat{I}_\Omega = -jA_{1\Omega}\left(A_2 + \frac{A_3}{\Delta_\Omega}\right)\hat{V}_\Omega + A_4'\frac{V_1^2 V_2}{\Delta_{\Omega c}} =$$

$$= -\left(\hat{Y}_\Omega \hat{V}_\Omega + jA_{1\Omega}A_4'\frac{V_1^2 V_2}{\Delta_{\Omega c}}\right) \tag{9}$$

where

$$\hat{Y}_\Omega = jA_{1\Omega}\left(A_2 + \frac{A_3}{\Delta_\Omega}\right) \tag{10}$$

is the resonator admittance at frequency Ω. It follows from equation (9) that the intermodulation current consists of a component dependent on admittance $\hat{Y}_\Omega$ and a component affected by control signals of levels V_1 and V_2.

Formulae given above make possible the selection of a suitable method for the measurement of the intermodulation signal magnitude. As the resistance R_L is known, we calculate the intermodulation current $\hat{I}_\Omega$ from the measured voltage $\hat{V}_L = R_L\hat{I}_\Omega$. Then we substitute into equation (8) and determine the voltage $\hat{V}_\Omega$. Its value can be verified by calculation if we know corresponding material parameters, frequency distribution and levels of exciting signals.

3 Intermodulation Products Measuring Set

A measuring station schematically shown in Fig. 2 was arranged for the intermodulation products measurement.

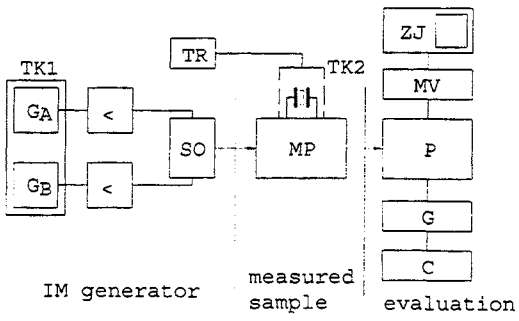


Figure 2: Measurement of intermodulation products of piezoelectric resonators

The measuring station contains generators in part IM. The used generators are crystal controlled oscillators of 10 MHz with mistuning of $\pm 1.10^{-3}$ and stability of the order $(10^{-6} \div 10^{-7})$. Generators G_A and G_B are in the thermostat $TK1$. Channel amplifiers produce output signal of controlled amplitude $(0 \div 3)$ V at impedance of 75 Ω with harmonic rejection of -38 dB at +5 dB in output.

Mixing circuit SO makes possible to increase the insertion loss of channel A with respect to generator G_B approximately to 40 dB and the second harmonic rejection increases to -85 dB. The forward direction insertion loss is approximately 1 dB.

The measuring equipment MP for piezoelectric resonator is made as a three-port. The transmission is assumed between ports $1-3$. The resonator connected to port 2 is positioned into the holder H. The crystal unit in the temperature chamber $TK2$ is hold at the reversing point temperature with an aid of thermoregulating device containing thyristor proportional regulation.

Evaluation part consists of commercial set Tesla BM 506 with measuring transmitter BM 502, generator BM 501, measuring receiver BM 503 and display unit BM 504. The set is completed by external normal of 5 MHz (precision crystal oscillator 1 PKO 13). Display unit of the set makes possible to visualize frequency characteristics in the band from 10 kHz to 30 MHz by the use of wobbulation regime. The expected levels in the range from -80 dB to -110 dB can be measured by the measuring receiver.

The key part of the control signals block is the mixing circuit that must provide a difficult co-operation of two generators operating in two very near frequencies at defined level of output voltage from each channel. The goal is the achievement of a small insertion loss in the transmission path from channel generator to the measuring circuit with resonator and on the other hand the achievement of a high insertion loss between both channels. Therefore, the mixed circuit arrangement utilizes crystal bandpass filters that improve furthermore the control signals quality by the increase of the level of rejection of harmonic distortion and increase the separation of white phase noise at generators output.

The mixing network consists of the crystal bandpass filters of centre frequencies $f_A = 9997.7$ kHz and $f_B = 10002.3$ kHz symmetrical with respect the nominal frequency 10 MHz. The 3 dB bandwidth is given by $B_3 = \pm 1$ kHz. Therefore mixing circuit selects the channel A bandwidth in range from $f_1 = 9996.7$ kHz to $f_2 = 9998.7$ kHz and channel B bandwidth from $f_3 = 10\,001.3$ kHz to $f_4 = 10\,003.4$ kHz.

Crystal bandpass filters contain simple symmetric lattice network with two resonators in connection that is equivalent to lattice network. They realize the transmission function with two poles. This type arrangement make possible to create filters of higher bandpass, high attenuation in the rejection band and high steepness of skirts of insertion loss characteristics.

To make only one transmission band, it is necessary the resonant frequency of the first branch to correspond to the antiresonant frequency of the second branch. In such a case the individual branch reactances have opposite signs. The result is that the maximum filter bandwidth is equal to twice of the difference of resonant and antiresonant frequency of one resonator. The distribution of poles and zeroes of transmission function of longitudinal and cross branch of lattice network is shown in Fig. 3.

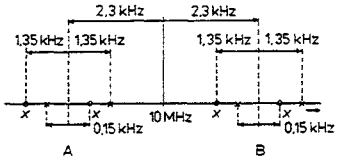


Figure 3: Positions of poles and zeroes of transmission functions of longitudinal and cross branch of lattice network type filters

.The crystal bandpass filter arrangement in the mixing circuit is shown in Fig. 4.

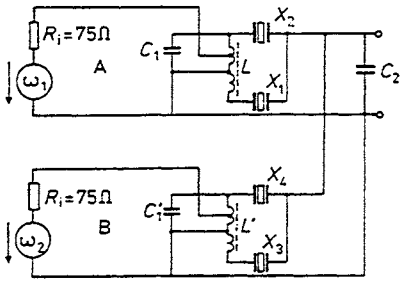


Figure 4: Crystal bandpass filter arrangement in the mixing circuit. Channel bandwidths are: A — $(9996.7 \div 9998.7)$ kHz, B — $(10\,001.3 \div 10\,0003.3)$ kHz

AT cut resonators of crystal bandpass filters have a form of plan-parallel circular plates with deposited electrodes. They are designed in order to achieve a sufficient level of suppression of anharmonic modes of thickness shear vibrations. To prevent the disturbing effects in insertion loss characteristics of the bandpass filter, the rejection at least of 40 dB is necessary.

Insertion loss characteristics of realized mixing circuit are in Fig. 5.

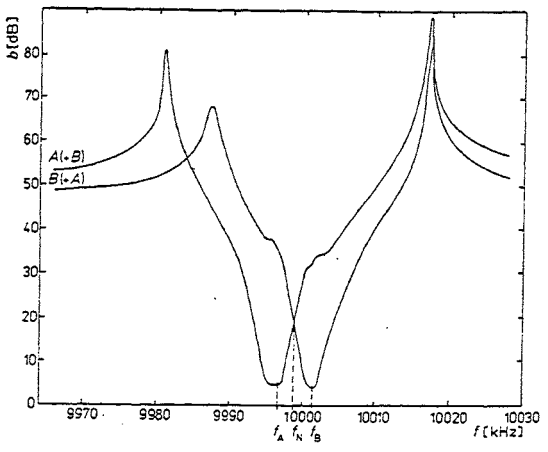


Figure 5: Measured insertion loss characteristics of mixing circuit

4 Experimental Results

Piezoelectric quartz resonators of parameters given in Table 1 were measured by measuring station in Fig. 1.

Some results as examples are given in Table 2.

Specimen	Type	Holder	Harm. order	Nominal frequency [kHz]	Cut
I1	Ú 41977	H	1	9 999.493	AT
I2	No 17	TO 8	1	9 999.765	AT
I3	403073	H glass	3	10 000.243	AT
I4	927359	H glass	3	10 000.157	AT
I5	SC 7	H glass	3	9999.897	SC
I6	SC 4	II glass	3	10 000.022	SC

Table 1: Basic parameters of measured resonators of resonant frequency of 10 MHz

Specimen	$V_{G1,2}$ [dB]	f [kHz]	f [kHz]	V_L [dB]
I1	4.5	±0.732	9997.24	-108
		±1	9996.5	-113
		±2	9993.5	-124
I4	4.5	±1	9995.35	-98
I5	4.5	+1.129	9991.52	-78
		-3.511		

Table 2: Intermodulation voltage measured at real impedance $R_L = 135\ \Omega$

5 Discussion

The measurement was performed with some problems. They were due to real properties of measuring apparatus.

First of all — frequency spectrum that was investigated at the wobbulation regime on display unit exhibited more pronounced signals in addition to basic signals f_1, f_2 of levels V_{L1} and V_{L2}. After the spectrum analysis their presence was explained by both the mirror frequencies of receiver and (it is exceptionally important) by intermodulation at input crystal filter of measuring receiver.

Further difficulty was based on the fact that namely for low levels of control signals the separation of intermodulation products level is already large and, in addition, is near the dominant control signal or its mirror image in given frequency band of input filter. At the level of V_L (supposed for example in the range from -80 dB to -100 dB) near the frequency of the signal with level of -5 dB, when the receiver sensitivity had to be considerably increased by switch, the negligible mistuning on one side of the bandpass had a consequence either the oscillation of input circuits or their damage.

In order to make the measurement easier, the suggestion was made to add further crystal filter to the measuring net. It should be between the measuring threeport and the measuring receiver. The goal of this change is to depress the level of control signals at the measuring receiver input and the use of its full sensitivity in narrow band.

The interpretation of measured results shown as an example takes some care. At the first insight, considerable separations of measured intermodulation products V_L from levels of control signals $V_{G1,2}$ are surprising.

However, the explanation follows from the measuring threeport circuit. By passing of signals of frequencies f_1, f_2 through input threeport divider a voltage decreases are made. The level of $V_{G1,2}$ decreases by the same amount. However, control voltages V_1 and V_2 at the resonator decrease, therefore also the intermodulation voltage V_Ω of frequency $\Omega = 2\omega_1 - \omega_2$ decreases. Value $V_{1,2}$ =0.805 V efficiently affects the level of the product V_Ω or V_L. For example, level $V_{G1,2} = +11.75$ dB corresponds to that level. However in a given range the measured values agree with computed ones.

Intermodulation voltage of SC specimens has not been verified by computation because the used model is valid for AT cuts. In a given case the problem is furthermore atypical due to its asymmetric distribution of frequencies f_1, f_2. We also suppose that the coincidence of intermodulation signal with some of mirror frequencies of measuring apparatus or SC cut mode took a place. In the case of such special cuts, as the SC cut is with no doubts, the unwanted nonlinear effects should be minimum.

6 Conclusion

The existence of intermodulation products of AT cut quartz resonators produced in Tesla Works, Hradec Kralove, Czech republic has proved experimentally. Their relative low value increases with increasing level of control signals and with decreasing separation from series resonant frequency. Measured values make possible to determine nonlinear forth order elastic module $c^E_{6666} = 76.10^{11}$ N.m^{-2} of AT cut quartz resonator.

References

[1] Tiersten, H. F.: J. Acoust. Soc. Amer., **57**, (1975), p. 667.

[2] Zelenka J.: Piezoelectric resonators and their use. Academia, Prague, 1983.

[3] Nosek J.: Slaboproudy obzor, **45**, (1984), p. 133 (in Czech).

[4] Nosek J.: Czech J. Phys., **A35**, (1985), p. 143-148.

The influence of surface structure on amplitude deviation and electrical parameters of quartz crystals

Eberhard Seydel , Andreas Landmann

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

1. Abstract

The influence of surface structure is described in this paper by two ways :
first - theoretical calculations of the frequency and amplitude deviation
second - experimental measurement of electrical parameters
1. The theoretical calculation describes the amplitude deviation of fundamental mode, inharmonic modes and unwanted modes for different geometrical outlines of the quartz crystal blanks. The influence on the amplitude deviation by the holder, surface roughness, various electrode materials and MESA-structures for high frequency quartz blanks can be calculated. Also the correlation between geometrical outlines and the electrical parameters (C1 and C0/C1) can be calculated.
2. The experimental part includes measurements of surface roughness and evaluation of micro-photographs. A correlation between surface roughness and electrical parameters C1, R1 and Q-value will be established. The influence on spurious suppression will be investigated. All measurements have been done on 22.9 Mhz AT-crystals. Surface roughness will be covered from coarse lapped, fine lapped to polished, as well as long time etched .

2. Theoretical background and calculation

The calculation of vibration-modes of quartz crystal resonators with different dimensions and surface contours is based on the publication of Beaver /1/.
The starting points are the differential equations

$$h(\Psi_1 + \Psi_{2,1}) - (\rho h^3 \omega^2 / D_{60})\Psi_1 = 0 \qquad (1)$$

$$(h(\Psi_1 + \Psi_{2,1}))_{,1} - (\rho k h \omega^2 / D_{60})\Psi_2 = 0 \qquad (2)$$

and the boundary conditions

$$\Psi_{1,1} = 0 \quad \text{at} \quad x_1 = \pm d/2 \qquad (3)$$

$$\Psi_1 + \Psi_{2,1} = 0 \quad \text{at} \quad x_1 = \pm d/2 \qquad (4)$$

In this way it is possible to calculate the vibrations of crystal resonators of various shapes (plano-parallel, bevelled, plano-convex, bi-convex, wedge-shaped), of different cut angles, with different electrode dimensions and at different temperatures.
The advantage of this approach over a finite element calculation is the simpler handling and the higher resolution in the considered direction /2//4/-/9/.

Therefore it is possible to find coupling modes of higher order than it would be with the FEM.
With this program variations in the amplitude distribution of each vibration mode 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 calculation works in one or three directions over the blank surface. The new version of the program takes into account the influence of the mounting system and the mesa structure of high frequency blanks. Figure 1 shows the results of calculation for a 55 MHz fundamental MESA-resonator with holder-clips for the first harmonic mode and inharmonic mode.

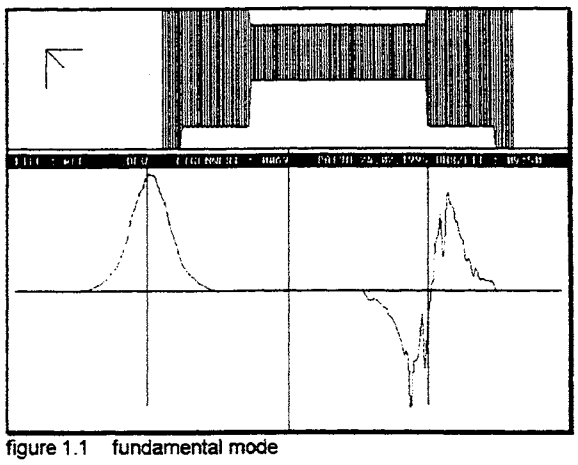


figure 1.1 fundamental mode

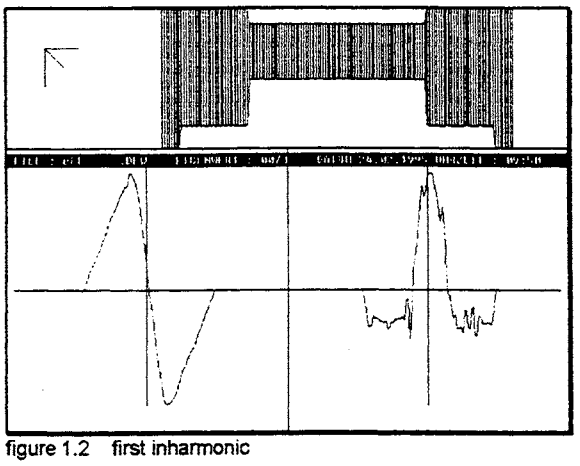


figure 1.2 first inharmonic

Another feature is the calculation of the electrical values C0 and C1 for each vibration mode. C0 is given by the electrode area and C1 by an integral over the amplitude deviation.

3. Theoretical results

These calculations show the deviation of electrical parameters versus the surface roughness or better versus the depth of the dislocations and cracks. Figure 2 illustrates the dependence of the normalized C0/C1-ratio upon the roughness.
The value of the C0/C1-ratio increases with a increasing surface roughness, the pulling range decreases for the same electrode area on the blank.

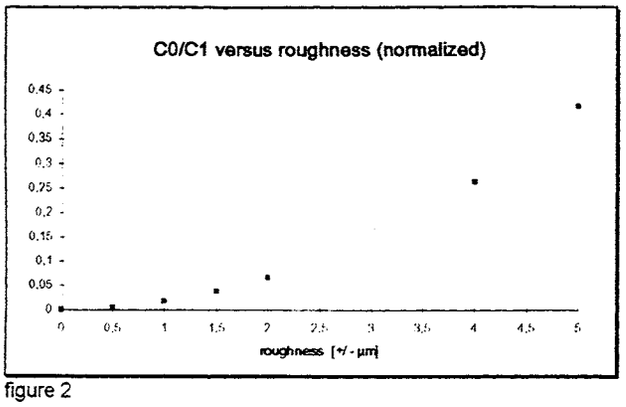


figure 2

With the surface roughness the number, the amplitudes and the frequencies of unwanted and inharmonic modes changes. See figure 3 and 4. This is caused by the scattering of the waves with different wavelengths by the rough surface.

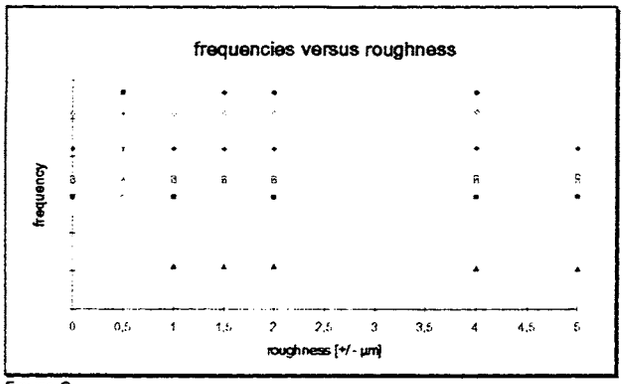


figure 3

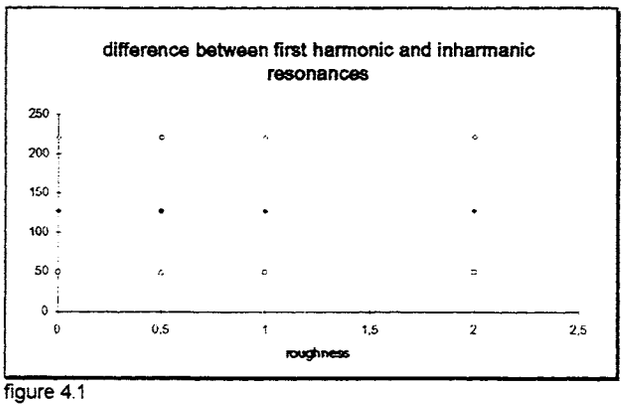


figure 4.1

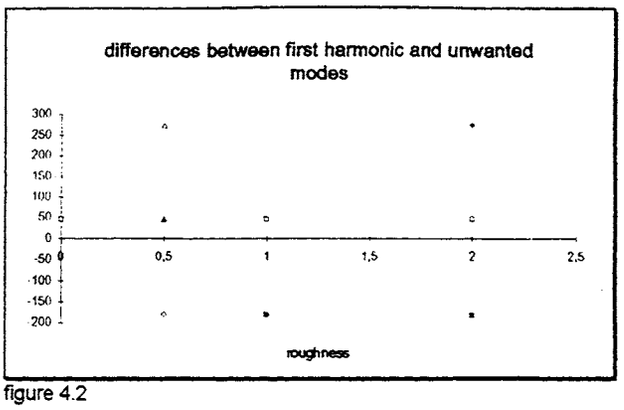


figure 4.2

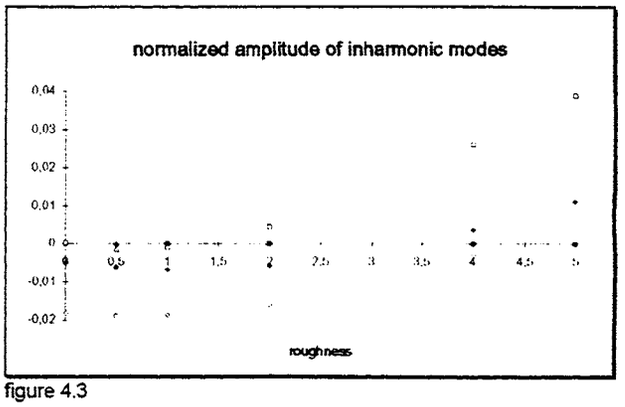


figure 4.3

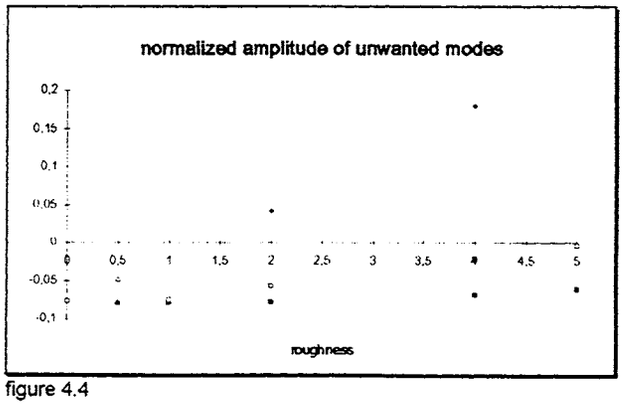


figure 4.4

4. Experimental methods

Experimental work includes investigation of surface roughness in relation to electrical data. The equipment used were the Nikon Optiphot2-Pol microscope, where surface roughness was measured by interferometry, and a network analyzer. Surface roughness varies from 0.05 to 2.75 µm, the crystal-blanks with different grain-sizes were generated by lapping and polishing and, in addition, with different etching times .

5. Experimental Results

Depending on the surface roughness of the crystal blanks, a different number of resonances (inharmonic modes) can be observed in the range of about 450 kHz above the main mode(see fig. 5).

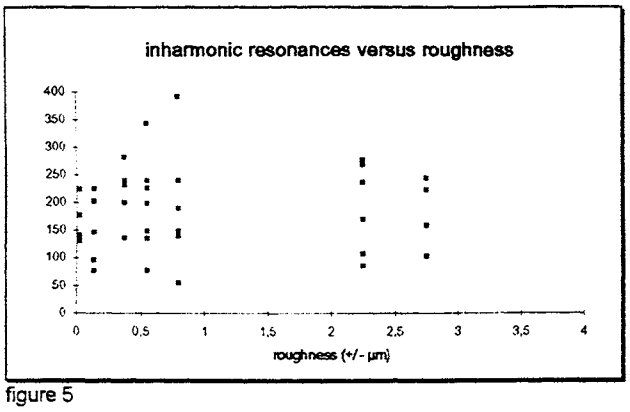


figure 5

For a surface roughness of 0.5 to 0.8µm a maximum of 7 inharmonic resonances were found. In addition the amplitude ratio of the inharmonic modes compared to the main mode varies(see fig. 6).

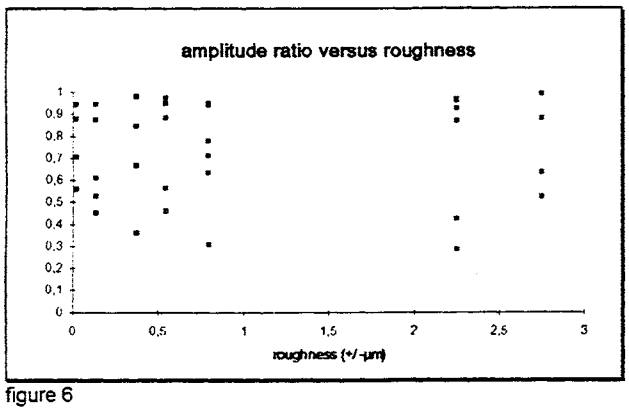


figure 6

For unetched blanks the correlation between C0/C1 ratio and spurious modes varies considerably within one production batch (see fig. 7).

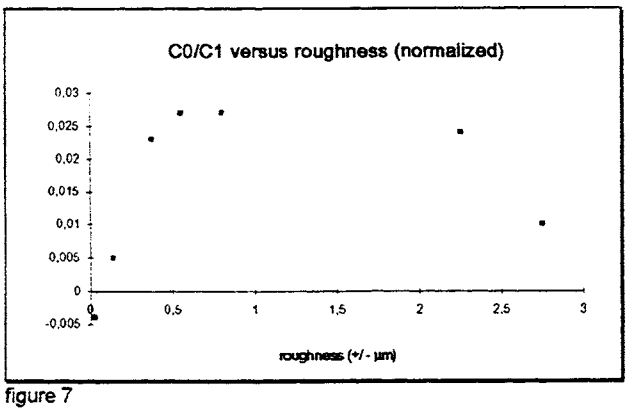


figure 7

After deep etching this spread is reduced dramatically. The reason for this effect is the large amount of surface damage (cracks, grooves) induced by the lapping process. The correlation between surface roughness and depth of cracks is not linear before etching.
The results show (Table 1) that the lapping powder with a grain size of 5 and 8µm generated different surface roughness which however resulted in almost the same depth of surface cracks.

6. Conclusions

The C0/C1 ratio of a crystal (keeping electrode dimensions constant) is not a function of the surface roughness of the blank but of the total surface damage (cracks, grooves etc.).

Comparing the calculated values with the experimental data allows us to determine the total surface damage of crystal blanks by means of the C0/C1 ratio (see fig. 8).

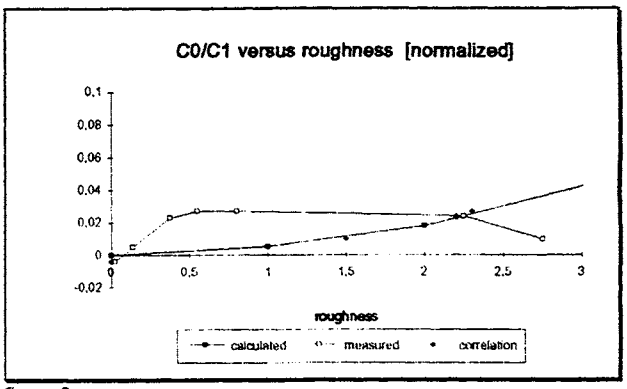


figure 8

Literature

/1/ W.D. Beaver
Design Equations for Bi- and Plano-convex AT-cut Resonators
27th Annual Frequency Control Symposium 1973 p.11

/2/ R.Lerch
Simulation piezoelektrischer Körper mit der Finite-Element-Methode
Fortschritte der Akustik -DAGA '86 p.797

/3/ J.Stoer
Einführung in die Numerische Mathematik
Springer-Verlag 1979

/4/ K.Trümpy
Störungsrechnung mit finiten Elementen angewendet auf Schwingquarze
Congrès de chronométrie '86

/5/ Yook Kong Yong, J.R.Vig
Resonator Surface Contamination
IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control 36,4(1989)452

/6/ R.C.Peach
A normal Mode Expansion for Piezoelectric Plates and certain of its Applications
IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control 35,5(1988)593

/7/ M.Naillon, R.H.Coursant, F.Besnier
Analysis of Piezoelectric Structures by a Finite Element Method
Acta Electronica 25,4 (1983) 341

/8/ D.C.L. Vangheluve
Finite Element Analysis of AT-cut Crystals
32nd Annual Frequency Control Symposium 1978 p.134

/9/ D.C.L. Vangheluve
The Frequency and Motional Capacitance of partial contoured Crystal Resonators
34th Annual Frequency Control Symposium 1980 p.412

/10/ B.Dulmet
A general Model for Quasi Thickness Vibrations in contoured Quartz Plates
5th EFTF 1991 p. 164

/11/ E.Benes, M.Gröschl, M.Schmid, W.Burger
Piezoelectric Frequency-output Sensors
8th EFTF 1994 p. 533

/12/ I.I.Postnikow
Computer Analysis of the Characteristics of High-Frequency Quartz Resonators
IEEE International Frequency Control Symposium 1994
p. 209

	polish/ optical	polish	fine lapped etched	lapped	lapped	coarse lapped etched	coarse lapped
measured (surface roughness)	0.025	0.1375	0.375	0.55	0.8	2.25	2.75
calculated (surface and cracks)	0	1	2.2	2.3	2.3	2.2	1.5
C0/C1 [normalized]	-0.004	0.005	0.023	0.027	0.027	0.024	0.01

table 1

Specific problems for quartz sensors in liquid media

C. Filiâtre*, A. Foissy**, F. Membrey**, J.-P. Valentin*, M. Valentin*

* Equipe de Physique et Electronique des Capteurs - ENSMM - Besançon
** Laboratoire d'Electrochimie des Solides - UFR Sciences - Besançon

Abstract

Quartz sensors in liquid media are used for measurements of deposited mass in situ and in real time. Impedance transformations on cascaded transmission lines yield an improved mathematical model of this quartz sensors. A quantitative knowledge of the resonator behaviour allows a proper interpretation of experimental results. This interpretation depends on the frequency at which the measurements are done : here the frequency f_m, at which the admittance modulus is maximum, is chosen. Effect of energy trapping is considered. The effective vibrating area of the quartz resonator is more important when the crystal is contacting a liquid media. To avoid this difficulty, quartz cuts with mesa shape may be used. In order to obtain practical results, the quartz resonance frequency may be measured by active or passive methods. The authors show that the active method generates mainly two inconvenients which falsify the results. Particularly for thick or viscoelastic films the passive method must be used. Experimental results are presented, essentially for copper films deposition. The range of the measurements is extented up to a thickness of 40 μm. Theoretical results are in good agreement with experimental results for all ranges of layer thickness.

I) Introduction

Sixty years ago quartz crystal users observed that the crystal resonance frequency may be modified by a slight change of the crystal mass. So broadcasting pioneers adjusted their crystals by weak pencil drawing. In 1959, Sauerbrey [1] gave the first relationship between the resonant frequency shift of a thickness-shear mode crystal and a mass deposited on the quartz crystal, operating in gas phase or in vacuum. This relation takes into account only the mass loading effect of the foreign layer. Therefore it is limited to infinitesimal films. In 1972, Lu et Lewis [2] developped a model which includes the density and the elastic shear modulus of the deposited film. Thus the resulting relationship extended the mass range.

Later, Electrochemists were interested in quartz crystal with simultaneous mass loading and liquid loading. The opinion which prevailed amid the quartz resonators experts was that the viscous damping, bound to liquid media, made to cancel crystal oscillations. Experiments have shown that quartz crystals can oscillate in liquid media and that the crystal resonance frequency is especially sensitive to mass variations. The first quartz crystal microbalance (QCM) is born from these experiments. Used in liquid media, the microbalance (LQCM) allows the measurement of deposited mass in situ and in real time. But the previous theories cannot longer be applied for quartz in liquid media.

II) Electrical and motional impedance of an electrochemical quartz microbalance : mathematical model

II.1) Impedance transformations along transmission lines.

In order to analyse and to modelize the behavior of the quartz crystal microbalance in liquid media, we propose a modelization which takes into account the properties of the deposited film and the properties of the contacting liquid. The theory is found on the analogy between the acoustic line-transmission line and on the impedance transformation of the cascaded transmission lines.

The closest analogy to microwave phenomena is to be found in sound waves where the wavelength is of the same order of magnitude as the dimensions of the apparatus in which they are used. The smallness of the wavelength made pratical the use of the distributed resonators as a replacement for the older resonant circuits composed of lumped elements. One of the most important concept in microwave is the concept of impedance and impedance transformations along the waveguide. If the load impedance Z_s of a transmission line m and the line characteristics (propagation factor γ_m and characteristic impedance Z_m) are known, the impedance Z_d at any point at distance d from the load, toward the input end of the line, may be calculated by means of the equation :

$$Z_d = Z_m \frac{Z_s + Z_m \tanh\gamma_m d}{Z_m + Z_s \tanh\gamma_m d} \qquad (1)$$

We have take the advantage of the impedance transformation concept on lines in order to propose an original analysis of the behaviour of a microbalance as a multilayered structure.

II.2) Electrical impedance Z_{LQCM} of an electrochemical microbalance

The electrical impedance Z_{LQCM} is expressed from the impedance transformation of three cascaded lines, characterized by acoustic line parameters $\gamma_m = j(\omega/\upsilon_m)$, $Z_m = \rho_m \upsilon_m$ for a layer m of thickness h_m (ρ_m is the mass per unit volume of the layer m, υ_m is the acoustic wave velocity in the layer and ω is the angular frequency). For a quartz (layer m = q) with simultaneous mass loading (layer m = f) and liquid loading (semi-infinite layer m = *l*), the input impedance Z_{IN} of the two cascaded layers f and *l* is :

$$Z_{IN} = Z_f \frac{Z_l + Z_f \tan\gamma_f h_f}{Z_f + Z_l \tan\gamma_f h_f} \qquad (2)$$

and Z_{LQCM} is expressed as [3]:

$$Z_{LQCM} = \frac{1}{j\omega C_o}\left[1 - K^2\frac{\tan x}{x}\frac{1-\frac{jZ_{IN}}{2Z_q\tan x}}{1-\frac{jZ_{IN}}{Z_q\tan 2x}}\right] \quad (3)$$

where $x = \frac{\omega h_q}{2v_q}$, $K^2 = \frac{e_q^2}{\varepsilon_q c_q}$. The quantities e_q, ε_q, c_q are the piezoelectric constant, the dielectric constant and the stiffness constant respectively, for the considered vibration mode of the crystal.

II.3) Additional motional impedance Z_{Ma} of electrochemical microbalance

When simultaneous mass loading and liquid loading affect the electrical impedance Z_{LQCM}, an additional motional impedance Z_{Ma} appears. After calculations Z_{Ma} is expressed as

$$Z_{Ma} \approx \frac{h_q^2}{4Ae_q^2} Z_f \frac{Z_l + Z_f \tanh\gamma_f h_f}{Z_f + Z_l \tanh\gamma_\gamma h_f} \quad (4)$$

where the parameter A is the effective vibrating area.

The effect of deposit thickness h_f in water on the real and imaginary parts of the motional impedance Z_{ma} has been evaluated from this mathematical model for aluminium oxide deposit, for nickel deposit and for gold deposit [3].

III) Resonance frequencies of a loaded quartz crystal

The resonance of a vibrating quartz crystal is characterized by six resonance frequencies [4]: - the frequency f_m at which the admittance modulus is maximum, - the frequency f_s at which the real part of the admittance is maximum, - the frequency f_r of zero phase and low resistance, - the frequency f_a of zero phase and high resistance, - the frequency f_p at which the real part of the impedance is maximum, - the frequency f_n at which the admittance modulus is minimum.

In air phase or in vacuum the frequencies f_m, f_r, f_s merge together. The frequencies f_a, f_m, f_p merge together also. In liquid media, all the frequencies are well distinguished because the additional motional resistance is high.

Moreover the frequencies are not only sensitive to mass loading but also to liquid properties. The three frequencies near the series resonance are affected by the liquid density and viscosity, while the three frequencies near the parallel resonance are furthermore affected by the liquid conductivity and dielectric constant and by parasitic capacitances in parallel with the static capacitance C_o [5].

For all these reasons, the interpretation of results depends on the frequency at which the measurements are done [6]. Thus the measured frequency must be chosen and specified.

IV) Energy trapping

An other experimental problem is due to energy trapping. Effect of a mass deposited on the quartz resonator surface is more important if the mass is deposited on an active area of the crystal. If the mass is deposited on a non-vibrating area, its effect is nul.

The vibrating energy repartition is not uniform, but settled between the electrodes for quartz resonators vibrating in gas or in vacuum. This effect of energy trapping is bound to the electrodes mass : the frequency resonance of the part covered by electrodes is lower than the frequency resonance of the free part. In the case of a 6 MHz quartz crystal, the difference is about 30 kHz (this result is obtained from Sauerbrey's equation [1], considering that the deposited mass corresponds to the electrode mass). Now the pass band of a 6 MHz quartz resonator is about 6 Hz, for a quality coefficient Q equal to 10^6. Under these conditions, only the covered part can vibrate when the frequency of the voltage applied to the electrodes corresponds to the resonance frequency of the crystal part covered by electrodes : the free part does not vibrate. So the mechanical vibration is located between electrodes without any coupling to the free part of the crystal. The middle zone located between the resonant part and the non resonant part corresponds to the edge of the electrode. The mechanical damping is essentially located in this middle zone. In the case of quartz resonators vibrating in gas or in vacuum, we may observe that the quartz holders are carefully dispatched on the periphery of the crystal, whereas the ideal holdering must be carried out on the middle zone corresponding to the edge of the electrode.

When a quartz resonator face is in contact with a liquid medium, the resonator quality coefficient decreases up to 10^3. In the case of a 6 MHz quartz crystal the pass band becomes about 6 kHz, when the gap between the resonance frequencies of the covered part and the free part stands about 30 kHz. Under these conditions a mechanical vibration piezoelectrically maintained in the covered part is partially coupled to the free part. The effective vibrating area of the quartz resonator is more important when the crystal is contacting a liquid media. The measurements realized in liquid media may be wrong if the effective vibrating area is not taken into account. To avoid this later difficulty, quartz cuts in mesa shape may be used (figure 1)

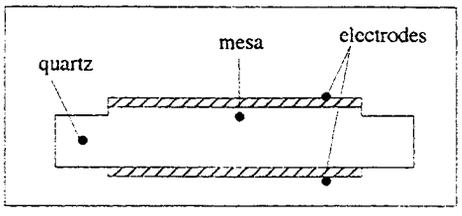


Figure 1 - manufactured cut in mesa shape

On the figure 1, the central part is covered by electrodes, and the resonance frequency corresponds to the resonance of this covered part. With few micrometers the mesa leads to a very efficient energy trapping located in the covered part. The cut may be easily manufactured by chemical etching.

V) Measurements and Results

V.1) Measurement methods

The quartz resonance frequency can be measured by two methods, the active and the passive methods. In the active method, an oscillatory sustains the quartz vibrations at the resonance while in the passive method the quartz is supplied by a sinusoïdal signal generator and the transmitted or reflected signal is analysed. Dedicated oscillatories for immersed quartz crystals have been studied [7].

Quartz crystals at resonance frequencies 2 MHz and 6 MHz are used for this study therefore two oscillatory kinds have been realized. A Colpitts oscillatory is used for 6 MHz quartz crystal having a low motional resistance R_m. But it is difficult to maintain oscillations when the motional resistance of the quartz crystal is higher than 1 kΩ. So an oscillatory with operational amplifiers (figure 2) has been used for the 2 MHz quartz crystal (R_m = 1,2 kΩ). The resistances R_1, R_2, R_4 are adjustable parameters permitting to start oscillations of the damped quartz.

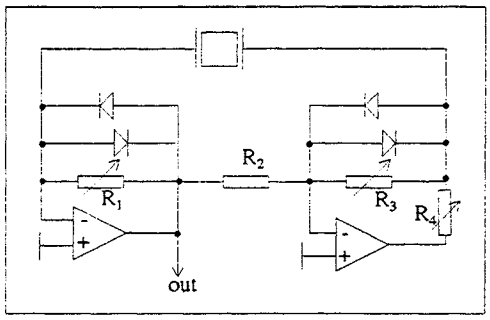


Figure 2 - 2 MHz oscillatory with operational amplifiers [8].

The advantages of an electronic oscillatory are the direct and continuous measurements of the resonance frequency and its low cost. But for damped quartz crystals, there are important limitations due on one hand to the electronic circuit and on the other hand to the experiments.

First of all, the frequency at which the quartz vibrates is not defined (f_m, f_r,...). And for mass measurements of thick films or for studies in highly viscous liquids, the working domain of the electronic circuit must be extend. In this case circuit simulations with PSPICE software show that the real and imaginary parts of the impedance (or admittance) of the amplifier circuit vary with the frequency [8]. The phase of the electronic circuit shifts with frequency. Otherwise the total phase in the loop quartz-amplifier circuit must be zero. So the quartz does not work at fixed phase and by this way the electronic circuit produces an additionnal frequency shift which is independant with any physical phenomenon.

Experiments introduce mainly two limitations to the active method. In the case of viscoelastic film deposit, the frequency shift due to the added mass cannot be distinguished from this one due to the film viscosity change and when the mass of the deposit increases, the oscillations stop.

Under these conditions, the active method for measurements of mass deposited on a quartz crystal in liquid media must be designed only for thin rigid films. For thick or viscoelastic films the passive method must be used. Moreover, this method allows the knowledge of the motional elements of the equivalent circuit and the frequency at which the measurements are done, is well defined.

V.2) Results

In the following study, all the measurements are done by passive method with a π-network and at the frequency f_m where the admittance is maximum. In the specific electrochemical cell built for measurements, the quartz crystal has one face in contact with the liquid medium and the π-network is very closed to the quartz crystal. The cell is described elsewhere [8].

The quartz crystal is an AT-cut plane parallel plate (diameter 16 mm) coated by two circular gold electrodes (diameter 9.7 mm). Its nominal resonance frequency is closed to 2 MHz. The quartz constants used are the following one : $e_q = 9.65\ 10^{-2}$ C m^{-2}, $c_q = 29.3\ 10^9$ N. m^{-2}, $\varepsilon_q = 4\ 10^{-11}$ F. m^{-1}.

Copper is deposited by galvanostatic way on the gold electrode of the quartz crystal at two current densities (10 and 20 mA/cm^2). The electrolytic bath has the following composition : H_2SO_4 0,5 M, $CuSO_4$ 0,5 M. Its density and viscosity are respectively equal to 1.15 and 1.6 cp at 20°C.

The mass of deposited copper is calculated by the Faraday's law assuming an efficiency of 100 %. During the deposition process, the frequency is measured at increasing times. The frequency shift in function of deposition time is represented figure 3 as well as the related mass of copper. The variation of the frequency f_m versus the copper mass is then deduced.

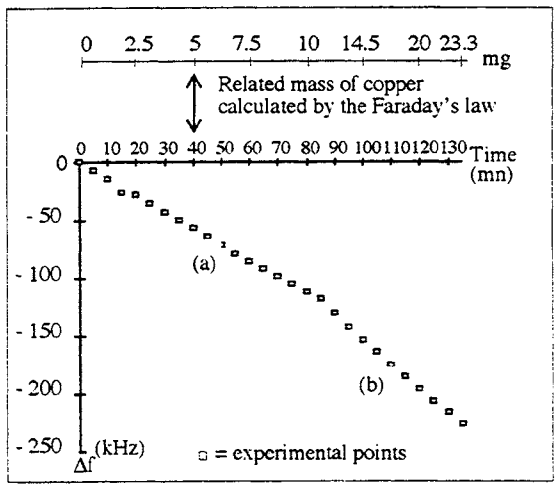


Figure 3 - measured frequency shift versus time and related mass of the deposited copper at two current densities a) 10 mA/cm^2, b) 20 mA/cm^2.

Using Mathematica Software, the theoretical modulus of the impedance Z_{LQCM} (3) is represented on a three-dimensional diagram versus the frequency and versus the deposited mass (figure 4). The bottom of the valley shows the frequency f_m at which the modulus of the impedance Z_{LQCM} is minimum, versus the increasing mass.

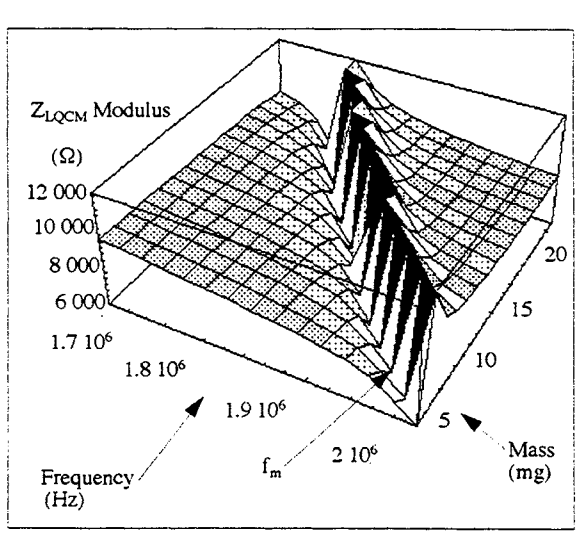


Figure 4 - modulus of the impedance Z_{LQCM} versus the frequency and versus the deposited mass.

Experimental and theoretical curves are compared on the figure 5. There is a good agreement between the two curves up to a mass equal to 15 mg. Beyond this value, a slight drift is observed and corresponds to the change of the current density. The increase of the deposition velocity probably changes the efficiency of the reaction. If a 95 % efficiency is considered all the experimental points are confused with the theoretical curve deduced from mathematical model.

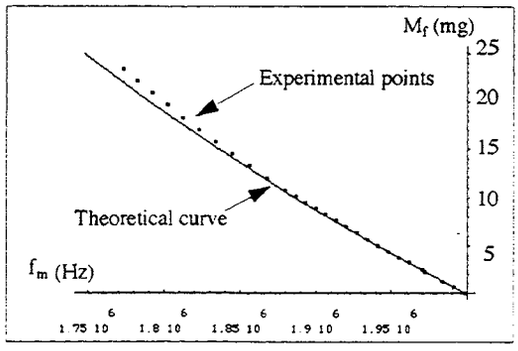


Figure 5 - mass of the copper deposit versus frequency : experimental (■) and theoretical (—).

In the previous experiment, a copper mass change up to 23 mg is measured. Assuming the copper density is 8.9 g.cm^{-3}, the corresponding thickness of the deposit equals 40 μm.

Using the model developped by Valentin and Filiâtre [3], the calculated results are in good agreement with experimental results, [8] for all ranges of layer thickness. For thin films ($h_f <$ twentieth of acoustical wavelength) the simplified equations are identical to those established by the means of other models (Sauerbrey model [1], Lu and Lewis model [2], Reed and al model [9], Crane and Fischer model [10]...).

VI) Conclusions

A mass loaded quartz crystal in liquid medium has been analyzed as one dimensional composite resonator and as three cascaded transmission lines. Impedance transformations on line have been used to obtain an analytical expression for electrical impedance and additional motional impedance with copper deposits as thick as 40 μm and with a semi-infinite contacting Newtonian liquid.

The weak quality factor (about 10^3 in liquid phase) decreases the energy trapping and leads to a more important effective vibration area of quartz resonator. Moreover it is sufficient to ensure frequency control using the passive measurement method. Specific oscillatories have been realized and several limitations must be taken into account in frequency measurement.

The proposed model is well suited for the analyzis of quartz crystal loaded by thick films, viscoelastic films, antireflecting films... It allows to distinguish changes in surface mass from changes in solution properties. Mixed terms due to the film and liquid loading appear in the mathematical model and yield to a best quantitative description of the sensor frequency dependence and sensor sensitivity for electrochemical, chemical and biological applications.

Bibliography

[1] SAUERBREY G., Verwendung von schwingquarzen zur wägung dünner schichten und zur mikrowägung, Z.Phys.,155, pp. 206-222, 1959

[2] LU C., LEWIS O., Investigation of film thickness determination by oscillating quartz resonators with large mass load, J.applied phys., 43, p. 4385, 1972

[3] VALENTIN M., FILIATRE F., Application of the impedance transformation on transmission lines to electrochemical microbalance, J. Phys. III, 4, pp. 1305-1319, 1994.

[4] IEEE Standard N° 177, 1966.

[5] BARNES C., Some new concepts on factors influencing the operational frequency of liquid immersed quartz microbalance, Sensors and Actuators A, 30, pp. 197-202, 1992.

[6] FILIATRE C., BARDECHE G., VALENTIN M., Transmission-line model for immersed quartz-crystal sensors, Sensors and Actuators A, 44, pp. 137-144, 1994

[7] BARDECHE G., FILIATRE C., VALENTIN J.P., Dedicated oscillator for immersed crystal sensor, 6th European Frequency and Time Forum, ESTEC (ESA), Noordwijk, march 1992, pp. 265-268.

[8] FILIATRE, C., Analyse et modélisation de la microbalance à quartz en phase liquide en vue de l'étude de dépôts épais, Thèse de Doctorat d'Université, 1994, Université de Franche-Comté - Besançon - France.

[9] REED C.E., KANAZAWA K.K., KAUFMAN J.H., Physical description of a viscoelastically loaded AT-cut quartz resonator, J. appl. phys. 68, pp. 1993-2001, 1990

[10] CRANE R.A., FISHER G., Analysis of a quartz crystal microbalance with coating of finite viscosity, J. Phys.D: Appl.phys.,12, pp. 2019, 1979

SIX-PORT TECHNIQUE IN REFERENCE MEASUREMENTS OF PIEZOELECTRIC RESONATORS PARAMETERS.

A. Lisowiec, M. Wójcicki
TELE-AND RADIO RESEARCH INSTITUTE
PIEZOELECTRIC DEPARTMENT
Ratuszowa 11, 03-450 Warsaw, Poland

(work sponsored by a grant nr 8S50100704 from State Committee for Scientific Research)

Abstract.

A special measurement system with six-port as a reflectometer has been developed at ITR. An analysis has been carried out on how the absolute accuracy of the measurement of basic piezoelectric resonator parameters depends on the accuracy of the immittance standards used for calibrating the six-port.

1. Introduction.

Six-port technique which has been used in microwave measurements for several years has recently been applied to the measurements of basic parameters of piezoelectric resonators. At Tele-and Radio Research Institute there has been developed a system for measuring piezoelectric resonators parameters by reflection method in which six-port is used as a reflection coefficient meter [1].

The detailed principle of the direct reflection method has been described in [1],[2],[4]; in general it consists of measuring certain characteristic points of reflection coefficient frequency characteristic in transmission line terminated with quartz resonator. The accuracy of measurement of the complex reflection coefficient Γ has direct influence on the accuracy of measurement of resonator parameters. As has been shown [1],[4], the basic measurement criterion for the series resonance frequency f is the equality of phase of the reflection coefficient Γ far from resonance region with the phase of Γ in the resonance region (possibly shifted by 180 degrees). Both phase values are measured at slightly different frequencies and moreover at different levels of the magnitude of the reflection coefficient Γ. The classical vector voltmeter used as the reflection coefficient meter introduces a subtle and difficult to evaluate nonlinear error connected with difference in levels of the reflection coefficient magnitude. As a concequence, the absolute accuracy of the measurement of the resonator parameters is hard to evaluate. When the six-port is used as reflection coefficient meter, it is possible, basing on its parameters and the parameters of the standards used for calibration to evaluate the absolute accuracy of reflection coefficient measurements. From the accuracy of reflection coefficient measurement the accuracy of the measurement of resonator parameters can be determined.

2. The measurement system.

The block diagram of the measurement system is presented in fig. 1. The principle of operation of the first laboratory version of this system has been presented in [1], [3]. Since then some efforts have been carried out to optimize the construction of the six-port circuit (enclosed by the dashed line) so as to obtain the highest measurement accuracy. The ACTIVE TYPE RMS DETECTOR, very linear by itself has been calibrated by means of precision attenuators. To make the construction more rugged, all the subsystems enclosed by the dashed line have been put into an enclosure made from one solid block of metal with one separate chamber for the SIXPORT together with BALANCED MIXERS, one for the SWITCH together with the band-pass filter BPF and ACTIVE TYPE RMS DETECTOR and one for the IEEE-488 interface circuit. The outer dimensions of this block are 80x180x220 mm. Thus the effects of electromagnetic interference have been minimized. To minimize the effects of temperature influence, the six-port block has been placed in a temperature chamber with temperature stabilized to 0.1 degree.

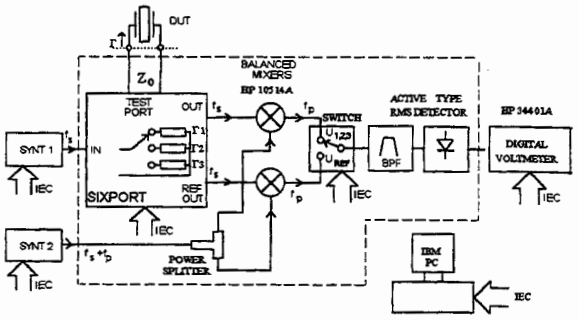


Fig. 1. The block diagram of the measurement system.

The transmittances from the input port IN to the output port OUT and from the input port port IN to the reference port REF OUT for three positions of the switch connecting three reflectors Γ1, Γ2 and Γ3 have been measured. From these transmittances the characteristic points q1, q2, q3 of the six-port over the frequency range 4-500 MHz have been calculated; they are presented graphically in fig.2. The magnitudes of the q-points lie between 1.4 and 2.5 and their phase differences, that is |arg(q1) - arg(q2)| and |arg(q1) - arg(q3)| lie between 80 and 150 deg. This confirms that the six-port circuit has been designed correctly [5].

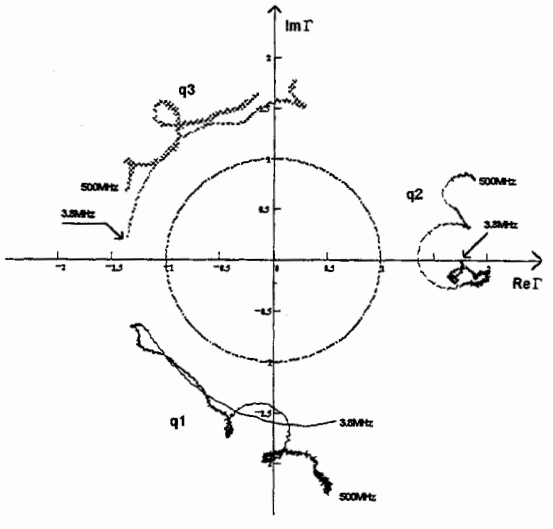


Fig. 2. The locus of the q points of the six-port in the complex Γ plane.

3. Accuracy analysis.

For the six-port circuit described above an effort has been taken to analyse how the accuracy of the immittance standards used for calibration and resolution of power measurements influence the absolute accuracy of reflection coefficient measurements. As a basis for the analysis it was assumed that the six-port circuit was calibrated by the Riblet-Hansson method [6] with 4 immittance standards of $|\Gamma| = 1$ and a match. The assumed standards were 14 mm coaxial air-line precision type. The model for the behaviour of the calibration standards is based on its longitudinal and cross-sectional dimensions and material properties. These characterising parameters are evaluated by either measurement or theoretical models and there is uncertainty in these evaluations [8]. The analysis has been carried out at the frequency 284 MHz.

In the first step it was determined how the uncertainties of the immittance standards parameters and errors in the measurement of power ratios influence the accuracy of determining the calibration matrix of the six-port. Then it was calculated how the errors of power ratios measurements influence the accuracy of determining the reflection coefficient during measurement. The calculated values of the errors versus magnitude of the measured reflection coefficient are presented in fig. 3 and 4. It should be stressed that the absolute values of the

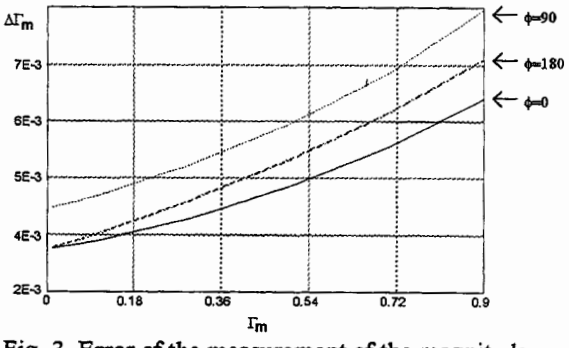


Fig. 3. Error of the measurement of the magnitude of reflection coefficient.

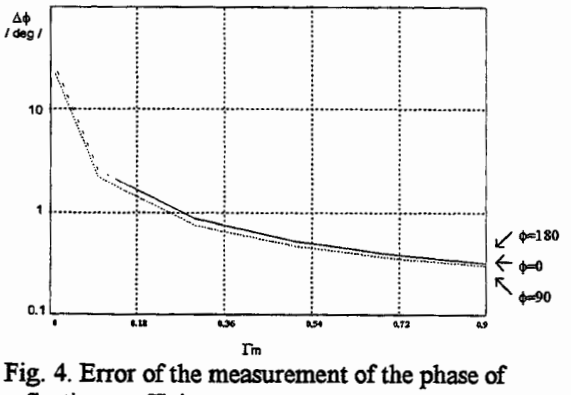
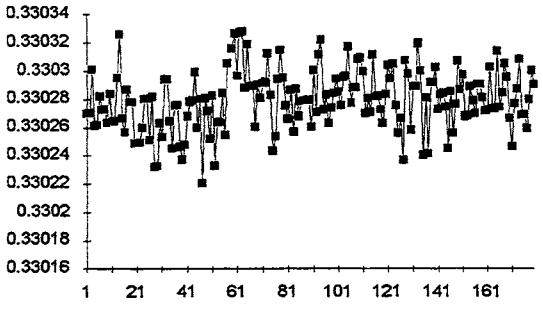


Fig. 4. Error of the measurement of the phase of reflection coefficient.

errors for the worst case condition have been

calculated. The analysis also has shown that with inaccuracies of power ratios measurement of the order of 10^{-4} (such was obtainable in the system), the main influence on the absolute accuracy comes from the inaccuracy of immittance standards used for six-port calibration. By the similar method the repeatability of the measurement of reflection coefficient has been calculated. The repeatability is 2 to 3 orders of magnitude better than the absolute accuracy. In fig. 5 a) and b) typical series of measurements of the standard immittance 100 Ohms are presented. The Allan variance for the



a)

b)

Fig. 5. A series of 180 measurements of reflection coefficient of 100 Ohm immittance standard; a) magnitude, b) phase (in deg)

magnitude is 2.2 10^{-5} and for the phase 7 10^{-3} deg.

Basing on the above analysis the absolute accuracy of determinig the series resonance frequency fs of the 284 MHz resonator can be estimated. The frequency characteristics of this resonator (magnitude and phase of reflection coefficient) are presented in fig. 6 and 7. From fig. 3 the maximum absolute error of determining the phase of

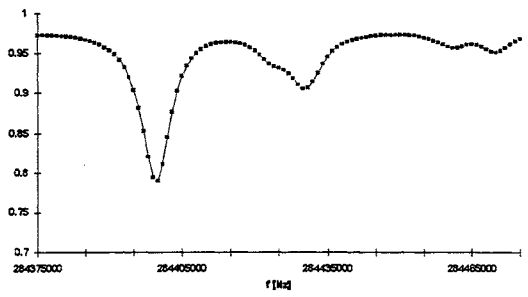


Fig. 6. Frequency characteristic of the magnitude of reflection coefficient of 284 MHz resonator.

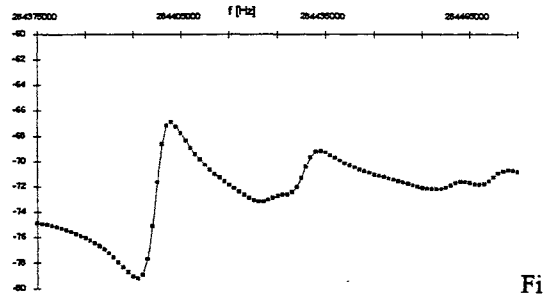


Fig. 7. Frequency characteristic of the phase of reflection coefficient of 284 MHz resonator.

reflection coefficient at this frequency is about 0.6 deg.(magnitude of reflection coefficient equal to 0.7). The reference phase (again from fig. 3) can be determined with absolute accuracy of about 0.4 deg.(magnitude of reflection coefficient equal to 1).The slope of the phase characteristics in the resonance region is about 1deg/1ppm. So in the worst case the maximum absolute error of determining fs is (0.6 + 0.4)/(1deg/1ppm) =1ppm.

The repeatability of the measurement of fs as can be seen from fig. 8 (taking into account temperature drift) is about 10^{-9} which is 3 orders of magnitude better.

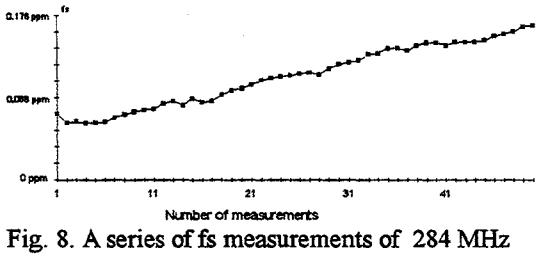


Fig. 8. A series of fs measurements of 284 MHz resonator.

4. Conclusions.

It has been confirmed that six-port technique can be used in the field of the measurements of piezoelectric resonators parameters. It offers potentially the highest absolute measurement accuracy limited only by the accuracy of the coaxial type immittance standards. Another advantages is the simplicity of the measurement circuit which makes attractive many applications in which low cost is mandatory

5. References.

1. M. Wójcicki, A. Lisowiec, *A Direct Deflection Method and Six-Port Measuring System for Accurate Determination of Piezoelectric Resonator Parameters,* Proc. 8-th European Frequency and Time Forum, Munich, 1994.
2. M. Wójcicki, A. Lisowiec, *Application of Direct Reflection Method in Quartz Resonators Metrology,* 8-th Piezoelectric Conference PIEZO'94, Zakopane, 1994.
3. A. Lisowiec, M. Wójcicki, *Six-Port Measuring System for Accurate Determination of Piezoelectric Resonators Parameters,* 8-th Piezoelectric Conference PIEZO'94, Zakopane, 1994.
4. M. Wójcicki, *A New Single Port Reflection Method for Measurement of Crystal Unit Equivalent Parameters in Wide Frequency Range,* Proc. Int. Conf. on Frequency Control and Synthesis, The University of Surrey, Guildford, 1987.
5. G. Engen, *The Six-Port Reflectometer: An Alternative Network Analyzer,* IEEE Trans. on Microwave Theory and Techn., 1977, MTT-25 12
6. G. P. Riblet, E. R. Bertil Hansson, *Aspects of the Calibration of a Single Six-Port Using a Load and Offset Reflection Standards,* IEEE Transactions on Microwave Theory and Techniques, vol. MTT-30, No 12, December 1982.
7. V. Bilik, V Raftaj, J. Bezek, *A NewExtremely Wide-Band Lumped Six-Port Reflectometer,* Proc. of 20-th Europ. Microwave Conf. Budapest, 1990
8. N. M. Ridler, *New VHF Impedance Measuring Instrument for UK National Standard,* IEE Proc. Set Meas. Technol., Vol. 141 No 1, January 1994

Computer arranged Crystal sets by Evaluation of Spurious Resonances and Motional Parameters, enabling professional Filter Production with Resonators of wider Tolerances

Markus Frey, Tele Quarz GmbH, Landstraße, 74924 Neckarbischofsheim, Tel. xx49-7268-801-157

Abstract:

Crystal filters for tight requirements in selectivity, group delay and/or time response require crystal units with extreme narrow tolerances of the static and motional parameters, also a high attenuation and sufficient frequency offset of spurious resonances, mainly for wide band applications.
If the crystals are specified in a "worst case" sense, i.e. such, that the crystals arbitrarily can be combined with each other without degradation of the filter specification, the yield in the production of such crystals is very low.
A much higher yield can be achieved, if the requirements are set much lower instead, and combined individually to matching crystal sets per filter.

This is done by a computerized method. The crystal parameters and, if required, its spurious resonances are measured using the method [1] described by B. Neubig (EFTF 1994). The results are fed into a circuit analysis module [2] (part of "TQ-WINFIL", see M.Frey, B.Neubig, EFTF 1992), with which the crystal set is selected, that fulfills the specified requirements.

The paper describes the overall concept. An exapmple application of a high selectivity broadband filter at 14,96 Mhz is given at the end of the paper, comparing the yield in the "worst case" method against this method.

Introduction:

The general filter circuit (see fig.1), which is assisted on the software, is the multiple cascaded half-lattice, containing arbitrary numbers of crystals in each branch.

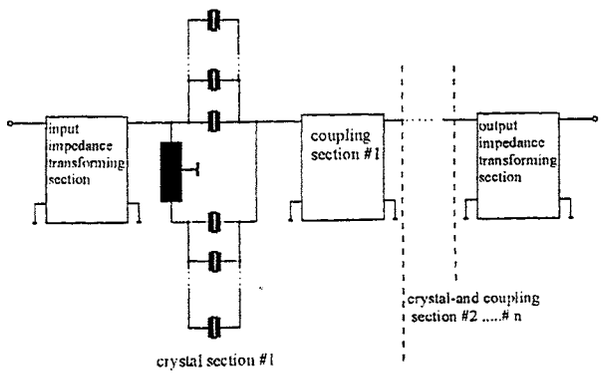


fig. 1

The electrical equivalent circuit (see fig.2) simplified describes the properties of the crystals, sufficiently. The main mode is generated by a loss imposed series resonance circuit in parallel to the static capacity. The spurious resonances are explained also by loss imposed series resonance circuits, connected in parallel to the main mode circuit.

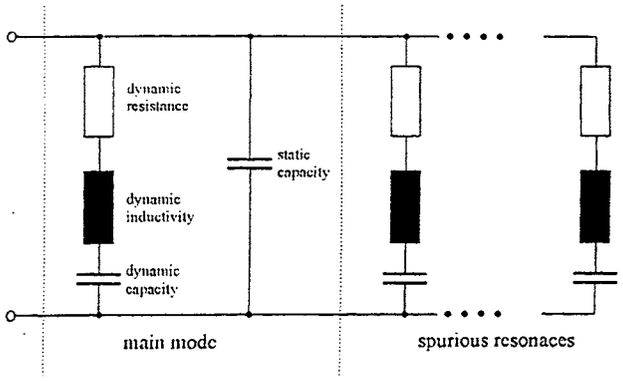


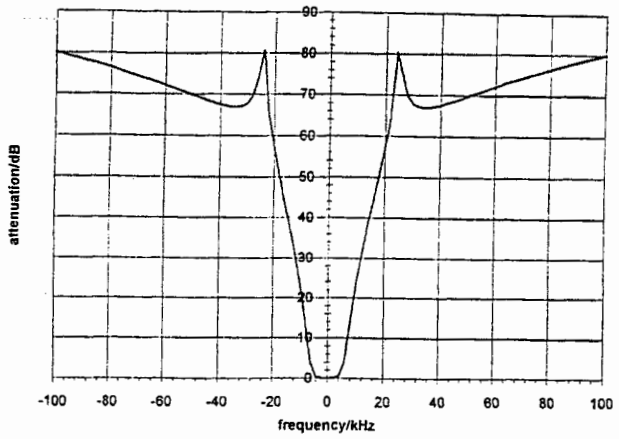
fig. 2

Generally, there are two most important properties of the crystals, that hide advantages and disadvantages with the realization of filters of half lattice structure.

- The Tolerance of the dynamic capacity C_1 of the crystals, mainly in correlation to their values of the crystals in the opposite branch.
- Cooperation between the spurious resonances of the crystals in the same branch on the one hand, and between the opposite branches on tha other hand.

Consideration of C_1 tolerance

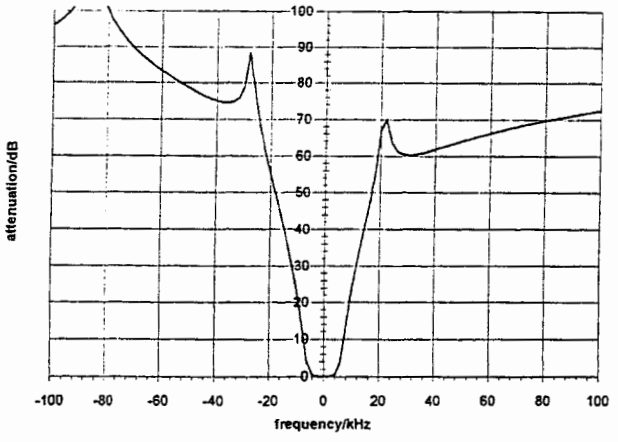
In the filter synthesis, the adequate C_1 values of the crystals are calculated refering to a C_1 value given by a choosen reference crystal. Due to this, absolute C_1 values are not reqired in order to design a certain filter characteristic, however the impedances are influenced. Consequently, the C_1 matching of the crystals to each other, entails unsymmetric shape factor by the appearance of additional infinite poles in the stopband, followed by a deficient stopband attenuation. Thus, wide C_1-tolerances are applicable (see fig.3), just in the case, that the C_1-tolerances relative to all crystals in the filter are kept in a necessary limit.



Synthesis datas:

Crystal 1:	C_1=4.0 fF
Crystal 2:	C_1=4.4 fF
Crystal 3:	C_1=4.2 fF
Crystal 4:	C_1=4.2 fF

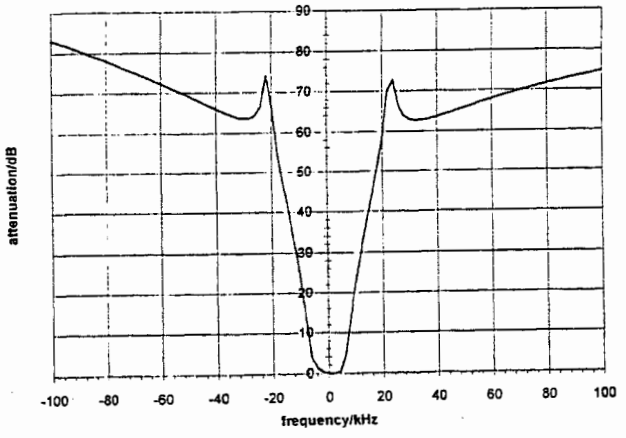
fig. 3a



Measured datas:

Crystal 1:	C_1=4.0 fF+5%=4.2 fF
Crystal 2:	C_1=4.4 fF- 5%=4.2 fF
Crystal 3:	C_1=4.2 fF- 5%=4.0 fF
Crystal 4:	C_1=4.2 fF+5%=4.4 fF

fig. 3b



Measured datas:

Crystal 1:	C_1=4.0 fF- 22%=3.1 fF
Crystal 2:	C_1=4.4 fF- 18%=3.6 fF
Crystal 3:	C_1=4.2 fF+21%=5.1 fF
Crystal 4:	C_1=4.2 fF+26%=5.3 fF

fig. 3c

Consideration of spurious resonances

Due to the phase shift of ideal 180° between the opposite branches of the half lattice circuit, spurious resonances in a very close frequency range, tend to compensate (see fig.4). However the quantitative evidence depends on the width of the spurious resonances given by their motional parameters.

In contrary to this, spurious resonances of crystals in the same branch decrease their resulting attenuation (see fig.4), if they are operating at the same or in a very close frequency range, in which this effect is the consequence of their impedances, that operate in parallel.

A "worst case" specification of the spurious resonances must consider the addition of the spurious admittances, however this tight specification is not necessary, if individual combinations of crystals are applied to filters.

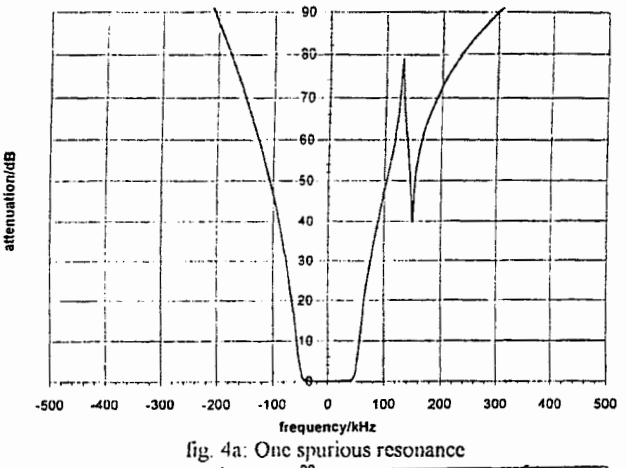


fig. 4a: One spurious resonance

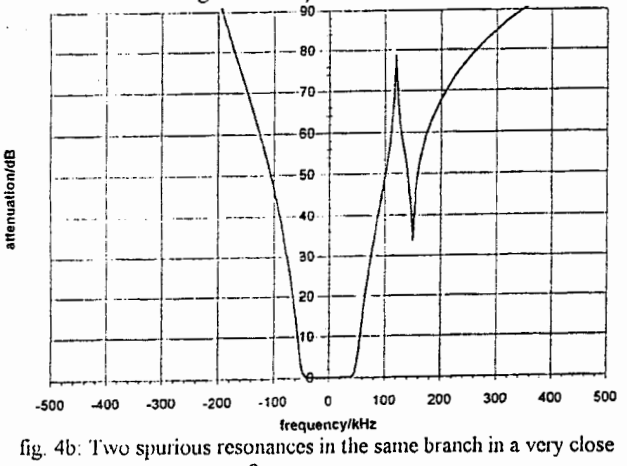


fig. 4b: Two spurious resonances in the same branch in a very close frequency range

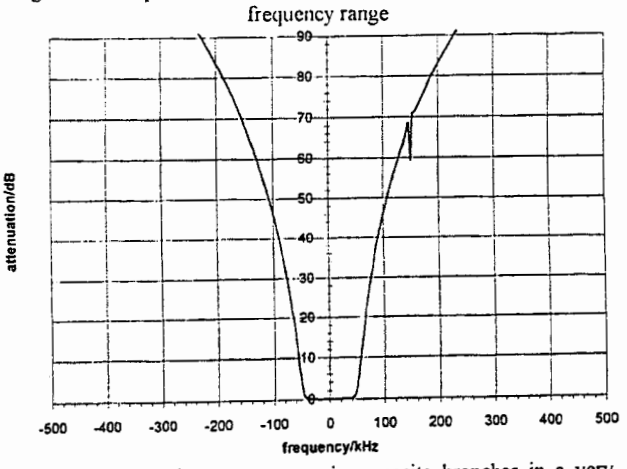


fig. 4c: Two spurious resonances in opposite branches in a very close frequency range

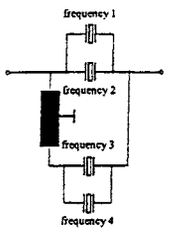
General aspects of combinating of crystals

In order to find out the maximum number of functional crystal sets of a certain quantity of crystals, first of all, the possible crystal combinations, due to the respective filter circuit equivalent, have to be ascertained. Consequently, if there are no contitions in order to reduce the number of possibilities systematically, all possible combinations have to be verified according to the filter specification.
In fig.5, two simplified samples are given in order to show the multiple possibilities of crystal variations, belonging to the respective filter circuit and a total amount of disposable crystals.

disposable crystals

row no	frequency 1	frequency 2	frequency 3	frequency 4
1	no. 11	no. 21	no. 31	no. 41
2	no. 12	no. 22	no. 32	no. 42

filter curcuit



all possible crystal arrangements due to the conditions above:

	combination #1	#2	#3	#4	#5	#6	#7	#8
branch #1	no. 11	no. 11	no. 11	no. 11	no. 11	no. 11	no. 11	no. 11
branch #2	no. 21	no. 22	no. 21	no. 22	no. 21	no. 22	no. 21	no. 22
branch #3	no. 31	no. 31	no. 32	no. 32	no. 31	no. 31	no. 32	no. 32
branch #4	no. 41	no. 41	no. 41	no. 41	no. 42	no. 42	no. 42	no. 42

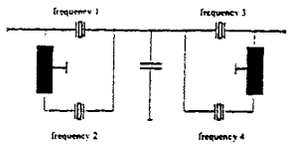
	#9	#10	#11	#12	#13	#14	#15	#16
branch #1	no. 12	no. 12	no. 12	no. 12	no. 12	no. 12	no. 12	no. 12
branch #2	no. 21	no. 22	no. 21	no. 22	no. 21	no. 22	no. 21	no. 22
branch #3	no. 31	no. 31	no. 32	no. 32	no. 31	no. 31	no. 32	no. 32
branch #4	no. 41	no. 41	no. 41	no. 41	no. 42	no. 42	no. 42	no. 42

fig. 5a

disposable crystals

row no	frequency 1	frequency 2
1	no. 11	no. 21
2	no. 12	no. 22
3	no. 13	no. 23

filter curcuit



all possible crystal arrangements due to the conditions above:

	combination #1	#2	#3	#4	#5	#6	#7	#8	#9
branch #1	no. 11	no. 11	no. 11	no. 12	no. 12	no. 12	no. 11	no. 11	no. 11
branch #2	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22
branch #3	no. 12	no. 12	no. 12	no. 11	no. 11	no. 11	no. 13	no. 13	no. 13
branch #4	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23

	#10	#11	#12	#13	#14	#15	#16	#17	#18
branch #1	no. 13	no. 13	no. 13	no. 12	no. 12	no. 12	no. 13	no. 13	no. 13
branch #2	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22
branch #3	no. 11	no. 11	no. 11	no. 13	no. 13	no. 13	no. 12	no. 12	no. 12
branch #4	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23

	#19	#20	#21	#22	#23	#24	#25	#26	#27
branch #1	no. 11	no. 11	no. 11	no. 12	no. 12	no. 12	no. 11	no. 11	no. 11
branch #2	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23
branch #3	no. 12	no. 12	no. 12	no. 11	no. 11	no. 11	no. 13	no. 13	no. 13
branch #4	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22

	#28	#29	#30	#31	#32	#33	#34	#35	#36
branch #1	no. 13	no. 13	no. 13	no. 12	no. 12	no. 12	no. 13	no. 13	no. 13
branch #2	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23	no. 22	no. 23	no. 23
branch #3	no. 11	no. 11	no. 11	no. 13	no. 13	no. 13	no. 12	no. 12	no. 12
branch #4	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22	no. 21	no. 21	no. 22

fig. 5b

In the samples shown in fig.5, the crystals are attached to the table by assorting each different crystal specification, i.e. crystal frequency, to a different column, whereat the row number is equal to the consecutive number of crystals with the same specification.
The calculation of the maximum number of crystal variations is done by multiplying the maximum number of variations inside each column with their belonging numbers of the other columns of the table.
Due to the sample in fig.5a, the result is

$$(2\times2\times2\times2=16) \text{ variations}$$

and due to the sample in fig.5b:

$$(6\times6=36) \text{ variations}$$

According [3], the maximum number of variations $V_n^{(k)}$, possible in one column of an arbitrary quantity of crystals, is given by

$$V_n^{(k)} = n\cdot(n-1)\cdot....(n-k+2)\cdot(n-k+1)$$

whereat k is the number of elements (crystals) forming one variation and n is the total number of elements (crystals) in the column.
In consideration of tables which contain more than one column, the maximum number of variations is given by multiplying the particular numbers of variations according to each column. Consequently, the maximum number of variatons according fig.5b is

$$V_3^{(2)}\cdot V_3^{(2)} = (3\cdot(3-1))\cdot(3\cdot(3-1)) = 6\cdot6 = 36$$

Applying this calculation to the practical example of the 14,96 Mhz broadband filter, which is given at the end of the paper, the maximum number of variations is gained in the range of 10^{39}. Thus a verification, i.e. analysis of all crystal variations is not applicable due to the time, even if it is done by a computer (CPU i80486/33 : time of calculation >10000 years).
Hence, simplifications have to be considered at the combining algorithm.

Computerized method of combinating crystals individually:

According to the tolerance of the C_1 values of the crystals and their properties in the filter circuit, which already are discussed above, the crystals that are fed into the combining algorithm, are assorted according to the magnitude of their C_1 values. Due to this, the probability for matching crystals is high at the beginning of the combining procedure, whereas it decreases strongly at further combinations. Thus, an interrupt condition and new starting conditions can be defined. Concerning the spurious resonances, yet no systematic combining method was found, that, in consideration of the C_1 assortment, reduces the calculation time. Hereafter, the principle of the combining algorithm is explained briefly, along with the example of the filter circuit of fig.6.

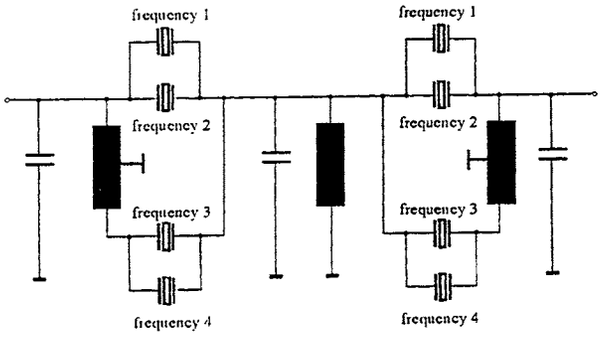


fig.6

A starting filter is set with the crystals no.1,3,7,67,4,15,8 and 44 of the table in fig.7. Thereafter, the filter is verified acc. the filter specification, using the complex network analysis of [2], and assessed sufficient or insufficient.

row no.	frequency #1	frequency #2	frequency #3	frequency #4	
1	no. 1	no. 3	no. 7	no. 67	crystals of starting filter
2	no. 4	no. 15	no. 8	no. 44	crystals of starting filter
3	no. 91	no. 27	no. 74	no. 74	crystals of exchanging filter
4	no. 3	no. 54	no. 21	no. 2	crystals of exchanging filter
5	no. 8	no. 9	no. 9	no. 7	
6	...	...	...	...	

fig. 7

If the filter with the topical crystal variation is assessed sufficient, i.e. it keeps the specification, the crystals and their position in the filter are saved, and the crystals are taken out of the table. Crystals in the row no.3 are replaced to row no.1, and crystals in row no.4 are replaced to row no.2, etc.
Else if the filter is assessed insufficient, i.e. it doesn't keep the specification, an exchange filter is set with the crystals of row no.3 and row no.4 of the table in fig.8.

row no.	frequency #1	frequency #2	frequency #3	frequency #4	
1	no. 1	no. 3	no. 7	no. 67	crystals of starting filter
2	no. 4	no. 15	no. 8	no. 44	crystals of starting filter
3	no. 91	no. 27	no. 74	no. 74	crystals of exchanging filter
4	no. 3	no. 54	no. 21	no. 2	crystals of exchanging filter
5	no. 8	no. 9	no. 9	no. 7	
6	...	...	...	...	

fig. 8

The variation process between the crystals of the starting filter and the exchanging filter is performed due to the theoretical considerations above, whereat each variation is verified using the circuit analysis program described in [2].

Three possibilities can be obtained:

1. The topical variation is assessed insufficient, but improved:
   The crystal variation in the starting filter is kept, and the combining of the crystals between the starting filter and the exchanging filter goes on as long as the possibility 2 or 3 applies.

2. The topical variation is assessed sufficient:
   The crystals of the starting filter are saved and taken out of the table. The gaps of the table are replaced by the crystals of the following rows and a new starting filter is set. The exchange procedure starts new.

3. After all possible variations between the two filters, no filter, which is assessed sufficient, was found:
   The crystal variation in the starting filter is kept, and the exchanging filter is set newly, with the crystals of the rows no.5 and no.6, see fig.9. The exchanging procedure between the starting filter and exchanging filter continues, however interrupt conditions are implemented in this case.

row no.	frequency #1	frequency #2	frequency #3	frequency #4	
1	no. 1	no. 3	no. 7	no. 74	crystals of starting filter
2	no. 4	no. 54	no. 21	no. 44	crystals of starting filter
3	no. 91	no. 27	no. 74	no. 2	
4	no. 3	no. 15	no. 8	no. 67	
5	no. 8	no. ...	no. ...	no. ...	crystals of exchanging filter
6	...	...	...	...	crystals of exchanging filter

fig. 9

If the third possibility also is completed, e.g. by an interrupt, or with the last two rows of the table, the table will be reorganized acoording fig.10. The complete combining procedure starts again. This is repeated until every row of the table once was placed at no.1, a logical interrupt appeared, or the process is stopped manually.

row no.	frequency #1	frequency #2	frequency #3	frequency #4	
1	no. 4	no. 15	no. 8	no. 44	crystals of starting filter
2	no. 91	no. 27	no. 74	no. 74	crystals of starting filter
3	no. 3	no. 54	no. 21	no. 2	crystals of exchanging filter
4	no. 8	no. 9	no. 9	no. 7	crystals of exchanging filter
5					
6					
...					
...					
	no. 1	no. 3	no. 7	no. 67	end of the table

fig. 10

This so called "rotation" of the table is applied due to the experience, that new starting conditions provide different results.
According to fig.11, which shows the functional description of the program, the crystal combinations, that provide the maximum number of the found and sufficiently assessed filters, are available at the a output of the software.

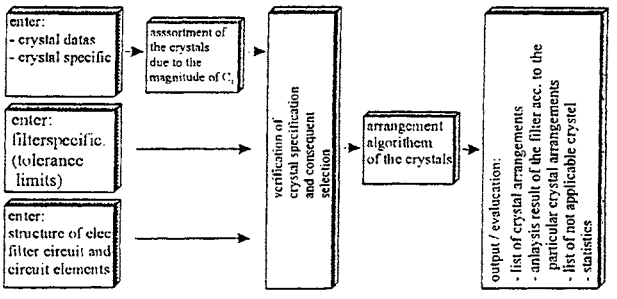


fig.11

Example 1:

Application of a high selectivitiy broadband filter see fig.12, at fc=14,96 Mhz.
Passband width: fc ±75 kHz min. at 1dB
Stopbandwidth: fc ±110 kHz max. at 70 dB
Spurious response attenuation:
general 70 dB minimum; attenuation of one defined spurious response has to be at least 65 dB.

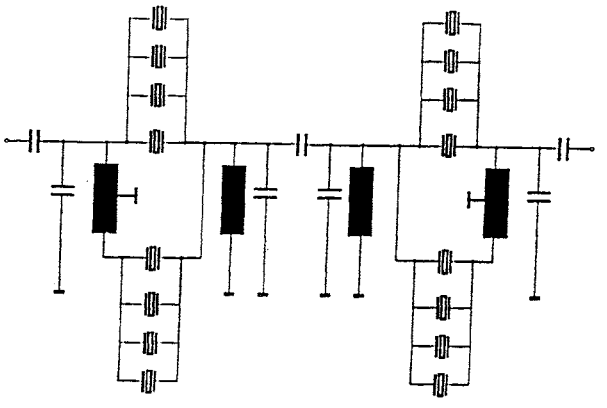
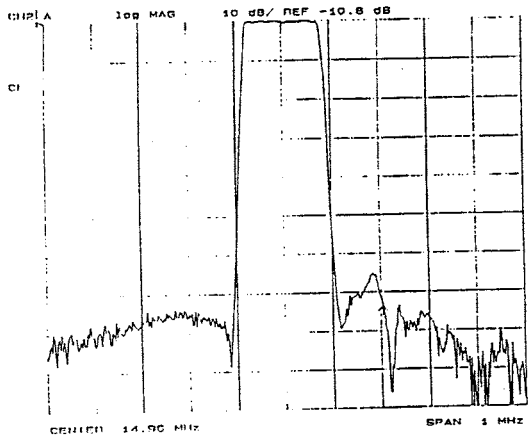
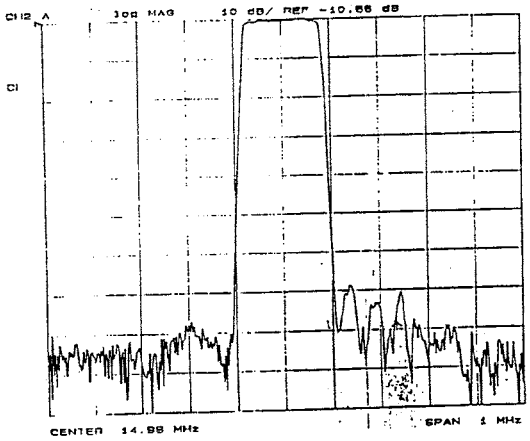
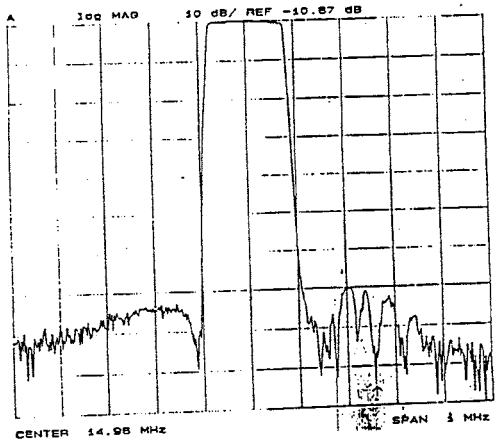


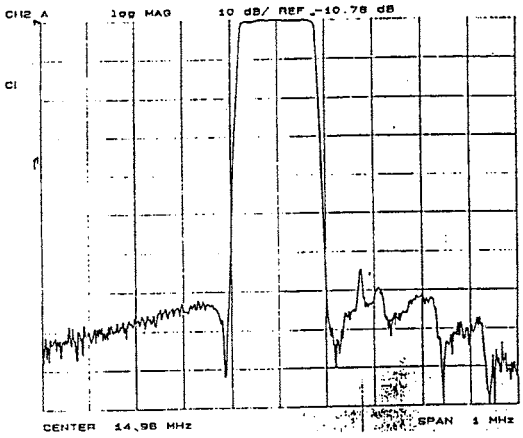
fig. 12

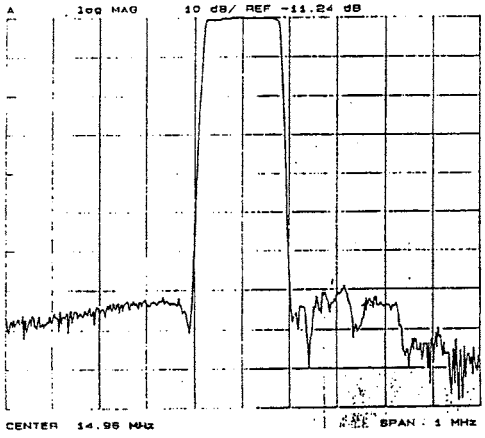
The disposable quantity of crystals in this example is 33 crystals of each frequency (this corresponds to 33 rows of the table), by eight different frequencies (columns of the table). Thus, 16 filters can be obtained maximum. Fig.13 shows the quantity and the results of obtained filters according to the specification, even though a manual selection and combination of the crystals was applied. A production of the crystals according to the necessary "worst case" specification is not known as possible, due to the spurious impedances in the range of mega ohms an C_1 tolerances in the range ±3% maximum. Hence, a filter production even without any crystal selection would typically not result in any filter solution fulfilling the specification in this case.


CH2 A
log MAG
10 dB/ REF -10.8 dB
CI
CENTER 14.96 MHz
SPAN 1 MHz


CH2 A
log MAG
10 dB/ REF -10.66 dB
CI
CENTER 14.96 MHz
SPAN 1 MHz


A
log MAG
10 dB/ REF -10.67 dB
CENTER 14.96 MHz
SPAN 1 MHz


CH2 A
log MAG
10 dB/ REF -10.78 dB
CI
CENTER 14.96 MHz
SPAN 1 MHz


A
log MAG
10 dB/ REF -11.24 dB
CENTER 14.96 MHz
SPAN 1 MHz

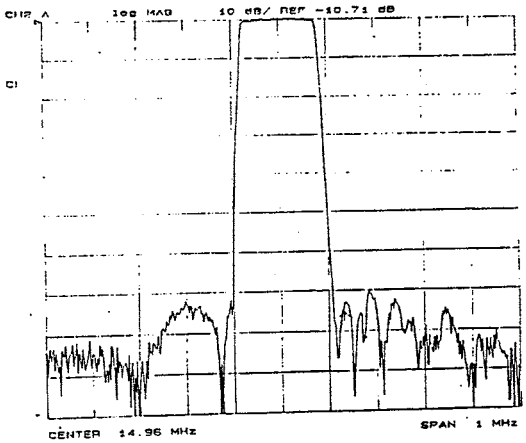
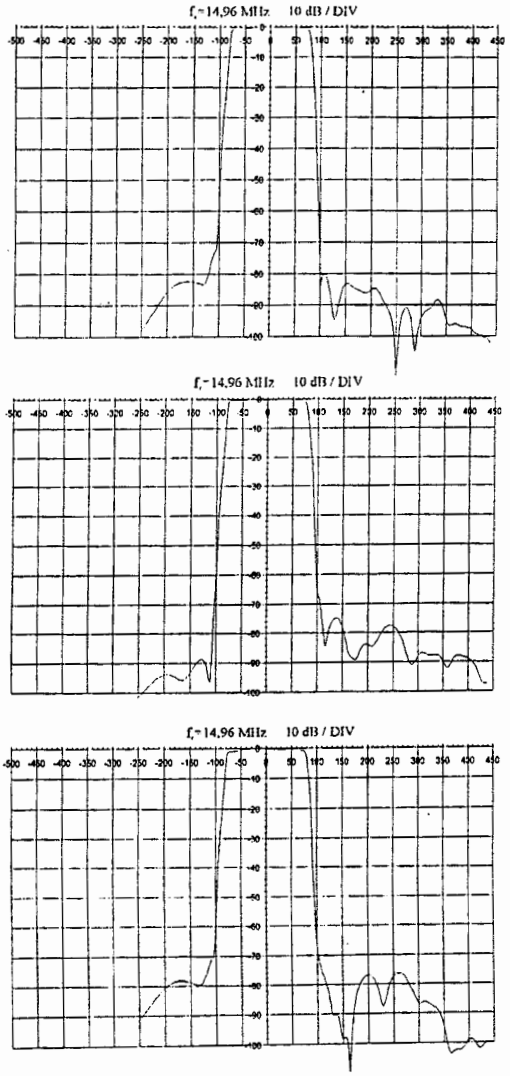
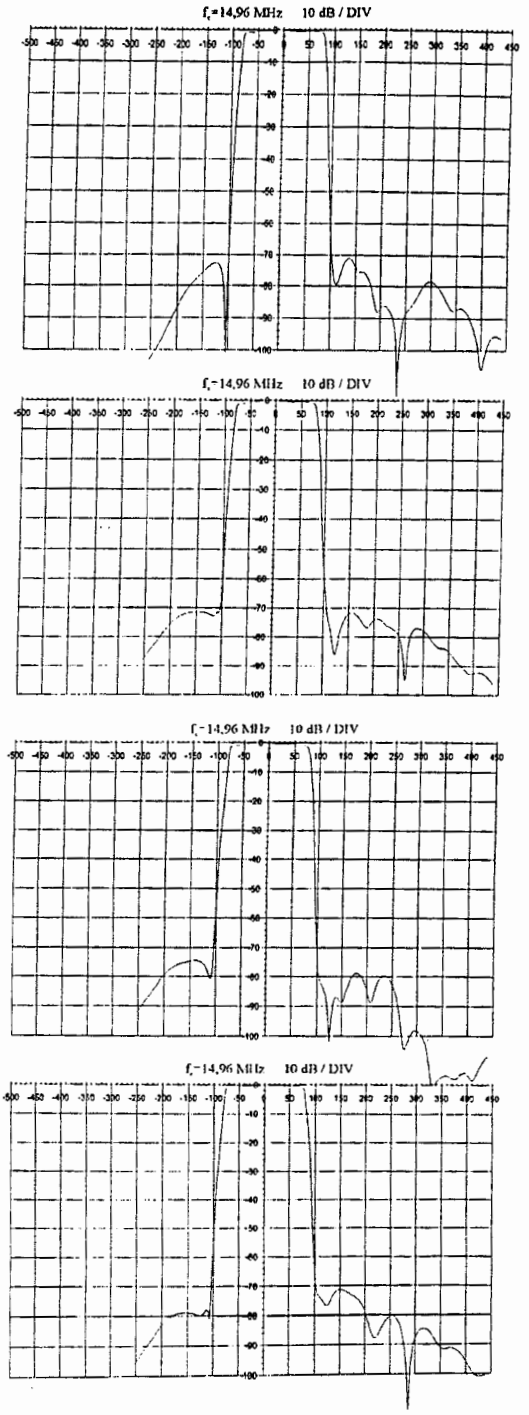
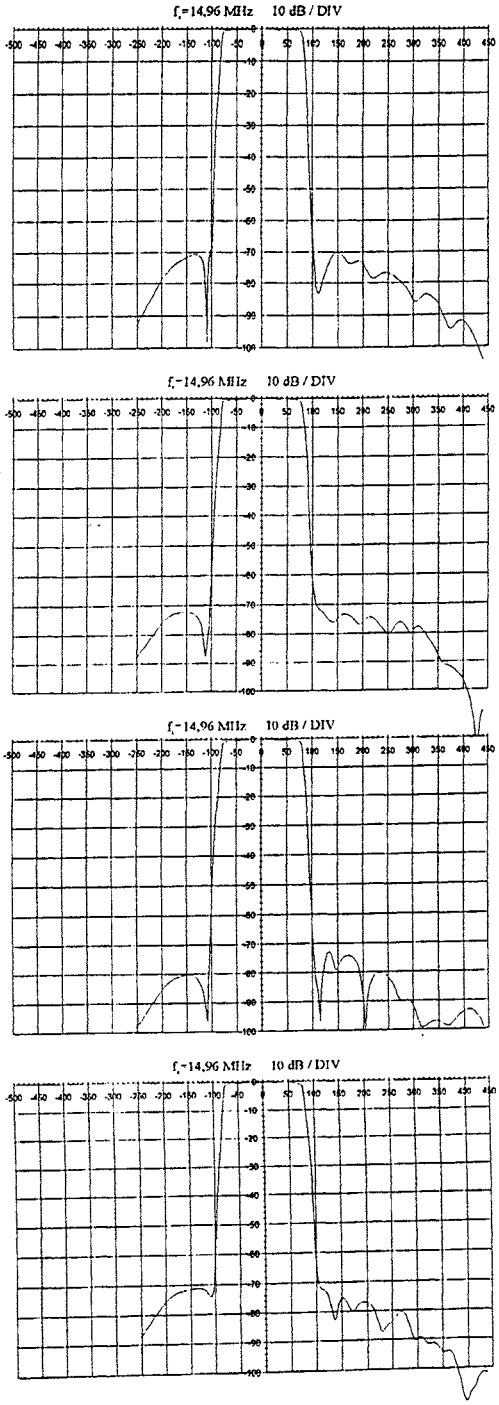

CH2 A
log MAG
10 dB/ REF -10.71 dB
CI
CENTER 14.96 MHz
SPAN 1 MHz

fig. 13

After applying the computerized method of individually crystal combination, 13 of 16 possible crystal combinations were found, and the results are shown in fig.14.







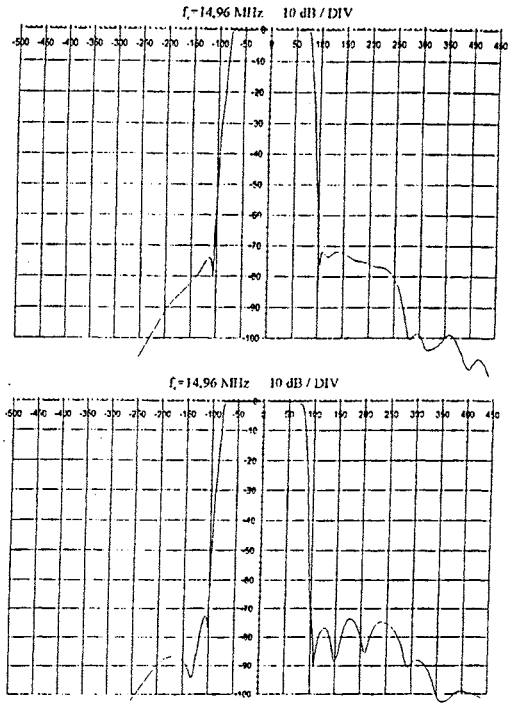


fig. 14

Summary:

The method of combining crystals to individual sets, usually is well known by the crystal filter manufacturers. However, far-going it is just applied on little quantities of filters and on filter types at whose "worst case" crystal specification is far away from the possibilities of realization. This is due to the teditious effort of time, that has to be spent by doing this work manually. The advantages of the computerized method are described by explaining the interrelations and the demonstration of examples. The results encourage to facilitate and widen up the crystal specification in order to increase the yield in the crystal production, whereat typically the yield of the realizable filters, that are produced of the less strong specified crystals, is not impaired negatively.

References

[1] Recent advances in the measurement of spurious resonances of crystal units. B.W. Neubig EFTF 1994

[2] A state-of-the-art crystal filter design package using Windows 3.0 techniques. M.Frey, B.W. Neubig EFTF 1992

[3] Taschenbuch mathematischer Formeln, Bartsch

Recent Advances In The Measurement of Spurious Resonances Of Crystal Units

Bernd W. Neubig

TELE QUARZ GROUP, D-74924 Neckarbischofsheim / Germany

1. Abstract

In the paper presented at the EFTF 1994 [3] we presented a new approach for the measurement of spurious responses with a network analyzer in conjunction with error correction techniques. It applies the standard test set-up (such as the π-network or a s-parameter test fixture) as it is used for the determination of the crystal parameters of the main mode [1]. In the meantime the proposed measurement algorithm has been improved and changed into a "mode subtraction" method, through which the measurement speed can be accelerated significantly. This paper describes the new algorithm together with measurement results achieved.

2. Measurement method

The measurement technique is based on a network analyser, the π-network according to IEC 444 and the error correction technique described in [5] and standardized in IEC 444-5 [2]. The same method can also be used in conjunction with a s-parameter test system and an adequate crystal test fixture.

In a π-network system an unknown admittance $\underline{Y}_x$ can be determined from the complex output voltage $\underline{U}_{Bx}$ and from the output voltages at calibration with reference short ($\underline{U}_{BK}$) and with reference resistor R_N ($\underline{U}_{BN}$) [complex variables are underlined]:

$$\underline{Y}_x = \frac{\left(\frac{\underline{U}_{BK}}{\underline{U}_{BN}} - 1\right)}{R_N \left(\frac{\underline{U}_{BK}}{\underline{U}_{Bx}} - 1\right)} \qquad (1)$$

The new measurement sequence is as follows:

Step 1: Calibration

Prior to the measurement of the crystal, the calibration is performed in three steps:

with reference short: $\underline{U}_{BK} = |U_{BK}|\,\varphi_{BK}$ (2)

with reference resistor: $\underline{U}_{BN} = |U_{BN}|\,\varphi_{BN}$ (3)

with "open": $\underline{U}_{Bo} = |U_{Bo}|\,\varphi_{Bo}$ (4)

The test fixture's cross-talk capacitance C_c is computed from the three calibrations:

$$C_c = \frac{\left(\frac{\underline{U}_{BK}}{\underline{U}_{BN}} - 1\right)}{j\,\omega_C R_N \left(\frac{\underline{U}_{BK}}{\underline{U}_{Bo}} - 1\right)} \qquad (5)$$

The frequency range for calibration is $\omega_C = 2\pi\,(f_{nom} \pm 20\%)$. At least 101 data points should be used. The sweep time can be as fast as the settling time of the instrument allows.

Step 2: Measurement of static capacitance C_0

The crystal units is inserted into the π-network and the complex output voltage $\underline{U}_{BC}$ is measured at 11 frequencies off resonance as defined in IEC 444-5. From this the static capacitance C_0 is computed:

$$C_0 = \frac{\left(\frac{\underline{U}_{BK}}{\underline{U}_{BN}} - 1\right)}{j\,\omega_{C0} R_N \left(\frac{\underline{U}_{BK}}{\underline{U}_{BC}} - 1\right)} - C_c \qquad (6)$$

The frequency range for $\omega_{C0} = 2\pi\,(f_{nom} \pm 20\%)$.

Step 3: Measurement of the main resonance

The nominal frequency f_{nom} of the main mode is assumed to be known. The expected resonance frequency can be estimated from f_{nom}. The parameters of the main resonance are determined by Neuscheler's method (described in IEC 444-5) from admittance values at two frequencies ω_1, ω_2 in the vicinity of the resonance frequency, preferably one below, one above it.

With

$$Y(\omega_1) = a_1 + j\,b_1 \qquad (7)$$

$$Y(\omega_2) = a_2 + j\,b_2 \qquad (8)$$

and

$$a_i^* + j\,b_i^* = \frac{1}{a_1 + j\,b_1} \qquad (9)$$

a first guess for the motional parameters can be calculated as

$$L_1^{(1)} = \frac{\omega_1 b_1^* - \omega_2 b_2^*}{\omega_1^2 - \omega_2^2} \qquad (10)$$

$$C_1^{(1)} = \frac{\omega_1^2 - \omega_2^2}{\omega_2^2\,\omega_1 b_1^* - \omega_1^2\,\omega_2 b_2^*} \qquad (11)$$

$$R_1^{(1)} = \frac{a_1^* + a_2^*}{2} \qquad (12)$$

The accuracy depends on how close the frequencies ω_1, ω_2 are located relative to the 3db (±45°) points of the admittance of the motional branch L_1, C_1, R_1. Even if these points are not well located, the series resonance frequency can already be computed to a fairly high accuracy (a few ppms) from the motional parameters, i.e.

$$f_{s1}^{(1)} = \frac{1}{2\pi\sqrt{L_1^{(1)} C_1^{(1)}}} \qquad (13)$$

From the Q-value derived herewith

$$Q_1^{(1)} = \frac{\omega_{s1} L_1^{(1)}}{R_1^{(1)}} \qquad (14)$$

the two corner frequencies $\omega_1^{(2)}$, $\omega_2^{(2)}$ - which are very close to the optimum 3db (±45°) points - are calculated

$$\omega_{1,2}^{(2)} = 2\pi f_{s1}^{(1)} \left(1 \pm \frac{1}{Q_1^{(1)}}\right) \qquad (15)$$

At these frequencies the measurement is repeated and an improved set of motional parameters $L_1^{(2)}$, $C_1^{(2)}$, $R_1^{(2)}$ and hence $f_{s1}^{(2)}$ and $Q_1^{(2)}$ for the main resonance is achieved with eq. (10) to (14). This can be considered to be sufficiently accurate. To avoid ringing effects, the settling time of at each frequency should be

$$t_{settle} \geq 5\,\frac{Q_1}{\omega_i} \qquad (16)$$

Step 4: Fast sweep through frequency range of interest

A fast sweep in the frequency range under consideration with a sweep time of approximately 1... 10 sec and with minimum 501 data points of amplitude $|U_B(f)|$ and phase $\varphi_B(f)$ is performed.

Using eq. (1) the output voltage values $\underline{U}_B(f)=|U_B(f)|\ \varphi_B(f)$ can be transformed to a complex admittance matrix $\underline{Y}_x(f) = |\underline{Y}_x(f)|\ \varphi_{Y_x}(f)$ with the influence of the static capacitance substracted:

$$\underline{Y}_x(f) = \frac{\left(\frac{\underline{U}_{BK}}{\underline{U}_{BN}} - 1\right)}{R_N * \left(\frac{\underline{U}_{BK}}{\underline{U}_B(f)} - 1\right)} - j\,\omega\,C_0 \qquad (17)$$

Step 5: Subtraction of the main mode admittance from $\underline{Y}_x(f)$

The admittance of the main mode $\underline{Y}_m(f)$ expressed in the element values and the series resonance frequency determined in step 3 is :

$$\underline{Y}_m(f) = \frac{1}{R_1 + j\frac{1}{\omega C_1}\left(\frac{\omega^2}{\omega_{s1}{}^2} - 1\right)} \qquad (18)$$

This admittance is now subtracted from the admittance matrix $\underline{Y}_x(f)$ of step 4. The residual admittance $\underline{Y}_r$ (f) after removal of the main mode is therefore:

$$\underline{Y}_r(f) = \underline{Y}_x(f) - \underline{Y}_m(f) \qquad (19)$$

Step 6: Search for spurious resonance peaks

The residual admittance matrix $\underline{Y}_r$ (f) from step 5 is now mathematically analyzed to select the resonance peaks of the spurs. This is done in the following sequence:

(a) Look for local maxima:
Compare $|Y_r(f_{i-1})|$ with $|Y_r(f_i)|$ and $|Y_r(f_{i+1})|$ for i = 2 ... 500.
If

$$|Y_r(f_i)| > |Y_r(f_{i-1})| \text{ and } |Y_r(f_i)| > |Y_r(f_{i+1})| \qquad (20)$$

then f_i is considered to be a candidate for a resonance peak $f_{peak\,i}$

(b) Distinquish between real peaks and fake peaks:
Fake peaks due to noise etc. can be identified by assuming a realistic maximum Q-value of the spurious resonances with respect to the main mode Q-value determined in step 3:

$$Q_{max} = \theta_{max} * Q_1{}^{(2)} \qquad (21)$$

A reasonable value of the multiplier is θ_{max}= 5 (recommended: 2 ... 10).
The minimum 3dB half bandwidth for a spurious peak is therefore:

$$\Delta f_{min} = f_{nom} / 2\,Q_{max} \qquad (22)$$

For each "candidate" for a resonance peak from (a) the data points next to $f_{peak} \pm \Delta f_{min}$ are inspected, if the amplitude at the sides is decreased not more than according to Q_{max}:

$$\frac{|Y_r(f_{peak})|}{|Y_r(f_{peak} \pm \Delta f_{min})|} \le \sqrt{2} \qquad (23)$$

Only if this condition is fulfilled, the peak is accepted as a further candidate.

(c) Given a minimum Q-value Q_{min} with respect to the main mode Q of

$$Q_{min} = \theta_{min} * Q_1{}^{(2)} \qquad (24)$$

A reasonable value of the multiplier is θ_{min}= 0,2 (recommended: 0,1 ... 0,5). The maximum 3dB half bandwidth for a spurious peak is therefore:

$$\Delta f_{max} = f_{nom} / 2\,Q_{min} \qquad (25)$$

For each remaining "candidate" for resonance peak after applying (b) the data points next to $f_{peak} \pm \Delta f_{max}$ are inspected, if the amplitude at the sides is decreased accordingly, i.e.:

$$\frac{|Y_r(f_{peak})|}{|Y_r(f_{peak} \pm \Delta f_{max})|} \ge \sqrt{2} \qquad (26)$$

If this condition is also fulfilled, the peak is accepted as a spurious resonance peak.

Note:
If spurious resonances very close to strong modes must be considered, it is recommended to use the 1 dB instead of the 3dB bandwidth in a.m. formulas. The term $\sqrt{2}$ is then to be replaced by the factor 1,122 and the Δf_{max} value has to be changed accordingly.

Step 7: Zooming of each identified spurious resonance peak

At each resonance peak determined in step 6 a new set of measurement data is taken by zooming a frequency range of $f_{peak} \pm \Delta f_{max}$ with at least 11 or 21 data points per sweep, and a sweep time of

$$t_{sweep} \ge \frac{10}{\theta_{min}} \frac{\sqrt{Q_{min}\,Q_{max}}}{\omega_0} \qquad (27)$$

should be used. The amplitude and phase data for each peak are transformed to admittance data $\underline{Y}_x(f) = |\underline{Y}_x(f)|\ \varphi_{Yx}(f)|$ as in step 4, eq. (17) , where the calibration data for $\underline{U}_{BK}$ and $\underline{U}_{BN}$ are derived from those of step 1 by interpolation.

Note:
This can only be done if the network analyser maintains its calibration in the narrower sweep ranges of the spurs.

Then the influence of the main mode is removed from the admittance matrices of all N measured spurs (j = 1 ... N) as described in step 5, eq. 15. This yields N intervals of residual admittance matrices $\underline{Y}_r(f^{(j)})$

$$\underline{Y}_r{}^{(j)}(f) = \underline{Y}_x{}^{(j)}(f) - \underline{Y}_m{}^{(j)}(f) \qquad (28)$$

In each of the N intervals the element with the maximum value of $|Y_r{}^{(j)}|$ is determined, yielding N values $|Y_{rmax}{}^{(j)}|$. In the next steps the intervals are analysed in a sequence according to the magnitude of $|Y_{rmax}{}^{(j)}|$.

Step 8: Computation of the spurious modes

(a) The strongest resonance peak, i.e. with the the highest $|Y_{rmax}{}^{(j)}|$ is selected first. Assuming, this is the k-th mode.
The parameters of this resonance are now determined by the same method as described for the main mode (step 3), i.e. from admittance values at two frequencies $\omega_1{}^{(k)}$, $\omega_2{}^{(k)}$ in the range of the spur's 3dB bandwidth, for which a first guess is derived from the Q-value of the main mode

$$\omega_{1,2}{}^{(k)} = 2\pi\, f_{spur}{}^{(k)} \left(1 \pm \frac{1}{Q_1{}^{(2)}}\right) \qquad (29)$$

In practice, the measured points next to these calculated frequencies are uses. Alternatively instead of $Q_1{}^{(2)}$ the root mean value $\sqrt{Q_{min}\,Q_{max}}$ can be used.

With the equations given in step 3 the parameters $f_s{}^{(k)}$, $R^{(k)}$, $C^{(k)}$, $L^{(k)}$, $Q^{(k)}$ of k-th spurious mode are computed.

(b) Subtraction of the spurious mode's admittance from $\underline{Y}_r(f)$

The admittance of the k-th spurious mode determined in (a)

$$\underline{Y}_k(f) = \frac{1}{R_2 + j\frac{1}{\omega C_2}\left(\frac{\omega^2}{\omega_{s2}{}^2} - 1\right)} \qquad (30)$$

is now substracted from the residual admittance matrix intervals of all remaining zoomed resonance peaks.

$$\underline{Y}_r^{(j)}(f) <= \underline{Y}_r^{(j)}(f) - \underline{Y}_k(f) \quad \text{for } j = 1 \ldots N, \quad j \neq k \qquad (31)$$

(c) Repetition for all remaining resonances

The residual admittance gained in step 8 (b) is the starting point for the computation of the next strongest spurious mode as in step 8 (a), which is then removed accordingly etc.
For weaker spurs and for spurs very close to other modes it is recommend to use frequency points closer to the peak than in eq. (29), e.g.

$$\omega_{1,2}^{(j)} = 2\pi f_{spur}^{(j)} \left(1 \pm \frac{1}{2\,Q_1^{(2)}}\right) \qquad (32)$$

The whole procedure of step 8 continues until all selected peaks are determined.

Final step: Evaluation of the validity of the analysis

Finally the admittances of all spurious modes including the main mode plus $\underline{Y}_{C0}$ are aummed together

$$\underline{Y}_{total} = \Sigma \underline{Y}_j + Y_m + \underline{Y}_{C0} \qquad j = 1 \ldots N \qquad (33)$$

and compared with the measured response

$$\underline{Y}_{error} = \underline{Y}_{meas} - \underline{Y}_{total} \qquad (34)$$

From the normalized sum of squares

$$Y^2_{error} = \frac{1}{\text{no. of points}} \Sigma(|Y_{i\,meas}| - |Y_{i\,super}|)^2 \qquad (35)$$

a measure for the quality of the data fitting can be derived.

3. Experimental results

A fundamental mode VCXO crystal with a load resonance frequency of 38,880 MHz was analyzed by the above described method. Fig. 1 to fig. 6 show the amplitude and phase of the crystal admittance $\underline{Y}_x(f)$ derived directly from the measured output voltages (solid line). The dotted lines show the value of the total admittance Ytotal summed from the resonance parameters computed from the meausurement according to eq. (33).

In fig. 1 the overall sweep is depicted, fig. 2 shows the main resonance, figs. 3 to 6 show the zoomed response of 4 spurious resonances (as denoted in fig.1).

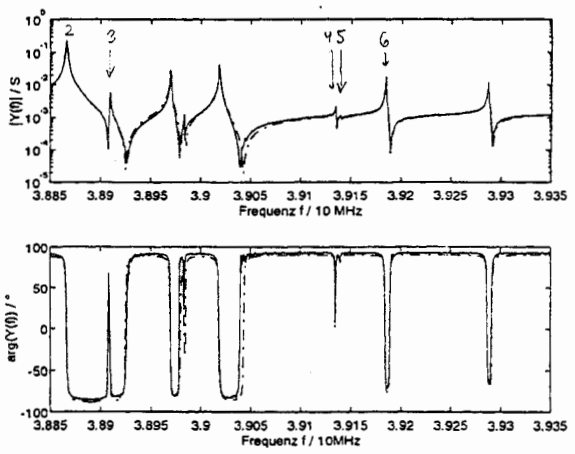


Fig. 1: complete resonance spectrum of crystal 38.880 MHz

The electrical parameters are:

resonance #	f_S[kHz]	R_1 [Ω]	C_1 [fF]	Q	C_0 [pF]
2	38.866,034	3,81	25,43	42.300	6,59
3	38.909,751	138	0,427	69.400	
4	39.134,844	431	0,127	74.300	
5	39.139,636	2810	0,0216	67.100	
6	39.184,517	53,1	1,15	66.500	

The figures show clearly, that the results match very close to the measured response.

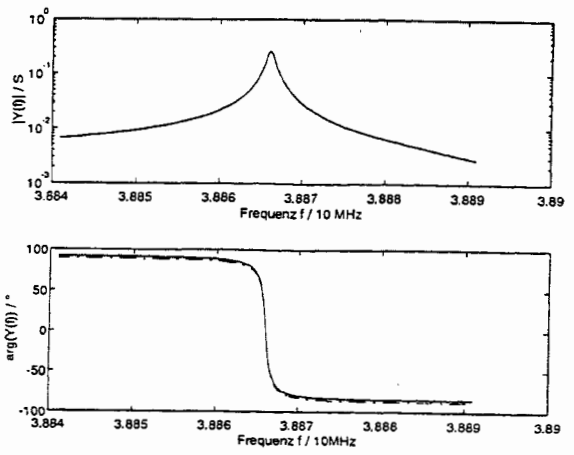


Fig. 2.: main resonance (# 2) 38.866,034 kHz

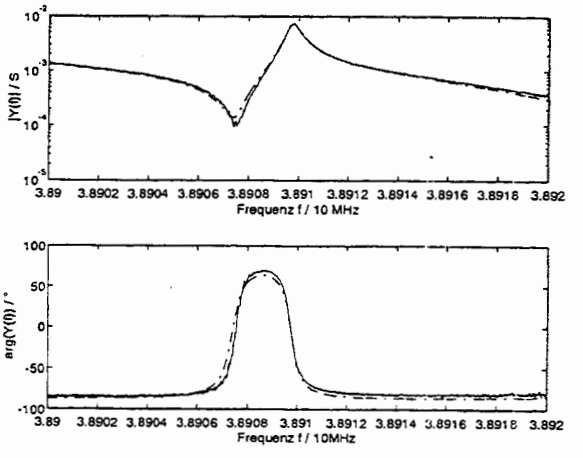


Fig. 3: spurious resonance #3 38.909,751 kHz

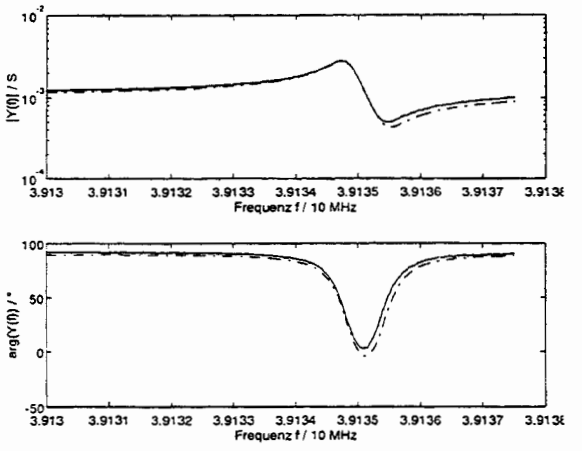


Fig. 4: spurious resonance #4 39.134,844 kHz

Special attention shall be drawn to figures 3 and 5. Fig 3 shows a resonance close to the main mode, where the admittance response is strongly effected by the latter, such that virtually the anti-resonance is below the series resonance. Fig. 5 is an extremely weak spurious resonance, whose response is nearly completely covered and "smeared" by the admittance of the static capacitance C_0 such that the total phase excursion is less then 20° rather than the "ideal" non disturbed 180°. The series resonance resistance of this spur is 2810 Ω, i.e. the resistance ratio to the main mode is about 57 dB !

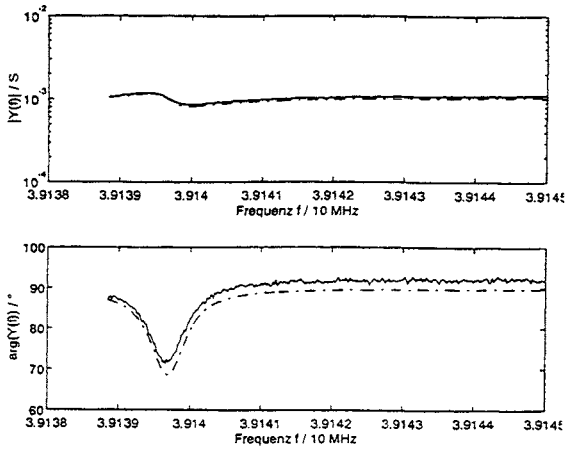


Fig. 5: spurious resonance #5 39.139,636 kHz

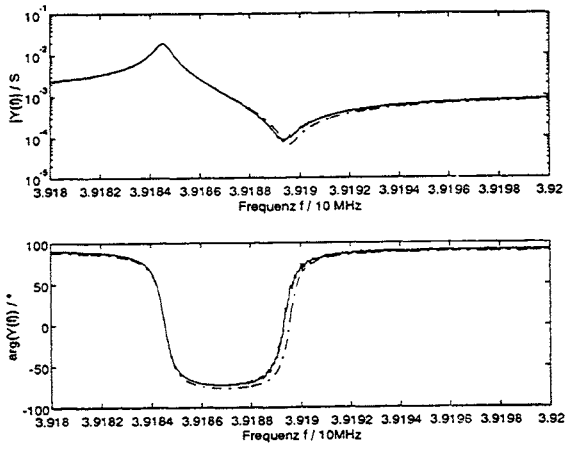


Fig. 6: spurious resonance #6 39.184,517 kHz

4. Summary

The paper describes an effective algrithm to measure spurious resonances of crystal units by using the p- network system. The measurement procedure is fast and the results show a very good agreement between the measured data and the response simulated with the computed crystal parameters. A dynamic range of 60 dB of the resonance resistances can be measured.

5. Acknowledgements

The author thanks Friedrich Krug and Jürgen Agrikola (University of Erlangen), for the good cooperation and the experimental verfication.

6. References

[1] Neubig, Bernd W.: Measurement of Quartz Crystal Units up to 500 MHz and Above by the Use of a Pi-Network with Error Correction; Proc. 11th Quartz Devices Conference, Kansas City (1989), Vol.1, p.60ff.

[2] IEC publication 444-5 (1993): Methods for the measurement of quartz crystal devices for the determination of equivalent electrical parameters using automatic network analyser techniques and error correction

[3] Neubig, Bernd: Measurement of Spurious Resonances of Crystal Units Using Network Analysers with Error Correction ;Proc. 8th European Frequency and Time Forum (EFTF) (1994), pp.103-109

SAW

Rotated Y-cut SAW Devices Immune to the Thermal Shock Due to Intermittent Laser Beam Irradiation

Nobunori Oura, Makoto Yamakita,
Michiaki Takagi* and Eishi Momosaki*

Precision and Intelligence Laboratory,
Tokyo Institute of Technology
4259 Nagatsuta, Midoriku, Yokohama 226 Japan
* R&D Division, Seiko Epson Corp. Suwa 392, Japan

1 Introduction

Although the demand for the stable frequency sources is growing with the fast development of the mobile communications [1], SAW oscillators are not used as the stable frequency source in such equipment. This is because the frequency stabilities of the SAW oscillators using existing SAW resonators are inferior to those of the crystal oscillators by one to three orders of magnitude. One of the reasons for this inferiority lies in the design concept of the resonators, in which the static frequency temperature characteristics are considered more important than the dynamic frequency temperature characteristics, namely the transient thermal frequency behavior of the resonators.

One of the ways to overcome the frequency instabilities due to the transient thermal behavior of SAW resonators is to find angles of cut immune to thermal shocks. There are several papers dealing with angles of cut immune to transient thermal shocks [2][3].

A method for measuring the thermal frequency excursion due to thermal shock using a laser beam as a thermal probe has been reported [4]. In that study, a series of maps of equi-frequency excursion, due to the thermal shocks similar to contour maps was drawn. Additionally, the thermal frequency behavior of a SAW device due to laser beam irradiation has been analyzed under several assumptions [5].

This paper describes a numerical analysis of the thermal behavior of a SAW device plate and the angles of cut immune to the thermal stress induced by laser beam irradiation obtained. The thermal stress immune angles of cut were obtained from the analysis with its underlying assumptions.

2 Theoretical

2.1 Resonance frequency of a SAW resonator without thermal stress

The resonance frequency f of a SAW resonator is determined by the propagation velocity v of the SAW and the pitch of the inter-digital exciting electrode p as follows;

$$f = v/2p \tag{1}$$

The velocity was obtained by conventional theory using the numerical computation method described in [6].

2.2 Resonance frequency of a SAW resonator under thermal stress

A He-Ne laser beam with a power of 20mW was employed to cause thermal stress in a quartz crystal plate. In the analysis, it is assumed that the diameter of the laser beam is 0.4mm and the absorption coefficient of quartz is 50%. In a previous study, the diameter of the beam was assumed to be infinitesimal [5]. For computation using FEM(MSC/NASTRAN), the quartz plate of 12 by 3 by 0.5mm was divided into elements of 0.3 by 3/16 by 0.25mm.

The thermal stress generated on the surface of the plate by a heat source of heat capacity Q at an instance $t = 0$ of the irradiation at the origin of the plate as an infinite half space elastic body is obtained from the heat distribution in the body using FEM, taking into account the anisotropic thermal conductivity of quartz.

Figure 1 shows the heat distribution on the plate after just 0.8s of irradiation by the laser beam at the center of the exciting electrode. The specimen plate is cut out with an angle of $(\phi = 0°, \theta = 96.51°, \psi = 33.79°)$ in right-handed Eulerian coordinates. The pattern of heat diffusion takes the form of an ellipse owing to the

anisotropy of the thermal conductivity of quartz. The axis of the ellipse tilts at the angle of cut of the plate with respect to the length of the plate.

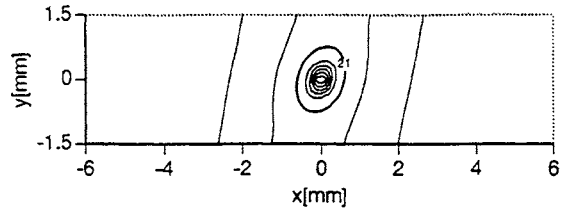


Figure 1: Heat distribution on the plate after 0.8s of the irradiation(This separation of the contour lines corresponds to 0.375°C)

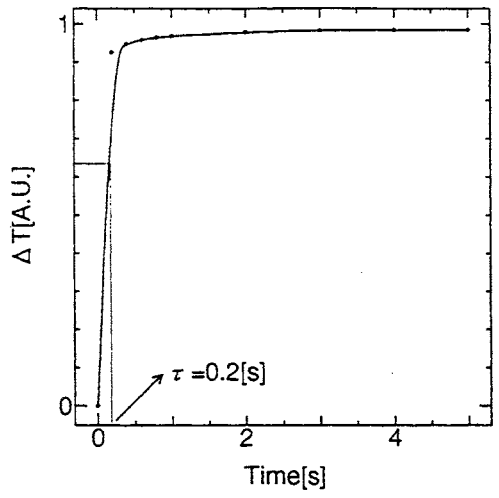


Figure 2: Step response of the plate to the irradiation

Figure 2 shows the temperature difference, which was obtained by computation, between the irradiation center, that is, the center of the electrode and a point separated by 1mm from the irradiation center as a function of irradiation time. The rise time of the curve is estimated as about 0.2s. The rise time indicates that the plate responds to frequencies of thermal fluctuation less than 5Hz.

The thermal stress distribution in the plate after the irradiation was computed by FEM using the heat distribution as a function of time, while taking into account the anisotropy of elasticity of the plate. The elastic constants under stress were calculated, using second and third order elastic constants, displacement gradients and strains [7]. The stress in the wave propagation direction at a point 1mm from the center of the exciting electrode was used as a typical stress for computing the wave velocity at 0.8s after the irradiation. The frequency excursion due to the irradiation was computed by subtracting the frequency without stress from that under stress.

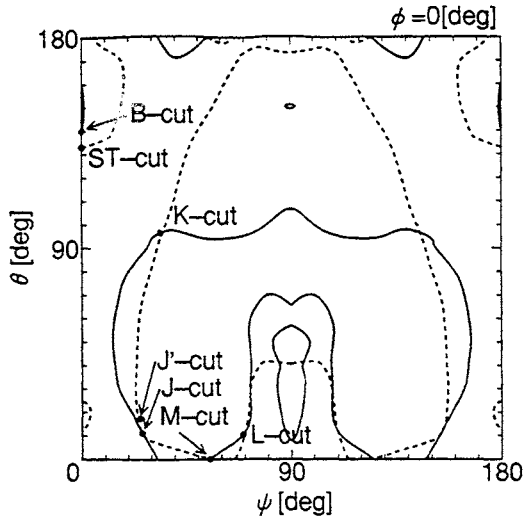


Figure 3: Loci of the angles of cut immune to the thermal stress

Figure 3 shows a chart of the loci of angles of cut immune to the thermal stress ,on which the loci of angles of zero-frequency temperature coefficient(ZTC) of the SAW at 20°C [8] are superposed.

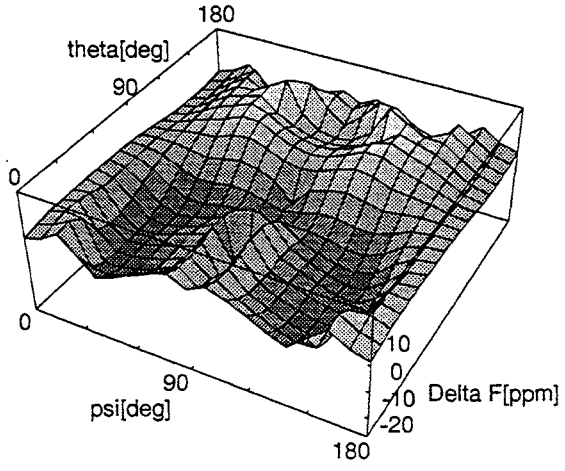


Figure 4: Three dimentional display of the loci

The sensitivity of the loci to the angles of cut must be cleared beforehand for making the specimen plates immune to the stress. A three dimentional display of this sensitivity is shown in Fig.4.

The loci are lines generated by crossing the frequency excursion plane of the plate with the zero-frequency excursion plane.

The gradient of the frequency excursion plane in the vicinity of the loci is the sensitivity with respect to the angle of cut on making the device plates. The looser the gradient is, the easier the manufacture of the corresponding device plate.

3 Experimental

The specimen plates immune to the stress are selected from Fig.3 from the standpoint of ease of manufacture and are shown in Table 1. The ZTC temperature is determined theoretically with respect to a plate in the free state. However, it is well known that the ZTC temperature depends upon the electrode film thickness. Several kinds of specimen plate were made based on these considerations.

Table 1: Specimens of K-cut group ϕ=0°,θ=96.51°

	Ka-cut	Kb-cut	Kc-cut
ψ[°]	29.36	32.50	33.79

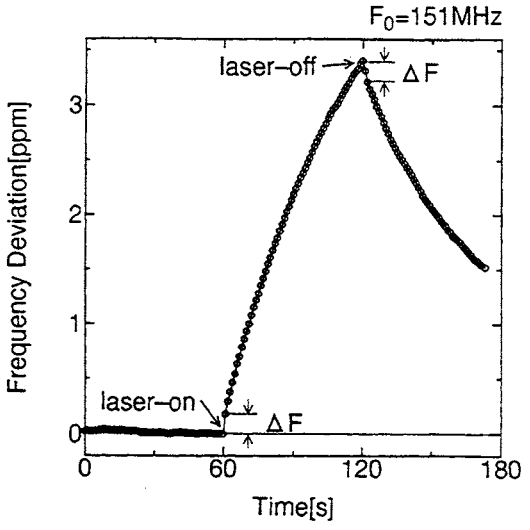


Figure 5: Thermal excursions of the specimen Ka-cut

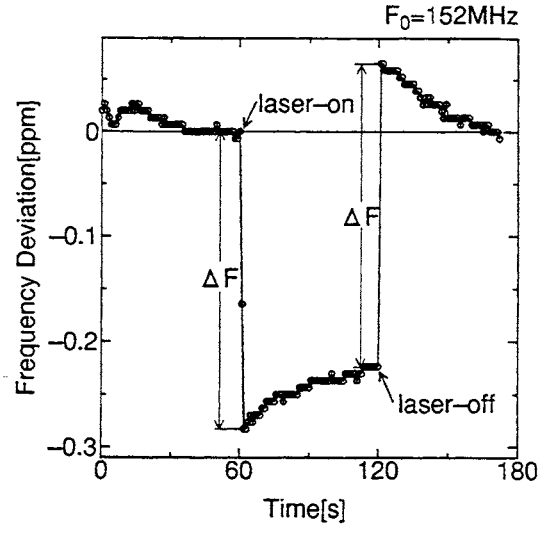


Figure 6: Thermal excursions of the specimen Kb-cut

Figures 5 and 6 show the frequency excursions due to laser beam irradiation on the center of the specimen plates with respect to angles of Ka-cut and Kb-cut, respectively. According to the computed frequency excursions, the magnitude of the excursion is in the following order: Ka-cut, Kb-cut, Kc-cut. The measured values of the excursion are in this order, too, although the absolute values of the excursion are different from those predicted by computaion. The main value of the computation is its predictive ability for determining the relative order of the frequency excursion magnitudes for different cut angles. Determination of absolute excursion magnitudes is less important.

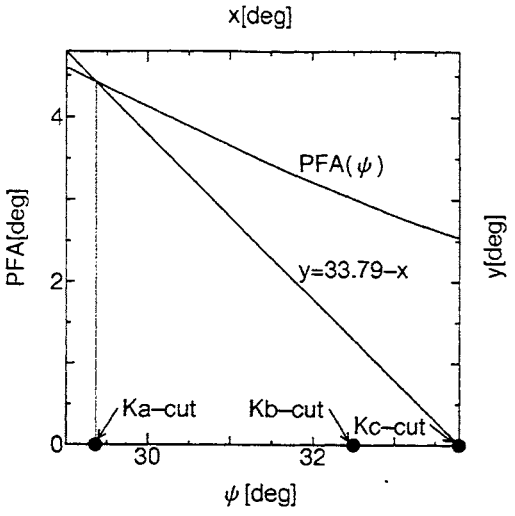


Figure 7: The loci of the angles and power flow angles

Figure 7 shows a plot of the excursion with respect to the angle of cut and the power flow angle. This plot indicates that there is an angle immune to the stress. It is considered that the discrepancy between the predicted angle and the specimen's angle results from the neglection of the power flow angle on selecting the angles of cut. Correction of the angle of cut immune to the stress is possible according to the results shown in Fig.7.

Figure 8 shows the oscillating frequency trends of the Ka-cut, Kb-cut plate and the ST-cut plate. Figure 9 shows Allan variance of the Ka-cut, Kb-cut plate and the ST-cut plate. The frequency stability of K-cut plate is superior to that of the ST-cut plate in a short term of τ less than a few seconds, where the environmental temperature fluctuation has great influence on the short term frequency stability.

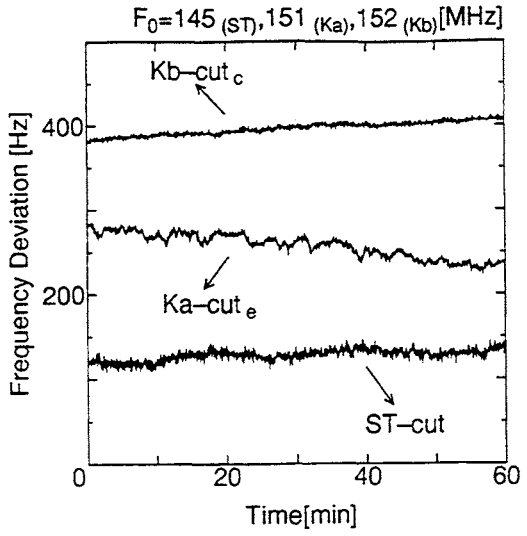


Figure 8: Oscillating frequency trends of the Ka,Kb and ST-cut plate(τ=2s)

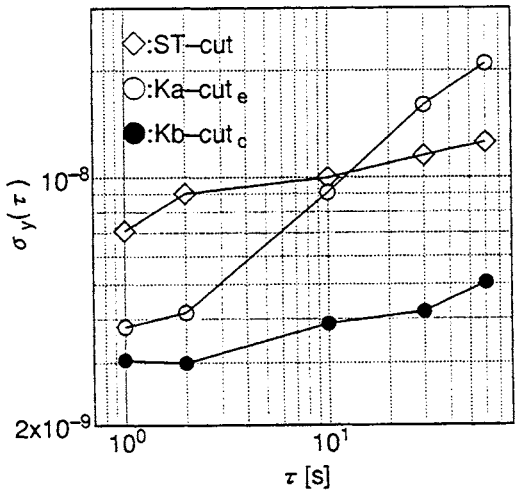


Figure 9: Allan Variance of the Ka,Kb and ST-cut plate

4 Conclusion

The loci of the angles of cut immune to the stress were obtained by computation with the FEM, taking into account the anisotropy of the physical constants of quartz and assuming that the phase velocity of the wave was governed only by the stress due to irradiation in the wave propagation direction.

Specimens prepared according to design considerations obtained from the FEM computations show frequency excursion characteristics immune to stress and oscillating frequency stabilities with respect to time one order of magnitude superior to those of ST-cut plates in a short term. By improving the transient thermal behavior of the plates, their possible use as mass sensors in hygrometers or as odor sensors may be realized.

References

[1] C.C.W.Ruppel et al.,"SAW Devices for Consumer Communication Application," IEEE Trans. UFFC, vol.40, No.5, pp.438-452, September 1993

[2] E.K.Sinha;"A Stress Temperature Compensated Orientation and Propagation Direction for Surface Acoustic Wave Devices," IEEE Trans. UFFC, vol.UFFC34, No.1, January, pp.6474,1987

[3] S.Ballandras and E.Bigler,"Doubly Rotated SAW Quartz Cuts with Low Sensitivty to Stress and Temperature Effects Exibiting a Power Flow Angle Close to Zero," J.Acoust.Soc.Am., No.95, Vol.3, pp.1390-1395, March 1994

[4] N.Oura et al.,"Transient Thermal Frequency Behavior of SAW Devices with a Laser Beam as a Thermal Probe," Proc. 7th EFTF,pp.239-242,1993

[5] N.Oura and T.Nakase;"Theoretical Analysis of Thermal Frequency Behavior of a SAW Device, "Proc. 1994 IEEE IFCS, pp.323-329, June 1994

[6] The 150th Committee of Japan Society for the Promotion of Science ed.;"Elastic Surface Wave Devices Handbook, " Ohom Press,Tokyo,1991 (in Japanese)

[7] R.N.Thurston et al.;"Third Order Elastic Coefficients of Quartz,"J.Appl.Phys., 371,pp.267-265,1966

[8] Y.Shimizu and Y.Yamamoto;"SAW Propagation Characteristics on α-Quartz with Arbitrary Cut," Trans. vol.J63A11, pp.799-806, 1980 (in Japanese)

INFLUENCE OF THE ALUMINIUM LAYER THICKNESS ON THE PROPERTIES OF BLEUSTEIN-GULYAEV WAVES IN $Li_2B_4O_7$ CRYSTAL

Waldemar Soluch

Institute of Electronic Materials Technology
Wolczynska 133, 01-919 Warsaw, Poland

ABSTRACT

Properties of Bleustein-Gulyaev waves (BGW) in $Li_2B_4O_7$ crystal were determined from the measured parameters of a delay line with double electrode interdigital transducers (IDT). Since the frequency response at the third harmonic was in good agreement with the SAW equivalent circuit model, it was possible to determine the BGW parameters. It was found, that the coupling coefficient of BGW is very sensitive to aluminum layer thickness variations. Good agreement was found between calculated and measured parameters of BGW from an extrapolation to the zero layer thickness.

1. INTRODUCTION

Lithium tetraborate ($Li_2B_4O_7$), belongs to the 4mm class of the tetragonal system. Bleustein-Gulyaev waves (BGW) exist in this class if the propagation direction is perpendicular to the [001] axis and if this axis is parallel to the propagation plane [1].

Let's assume, that the BGW propagate in the x_1 direction (Fig.1), the x_3 axis is parallel to the [001] crystal axis, and x_2 axis is perpendicular to the propagation plane. Then, the mechanical displacement u_3 can be written as [1]

$$u_3=u_o\exp(-\alpha k x_2)\exp(-jkx_1), \qquad (1)$$

where, u_o is the amplitude of mechanical displacement at the crystal surface ($x_2=0$), α is the decay constant ($\alpha>0$ for $x_2>0$), $k=2\pi/\lambda$ and λ is the wavelength. The decay constants α_f and α_m, for free and metallized surface, respectively, are given by the following expressions

$$\alpha_f=\frac{k_{15}^2}{1+\varepsilon_{11}^S/\varepsilon_o}, \qquad (2a)$$

$$\alpha_m=k_{15}^2, \qquad (2b)$$

where

$$k_{15}^2=\frac{e_{15}^2}{\bar{c}_{44}\varepsilon_{11}^S} \qquad (3)$$

and

$$\bar{c}_{44}=c_{44}^E+e_{15}^2/\varepsilon_{11}^S. \qquad (4)$$

Here k_{15} is the electromechanical coupling coefficient for bulk shear wave, e_{15} is the piezoelectric constant, c_{44}^E is the elastic stiffness at constant electric field E, and ε_{11}^S is the dielectric permittivity at constant strain S.

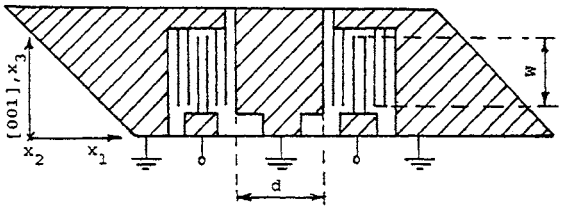


Fig.1 BGW delay line on $Li_2B_4O_7$ crystal.

The BGW velocities v_f and v_m, for the free and metallized surface (infinitesimally thin perfect conductor), respectively, can be determined from the expressions

$$v_f=v_S\sqrt{1-\alpha_f^2}, \qquad (5a)$$

$$v_m=v_S\sqrt{1-\alpha_m^2}, \qquad (5b)$$

where

$$v_S=\sqrt{\bar{c}_{44}/\rho}. \qquad (6)$$

Here v_S is the shear bulk acoustic wave velocity and ρ is the mass density. The penetration depths, d_f and d_m, for free and metallized surface are defined as

$$d_f=\lambda/(2\pi\alpha_f), \qquad (7a)$$

$$d_m=\lambda/(2\pi\alpha_m). \qquad (7b)$$

The coupling coefficient k_{BG} for BGW can be determined from the expression

$$k_{BG}^2=2(v_f-v_m)/v_f, \qquad (8a)$$

which is equivalent to [2]

$$k_{BG}^2\cong k_{15}^4. \qquad (8b)$$

Using the $Li_2B_4O_7$ constants reported recently [3]: mass density $\rho=2432kg/m^3$, relative dielectric permittivity $\varepsilon_{11}^S/\varepsilon_o=8.66$, $c_{44}^E=57.1GPa$ and $e_{15}=0.39C/m^2$, it was calculated, that: $v_S=4928.9m/s$, $k_{15}^2=0.033$, $k_{BG}^2=0.0011$, $v_f=4928.9m/s$, $v_m=4926.1m/s$, $d_f\cong45.5\lambda$ and $d_m\cong4.7\lambda$.

If we compare the above calculations with published experimental results [4]-[7], we see that good agreement exists for velocity of BGW, while large differences are for the coupling coefficient. For example, the measured BGW velocities were about 4900m/s [4] (Fig.6, θ=0) and 4920m/s [5]. The mesured square of the effective coupling coefficients were equal to about 0.0028 [4] and 0.0027 [5]. It was also shown, that the thickness of the aluminum layer has very large influence on the coefficient [6]-[7].

It is the purpose of this paper to determine the BGW parameters from the measured frequency responses of a delay line fabricated on the $Li_2B_4O_7$ substrate.

2.EXPERIMENTAL PROCEDURE AND RESULTS

A delay line, which consisted of two identical, unapodized, double electrode interdigital transducers (IDT), was made by the lift-off method (Fig.1). The whole crystal surface was metallized to obtain small penetration depth and the ends of the substrate were angled by 45° to reduce reflections. The electrical period of the IDT was equal to 48μm, electrodes were 6μm wide and the number of electrodes was equal to 202. The aperture W and the distance d between the IDTs were equal to 2mm.

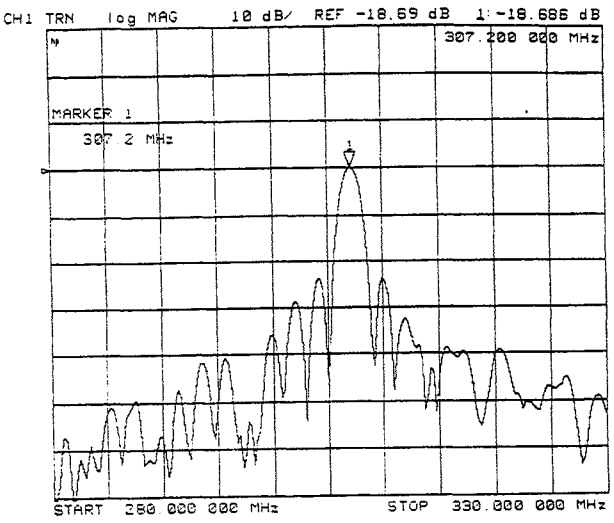
Periodical grooves were cut at the bottom surface and epoxy resin was applied at the ends of the substrate to scatter and attenuate the reflected bulk and BGW, respectively. To estimate the influence of the aluminum layer thickness h on the BGW parameters, measurements were done for h=0.20μm and h=0.12μm.

The frequency responses of the delay line were measured in a 50Ω system (HP Network Analyzer 8752A) at the third harmonics, since at the first ones the responses where deformed by bulk waves [5]. The measured frequency responses are shown in Fig.2. The passbands of these responses are in good agreement with the SAW equivalent circuit model, therefore it is possible to determine velocity and effective coupling coefficient of BGW. Velocity is directly related to the central frequency and coupling can be evaluated from the measured insertion loss.

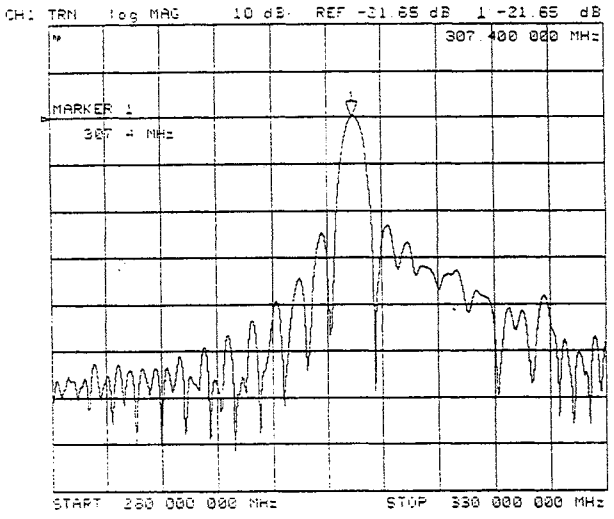
The ratio of the output voltage for the case when the delay line is inserted between source and load, to the output voltage for the case when the load is directly connected to the source, can be written as [8]-[9]

$$A=\frac{(y_1+y_2)y_{12}}{y_{12}^2-(y_1+y_{11})(y_2+y_{22})}, \qquad (9)$$

where y_1 and y_2 are the source and load admittances, respectively, y_{11} and y_{22} are the input admittances of the IDTs, and y_{12} is the transfer admittance. At the central frequency f_o, we can write [10]



(a)



(b)

Fig.2 Frequency responses of the delay line at the third harmonic. (a) h=0.20μm. (b) h=0.12μm

$$y_{11}=G_oN^2+j\omega_oC_o, \qquad (10)$$

where G_o is a function of substrate and IDT parameters, N is the number of electrode overlaps, and C_o is the static capacitance. In the case of a delay line with two identical IDTs, we have $y_{22}=y_{11}$ and $y_{12}=G_oN^2$ [9].

Since insertion loss IL is usually expressed in dB, then

$$IL=-20\log|A|. \qquad (11)$$

The results of measurements and calculations are presented in Table 1. Here, IL_o is the insertion loss measured at the central frequency f_o. Velocity v and the square of the effective coupling coefficient k_e^2 were determined from the condition of agreement between

Table 1. Results of measurements and calculations of the BGW delay line parameters (λ=16μm).

h[μm]	0.20	0.12	0
h/λ	0.0125	0.0075	0
f_o[MHz]	307.2	307.4	307.7
IL_o[dB]	18.5	21.5	28
v[m/s]	4915.2	4918.4	4923.2
k_e^2	0.0025	0.0019	0.00083

measured and calculated f_o and IL_o for a given λ and h. The values of f_o, v and k_e^2 for h=0 were obtained from a linear extrapolation.

To determine the v_f and v_m velocites from the results for h=0, we assume, that for the IDTs with equal gap and electrode widths $v=2(v_f v_m)/(v_f+v_m)$. We also assume that $k_e^2=2(v_f-v_m)/v_f$ (for h=0). It means that $k_{BG}^2=k_e^2$ in this case. From the above expressions we obtain v_f=4924.2m/s and v_m=4922.2m/s.

From (8), (2) and (7) we obtain k_{15}^2=0.0288, d_f=53.4λ and d_m=5.5λ, respectively. Finally, from (5), (4) and (3) we have v_s=4924.2m/s, c_{44}^E=57.3GPa and e_{15}=0.37C/m^2, respectively. The above values of c_{44}^E and e_{15} are in very good agreement with other published data [3].

4. CONCLUSION

The third harmonic frequency response of a BGW delay line with double electrode IDTs was in good agreement with the SAW equivalent circuit model. Therefore, it was possible to determine the BGW parameters from the measured response. It was experimentally confirmed, that the effective coupling coefficient of BGW in $Li_2B_4O_7$ is very sensitive to the aluminum layer thickness variations. For h/λ=0.0125, the coefficient is about 3 times larger than that for h=0. It can be explained by a large difference between the shear acoustic wave velocities in aluminum (about 2920m/s) and in $Li_2B_4O_7$ substrate (about 4920m/s). Good agreement was found between the c_{44}^E and e_{15} constants obtained from BGW parameters and measured by other methods. It means, that for h=0, the expression $k_{BG}^2=2\Delta v/v$ for the coupling coefficient is also correct for BGW.

High velocity and reasonable coupling make the BGW in $Li_2B_4O_7$ crystal promising for chemical sensor and high frequency filter application.

ACKNOWLEDGMENTS

I would like to thank Z.Luczynski, T.Lukasiewicz, J.Kisielewski and W.Hofman for lithium tetraborate crystal and substrate fabrication, and to M.Teodorczyk, E.Lipinska, H.Majewska, P.Naglowski and T.Wrobel for delay line fabrication and measurements.

This work was supported by the State Committee for Scientific Research (KBN) under grant number 8 8009 92 03.

REFERENCES

[1] J.L.Bleustein,"A new surface wave in piezoelectric materials," Appl. Phys. Lett., vol.13, No.12, Dec. 1968, pp.412-413.

[2] W.Soluch, "Theory of amplification of Bleustein-Gulyaev waves in tetragonal and hexagonal crystals," IEEE Trans. Sonics and Ultrasonics, vol.SU-24, No.1, Jan. 1977, pp.43-48.

[3] J.A.Kosinski, A.Ballato, and Y.Lu, "Pure-mode measurements of dilithium tetraborate material properties," in 1993. Ann. Freq. Cont. Symp. Proc., 1993, pp.359-370.

[4] M.Adachi, T.Shiosaki, H.Kobayashi, O.Ohnishi,and A.Kawabata, "Temperature compensated piezoelectric lithium tetraborate crystal for high frequency surface acoustic wave and bulk wave device applications, "in IEEE Ultrason. Symp. Proc., 1985, pp.228-232.

[5] W.Soluch, "Properties of Bleustein-Gulyaev waves in $Li_2B_4O_7$ crystal," IEEE Ultrason.Symp.Proc., 1993, pp.237-241.

[6] T.Sato and H.Abe, "BGS waves on $Li_2B_4O_7$ substrate," 1993 Autumn Meeting of the Institute of Electronics, Information and Communication Engineers of Japan, p.1-251.

[7] T.Sato and H.Abe, "Temperature stable propagation properties of BGS waves on $Li_2B_4O_7$ with SiO_2 film overlays," Technical Report of IEICE, 1994,pp.9-15.

[8] G.L.Matthaei, D.Y.Wong, and B.P.O'Shaughnessy, "Simplification for the analysis of interdigital surface-wave devices," IEEE Trans. Sonics and Ultrason., vol.SU-22, Mar.1975,pp.105- 114.

[9] A.J.DeVries, "Surface wave bandpass filters" in *Surface Wave Filters*, H.Matthews, Ed., New York: John Wiley & Sons, 1977, p.263.

[10] H.Engan, "Surface acoustic wave multielectrode transducers," IEEE Trans. Sonics and Ultrason. vol.SU-22, Nov. 1975, pp.395- 401.

Design and Optimization of a Low Cost SAW Stabilised Fixed Frequency Oscillator

Uwe Stehr*, Bernd Adler**, Heinold Kostner**

* Technical University of Ilmenau, D-98684 Ilmenau, Germany
** TELE QUARZ Group, D-74924 Neckarbischofsheim,Germany

Contents of this article is the analysis and optimization of a Low Cost Fixed Frequency Oscillator with SAW-Resonator (PSO - Packaged SAW Oscillator). Not only good electrical characteristics (low phase noise, low harmonic distortion without additional filter, low ageing, low current consumption, moderate output, optional possibility for ASK and FSK) were required but also technological parameters (low costs of elements, little space required, low or minimum adjustment in production, high yield in production).

Concept of circuit

There are various circuit concepts for the design of SAW oscillators /1-4/. You can feedback an amplifier (transistor or MMIC) with a SAW resonator and a phase-matching network to start the oscillation.

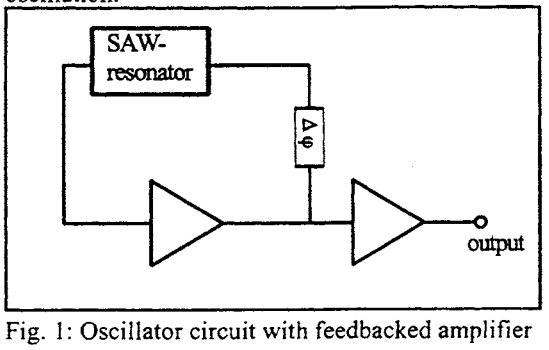


Fig. 1: Oscillator circuit with feedbacked amplifier

For practical purposes this concept, however, requires more work to adjust the loop phase to the necessary phase shift of $\varphi = 0°$ or $n * 360°$. A phase shifting network is usually necessary. Moreover the attenuation of the harmonics is relatively small (typically 10dB...15dB) and for the point of signal output a buffer amplifier is required. Besides the SAW drive level is relatively high since the SAW resonator is directly placed in the feedback circuit of the oscillator amplifier. Another possibility is using the Collpits circuit with a transistor in common base circuit. As a consequence of this design the SAW resonator will ground the base of the transistor what finally ends in oscillation. This circuit generally fulfils the requirements of the PSO.

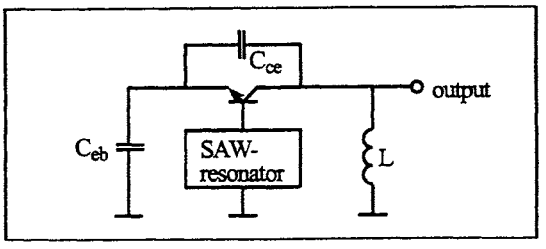


Fig. 2: Collpits oscillator circuit

The low SAW drive level secures low ageing, through the selective feedback over a parallel LC circuit a high harmonic distortion is achieved. No additional buffer amplifier for the point of signal output is required so there are less components needed which can be easily mounted on a small PCB. Thus the Collpits circuit seems to be the most favourable concept for the realisation of a Low Cost Fixed Frequency Oscillator with SAW resonator.

Analysis of circuit

The analysis of circuit was made by the CAE simulation programme MDS (Microwave and RF Design System) of Hewlett-Packard. This analysis was aimed at optimizing the circuit referring to electrical characteristics and based on this, a tolerance analysis to optimize the yield in production.
For the analysis the circuit was considered as an open oscillator loop (fig. 3). In this case the oscillator criteria are loop gain > 1 with loop phase = 0°.

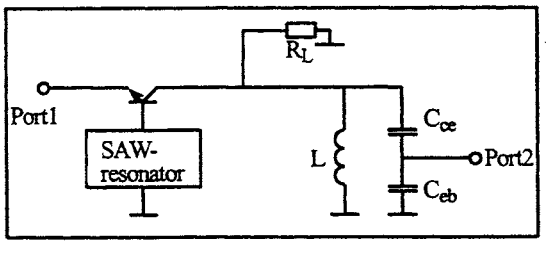


Fig. 3: Collpits Oscillator model for small signal analysis

As equivalent circuit for the SAW resonator measurement results from the network analyser were used. The results of the analysis were compared to

values of an empirical design of the oscillator. For this initial design the forward transmission gain of the open oscillator circuit was measured with the network analyser to compare the theoretical simulation results.

Small signal analysis

First, a small signal analysis (S-Parameter analysis) was made to analyse the start up characteristic of the oscillator. Please take the results of this first analysis from fig. 4.

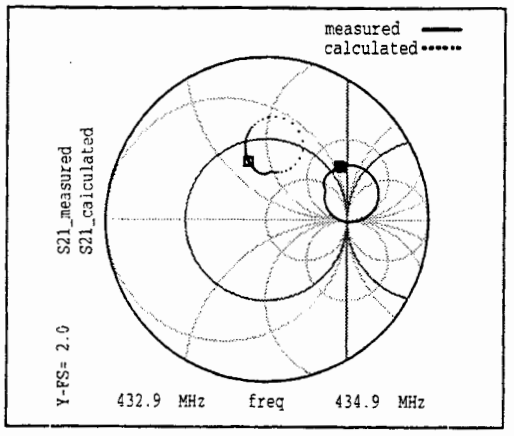


Fig. 4: Small signal analysis of the open loop

As shown in fig. 4 the oscillation point (S_{21}= 1) is not covered by the calculated curve so that the criterion of oscillation are not fulfilled according to the analysis. The measured diagram of the loop gain covers the point S_{21} = 1.
This means that the model of the oscillator used in the analysis does not fit enough. Possible sources of error are absent parasitics in the circuit layout and signal delays of the printed circuit.
Since the parasitic capacities of the transistor are already taken into consideration by the existing device model, especially the parasitic elements of the coil are of importance. For the inductivity measurements revealed a series resistance of 0.2Ω and a shunt capacitance of 0.4pF.

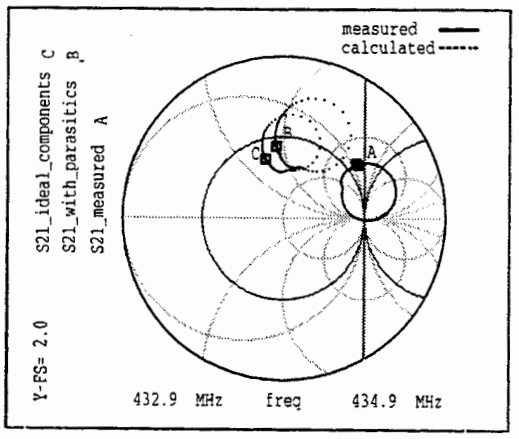


Fig. 5: Small signal analysis with parasitics

Fig. 5 shows the result of the small signal analysis taking into consideration these parasitic elements.
The simulation shows that taking the parasitic elements into account is not enough to get an exact description of the circuit. Therefore the mechanical dimensions of the printed layout were included into the analysis. Especially the length of line between the transistor output (collector) and the feedback circuit determines the loop phase substantially.
Fig. 6 shows simulation results for various line lengths. Best accordance between measurement results and simulation results were achieved with a line length of 11.1mm.

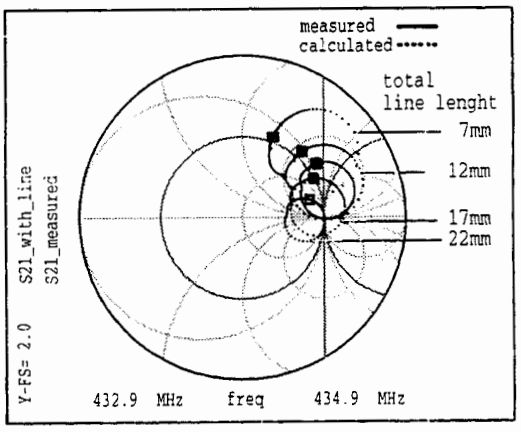


Fig. 6: Small signal analysis including the transmission lines of the PCB

Large signal analysis

A large signal analysis of the oscillator circuit was made for determination of the output power and the harmonic distortion of the oscillator. First a description of the SAW resonator had to be found which is fitted in the whole frequency range of the harmonic distortions. A good way for more complex SAW elements is to calculate an equivalent circuit from the measured data **/5/.**

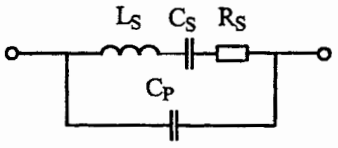


Fig. 6: Equivalent circuit of the SAW resonator

Because of the relative simple equivalent circuit of the SAW resonator, the following possibility was chosen: The frequency range is split in a near resonance and wide up range. The resonant frequency range (f_{res} ± 1MHz) is simulated with a measured set of S-parameters which is embedded as a two-port in the simulation circuit. The wide up range is described as a simple shunt capacitor.
For the simulation of the large signal behaviour of the open loop a Harmonic Balance analysis was made with variable input power. From the simulated port voltages and currents S-parameter sets were

calculated for the fundamental frequency. For checking the simulation results, these power dependent parameters were compared with S-parameters of the open oscillator loop, measured with the network analyser for variable input power. Good fitting of the curves can be seen.

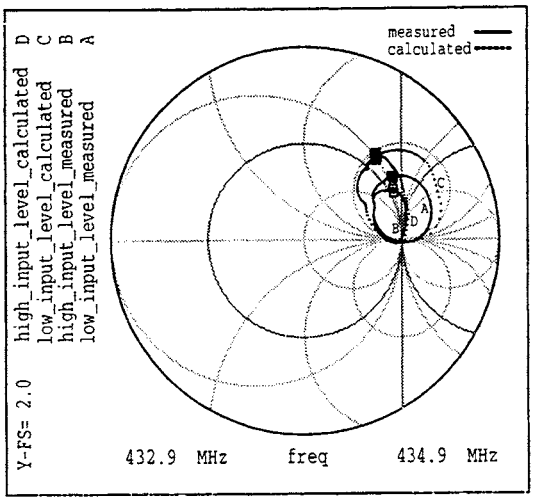


Fig. 7: Harmonic Balance analysis of the open loop

This equivalent circuit of the SAW oscillator serves as a base for an oscillator simulation with the help of the oscillator simulation tools in the simulation program. The oscillation frequency, output power and harmonic distortion of the closed loop are calculated by means of a special Harmonic Balance procedure.
With the help of this algorithm the following results of the test oscillator were calculated:
f_{osc} = 433.8954 MHz, P_{out} = 11.2 dBm,
$P_{1.OT}$ = -19.2 dBm.
The results of the oscillation frequency and the output are in accordance with the measured values of the test oscillator (f_{osz} = 433.8949 MHz, P_{out} = 12.2 dBm) from the spectral analyser. The measured output power of the first overtone ($P_{1.OT}$ = -10.1 dBm), however, is higher than simulated.

We can proceed from the fact that the equivalent circuit we found is suitable for the analysis of the SAW oscillator and can be used for a tolerance analysis.

Statistical analysis

For the statistical analysis C_{eb}, C_{ce} and L were included because they have the most important influence on the behaviour of the oscillator. Fig. 8 shows the dependence of oscillating frequency and output power from varying one of these elements.

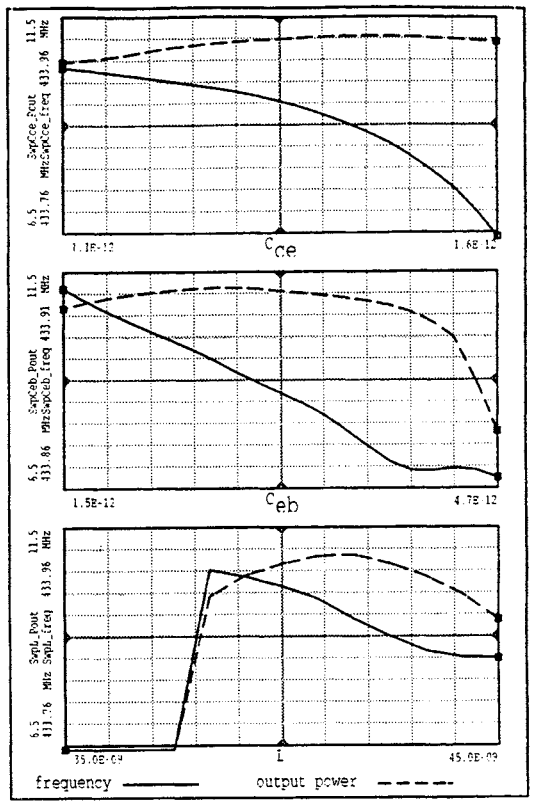
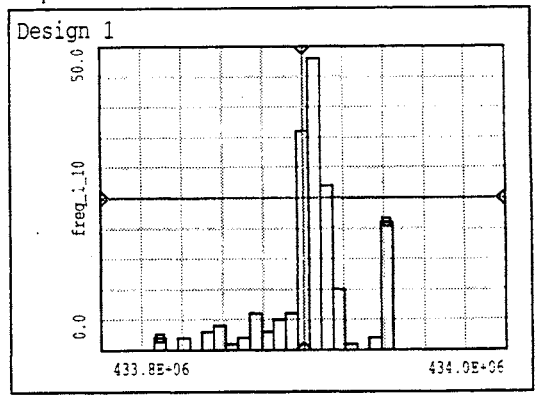


Fig. 8: Dependence of oscillating frequency and output power from varying C_{ce}, C_{eb} and L

In a first step, an analysis of the oscillator circuit with several values of these components was made. Three designs with good (and similar) frequency and output power parameters at nominal values of C_{eb}, C_{ce} and L were used for the statistical analysis. The calculated results of these designs were:

design 1: f_{osc}=433.906MHz
P_{out}=10.7dBm
design 2: f_{osc}=433.901MHz
P_{out}=10.9dBm
design 3: f_{osc}=433.913MHz
P_{out}=10.7dBm.

A gaussian model was used to describe the statistical distribution. For the capacitors a 5% tolerance range was assumed where the values with a tolerance smaller then 1% were excluded. The tolerance range of the coil was assumed to be 10%. Figure 9 shows the distribution of the calculated oscillating frequencies of 201 sets of values.



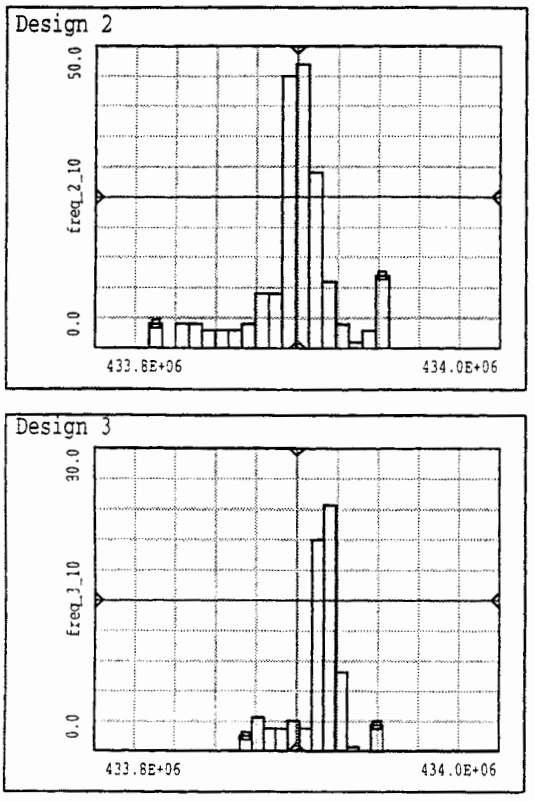


Fig. 9: Distribution of the calculated oscillating frequency of 201 sets of values

As criterion to perform a good-bad decision (yield analysis) two equations were used:

$$|f_{osc}-f_{res}| < 10ppm$$
$$P_{out} > 10dBm$$

(f_{res} center frequency of the SAW resonator)

With this criterion the calculated yield of the different designs was the following:
design 1: 70.1%
design 2: 78.1%
design 3: 89.6%.
It is to be seen that there is a significant difference in the yield of the design variants though they had similar behaviour at nominal values. Design 3 offers the best yield but for a low cost production a yield of about 90% is already too small. So design 3 was calculated again using now a 5% tolerance value of the coil. With this circuit configuration a yield of 99.5% was calculated. That means it is necessary to use a coil with a 5% tolerance range to fit the requirements of a low cost production. Another way to increase the production yield is to adjust the circuit. A worst case calculation of design 3 showed that it is possible to fulfil the yield equations at any case with a one-step adjustment of one of the capacitors but with the loss of up to 5dB of harmonic distortion compared to the nominal value.

Conclusions

The results above show that a statistical analysis is very useful to find a design variant that offers both: a good electrical performance (fig. 10) and also a high yield in the production process which is very important for low cost products.
With an optimal design an increase of the production yield from about 70% of the initial design to nearly 100% could be reached for the investigated fixed frequency SAW oscillator.

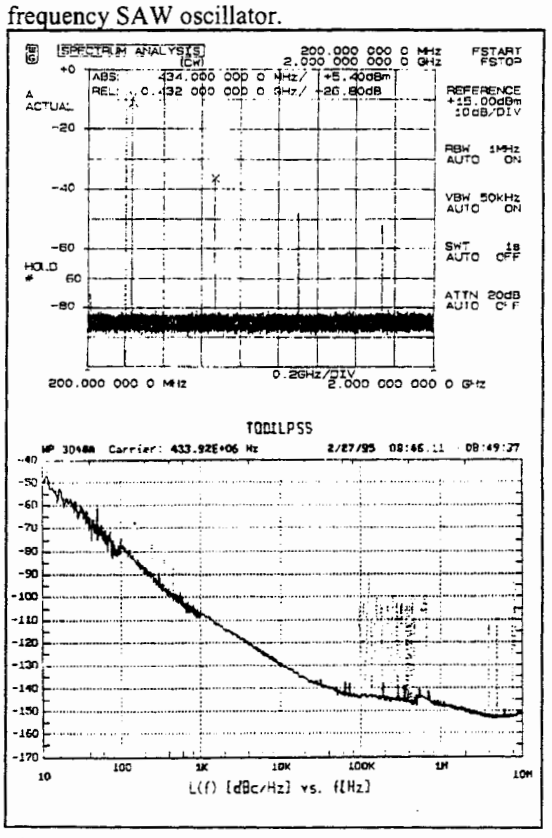


Fig. 10: Typical plots of output spectrum and phase noise of the investigated SAW oscillator

Literature

/1/ Colvin, R.D.: UHF acoustic oscillators, Microwave Journal, Nov. 1980, pp. 22-23

/2/ Westbrook, R.M.: The Potential of SAW Resonators in RF Oscillators, Electronics Test, December 1982, p. 136

/3/ Wilkinson, V.J.; Durrant, P.A.: SAW Oscillators as receiver Components, Fifth International Conference on Radio Receivers and Associated Systems, London 1990

/4/ Parker, T.E.: Precision SAW Oscillators, IEEE Ultrasonics Symposium Proceedings 1982, pp. 268-274

/5/ Beier, J.; Neubig, B.: Measurement of the Equivalent Circuit Parameters of Transversely Coupled SAW Resonator Filters, IEEE Ultrasonics Symposium, Cannes, France, 1994, Paper CC-6

SPACE APPLICATIONS SYNCHRONIZATION

PRECISE TWO-WAY MICROWAVE TIME TRANSFER TO SPACE-BORNE ATOMIC CLOCKS

S. BEDRICH, J. HAHN

DLR Oberpfaffenhofen, Institute of Radio Frequency Technology, D-82230 Weßling

The paper presents a new system for highly precise two-way time transfer. Possible time transfer techniques are introduced which include all error corrections. The system also offers the potential for one-way time dissemination. In combination with a polar-orbiting H-maser it fulfills all requirements for various applications in the fields of time metrology, navigation, fundamental physics, and as a basic element in a future GNSS II.

1 Introduction

For any future satellite, space-borne communication network, or navigation satellite system, a reliable, ultra-stable, onboard time reference will become indispensable. Equally important is the accurate connection to terrestrial UTC to monitor and regulate short- and long-time performances. Deviations are to be measured with utmost accuracy. They have to be considered by routine data evaluation and transmission to allow the user to synchronize its own clock to such a space-clock (S/C).

The latest generation of the up to now most precise clocks in the mid-term range between 10^2 and 10^5 sec, H-maser atomic clocks, is at present under preparation for use in space. Development efforts in the U.S., in Russia, and also in Europe on behalf of ESA promise soon experimental satellite systems carrying such clocks [2].

To make those orbiting, ultra-stable time scales available at remote sites, either on ground or in space, radio frequency links in the daytime- and weather-independent microwave range have to be established. The microwave system locked coherently to the atomic oscillation, highly precise time transfer becomes possible.

2 Microwave Signal Transmission

The microwave signal transmitting the frequency and time information is subject to diverse disturbances on its way from clock to clock, what fig. 1 gives an impression for.

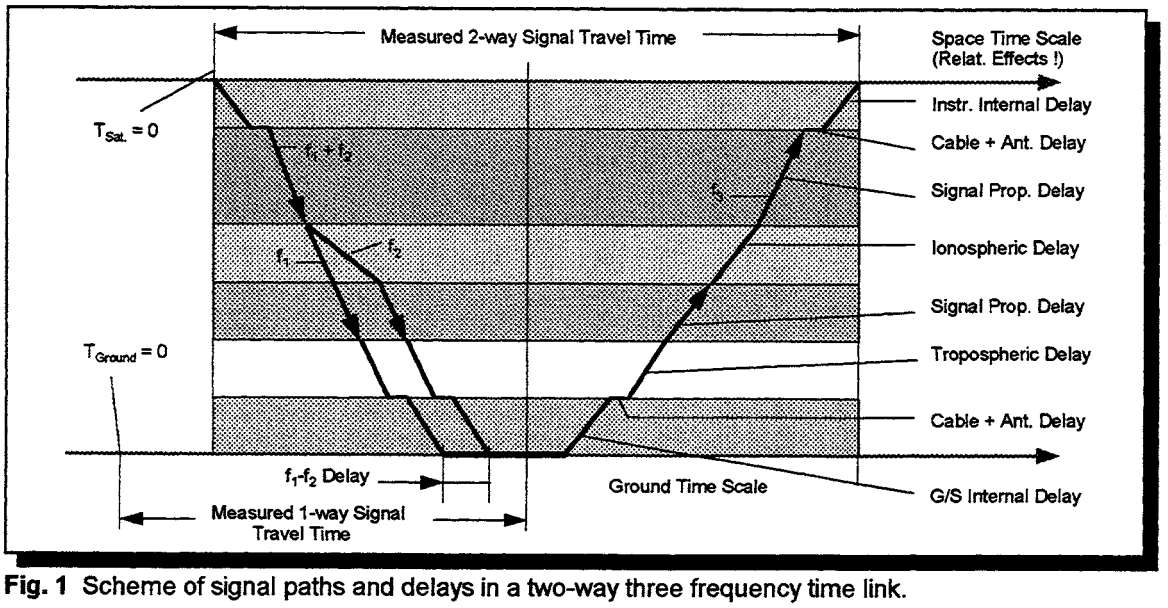


Fig. 1 Scheme of signal paths and delays in a two-way three frequency time link.

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. In addition, several modes of signal transmission can be distinguished which are different concerning their technical complexity, but, again, concerning the synchronization accuracy.

Two-way principles (fig. 2) allow to measure the desynchronization D of the involved clocks, if the asymmetry of the two paths is determined. This is fulfilled by measuring the signal travel time simultaneously at the receiving clock (one-way travel time Δt_{Rx}) and at the reference clock (round-trip travel time Δt_{Ref}). This method we call *time-ranging mode* with:

$$D = \Delta t_{Rx} - \frac{\Delta t_{Ref}}{2} \quad . \qquad (1)$$

At the Institute of Radio Frequency Technology, an alternative precision method, the *time comparison mode*, for desynchronization determination has been developed [1]: If not only the travel time, but also the time-of-arrival of the one-way signal at the receiving clock is noticed, and this information is then used at the reference clock for stopping the travel time measurement of the backward signal, the two-way measurement is changed to a twice one-way measurement which offers an again improved accuracy.

$$D = \frac{\Delta t_{Rx} + \Delta t'_{Ref}}{2} \quad . \qquad (2)$$

The practical realization of such a pretentious measurement configuration is presently under investigation; this method, however, represents the ultimate remote clock comparison technique.

At the same time, the round-trip measurement can provide the slant range r between the two receivers by the basic equation

$$r = \frac{c}{2} \cdot \Delta t_{Ref} \quad . \qquad (3)$$

Exchange of the measurement results between the clocks has to be done by data transfer on a separate data link. The main advantage of two-way measurements is the fact that the two results are achievable one without the other, i.e. the clock desynchronization D is obtained without knowledge of the clocks' distance, and the slant range r is obtainable without information about the clocks' deviation.

Synchronization of remote ground-clocks (G/C's) mutually is possible by "Common View" (CV) observation of the space-clock, or even by, as

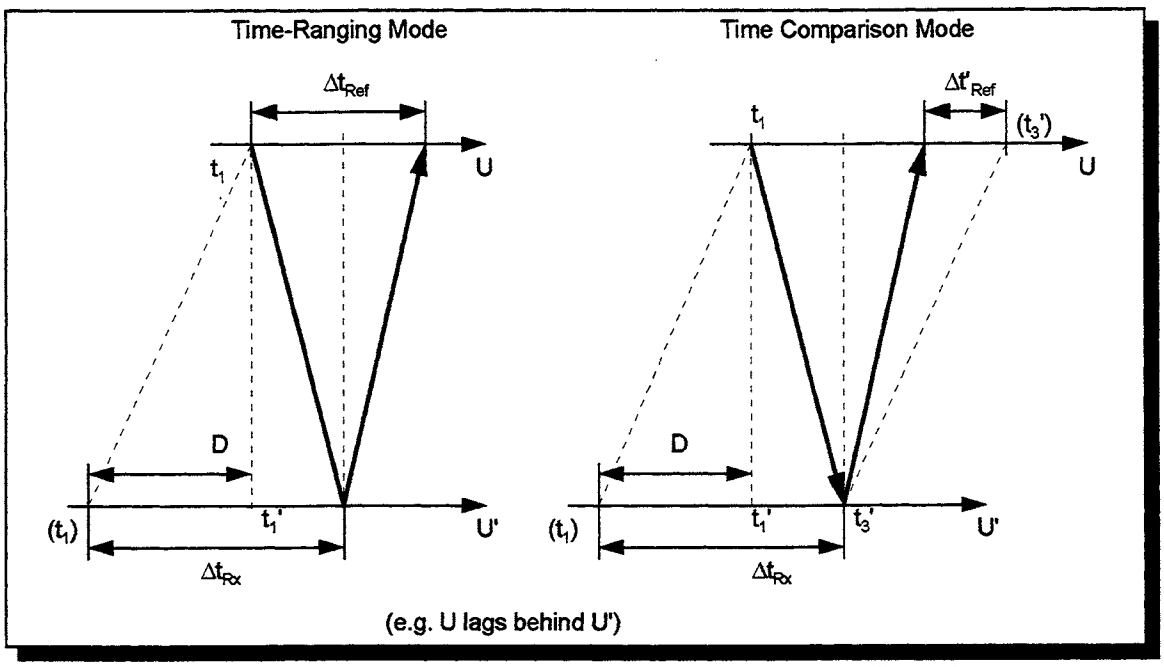


Fig. 2 Simplified picture of time-ranging and time comparison mode (symmetric case).

we call it, "Extended Common View" (E-CV) mode with separate observation times due to (a) the very high mid-term stability of a space-maser clock and (b) missing intentional degradation (cf. fig. 3).

If the microwave signal travels the ionosphere, a coherent signal on a second frequency with correlated reception gives the possibility to correct very precisely the ionospheric refraction. Tropospheric corrections have to be included by meteorological measurements and appropriate modelling; all other influences depend individually on the used system and the calibration efforts.

3 PRARE / PRARETIME System Description

A microwave system which fulfills all requirements for transmission of ultra-stable time information is the Precise Range and Range-Rate Equipment PRARE in its modified version PRARETIME. The PRARE instrument has been developed during the last decade in Germany for geodetic purposes. It is presently in use on polar-orbiting satellites: the Russian meteorological Meteor-3#7 spacecraft (launched in january 1994), and the European ERS-2 (to be launched in 1995) [8].

The instrument is a two-way (signal flow: space-ground-space), three-frequency (X-/S-band) system, working with pn-code correlation, spread spectrum data modulation, and carrier doppler measurement. These functions are derived coherently from one ultra-stable reference oscillator and assembled in a compact (21 kg, 400 x 250 x 180 mm^3) box which is, together with a small antenna, to be installed on the host satellite [7].

The main advantages of the PRARETIME system are - beside its two-way characteristics - its relatively high code and carrier frequencies which result in very low residual noise. Additionally, the high data transfer rates make it easy to transmit additional information in both directions. The data structure is organized within 1-sec frames whose construction shows fig. 4. As it can be recognized, the PRARETIME principle is not dependent on comparison of single 1-pps signals with their known high residual uncertainty, but makes consequently use of digital comparison of second-information which are spread onto the pn-codes.

The two main - nevertheless marginal - points of system modification to evolve from PRARE to PRARETIME are the following: the time transfer aspect in the present PRARE configuration being secondary behind the geodetic applications, only the master and monitoring ground-stations of the system, located in Oberpfaffenhofen, Potsdam, and Stuttgart, have the intention for the necessary one-way signal delay measurement facilities. Second, routine data evaluation and data exchange of the desynchronization results between space- and ground-segment is not implemented.

4 Calculated Time Transfer Accuracy

At the Institute of Radio Frequency Technology, simulations of the achievable clock synchronization accuracy with a modified PRARETIME instrument, connected to a space-

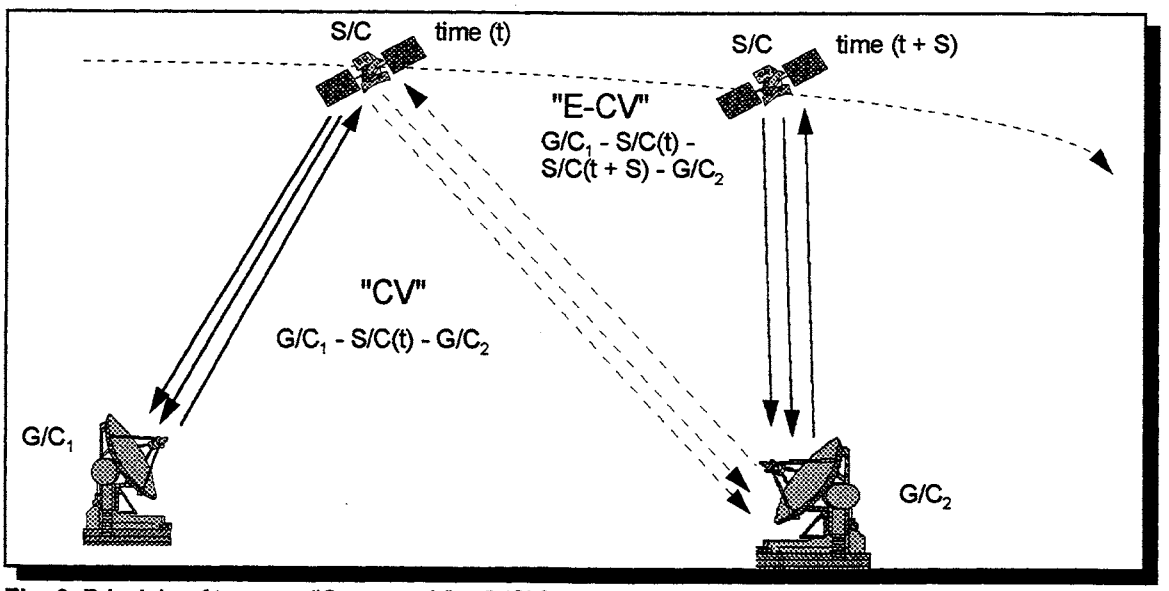


Fig. 3 Principle of two-way "Common View" (CV) and Extended CV (E-CV) techniques.

qualified H-maser clock and orbiting the Earth at around 1000 km, have been carried out. Individual error models and corresponding error budgets (1σ) for each of the two-way time transfer modes mentioned above have been derived, which include the parameters listed in tab. 1.

The quoted σ-values are developed in detail in [2]. It is important to mention that most of the numeric values are valid for observation in the neighborhood of culmination (= point of closest approach during satellite overflight), where degrading asymmetry influences are minimum. As an example we give the equation to evaluate

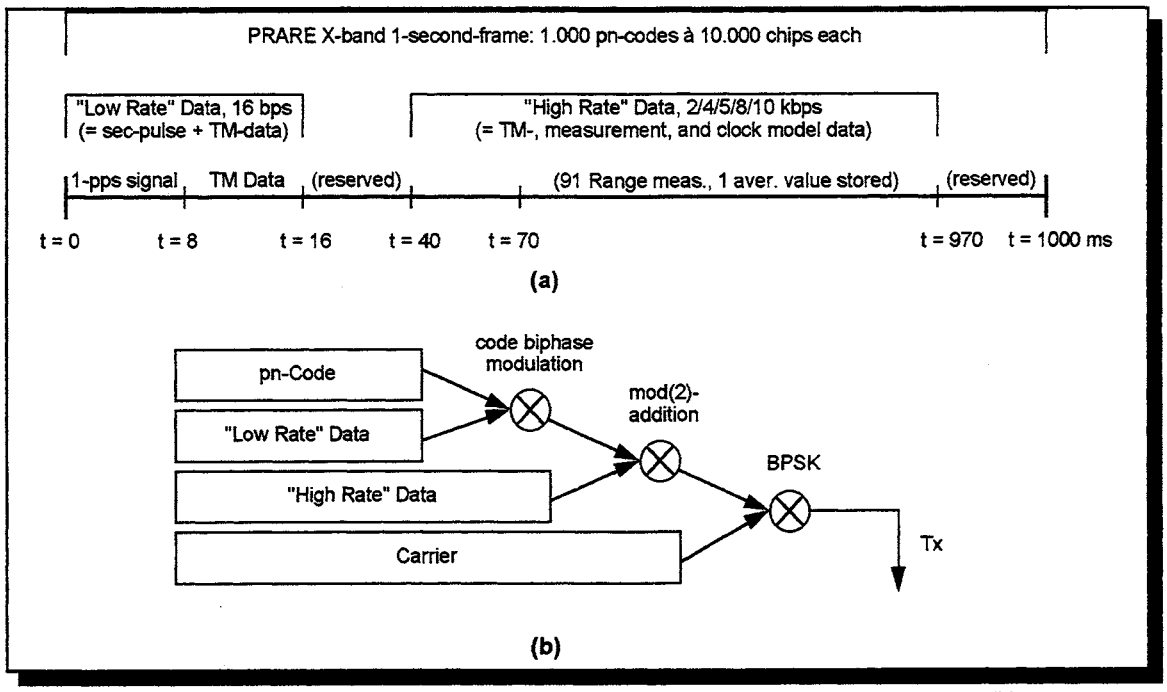


Fig. 4 PRARE system second-frame structure (a) and overall data modulation scheme (b).

Measurement principle	- PRARE method time interval measurement accuracy, $\sigma_{\Delta t}$ = 10 ps - PRARE hardware calibration accuracy, within specified operational limits σ_{cal} = 20... 50 ps - Tropospheric uncertainty (in one-way modes only), σ_{trop} = 50 ... 250 ps - Ionospheric uncertainty, determined by carrier frequency ratio, σ_{ion} = 50 ps
Onboard instrumentation	- H-maser oscillation error, specified e.g. in [3] with $\sigma_{s/c\ loc}$ = 0.15 ps for 1 s, $\sigma_{s/c\ loc}$ =15 ps for 10000 s - Frequency distribution system error, with specific design, σ_{freq} = 30 ... 100 ps - PRARE signal generation and correlation electronics noise, depending on pn-code length and signal amplitude [4], $\sigma_{S/C\ pps}$ = 30 ... 300 ps
Ground-station instrumentation	- Atomic clock oscillation error, specified e.g. in [5] with $\sigma_{G/C\ loc}$ = 5 ps (H-maser), $\sigma_{G/C\ loc}$ = 10 ps (Cs) for 1 s integration time - Time measurement system error, with high-quality equipment $\sigma_{G/C\ scale}$ = 100 ...300 ps - PRARE signal generation and correlation electronics noise, depending on pn-code length and signal amplitude [4], $\sigma_{G/C\ pps}$ = 60 ... 600 ps
Other	- Antenna phase center variation, depending on external conditions, σ_{ant} = 30 ... 70 ps - Antenna position uncertainty (in one-way modes only), σ_{pos} = 1 ns - Residual relativistic influences for a 1000 km polar orbit [6], σ_{rel} < 1 ps

Tab. 1 PRARETIME error contributions.

the total time transfer accuracy σ_T of the two-way time-ranging mode for S/C - G/C comparison:

$$\sigma_T^2 = \sigma_{\Delta t}^2 + \sigma_{cal}^2 + \frac{1}{4}(\sigma_{S/C\ loc}^2 + \sigma_{S/C\ freq}^2 + \sigma_{S/C\ pps}^2) + \sigma_{G/C\ loc}^2 + \sigma_{G/C\ scale}^2 + \sigma_{G/C\ pps}^2 + \sigma_{ant}^2 + \sigma_{rel}^2 \quad (4)$$

The ionospheric delay uncertainty σ_{ion} is contained in the method time interval measurement accuracy $\sigma_{\Delta t}$.

The results of the total time transfer accuracy simulations of the described PRARETIME two-way modes are presented in tab. 2. A detailed analysis of all time transfer modes can be found in [2].

Moreover, we estimated the time synchronization accuracy of different possible PRARETIME one-way modes, and compared it to the accuracies of corresponding GPS modes. Tab. 3 presents the accuracy of simulated time dissemination with PRARETIME + H-maser and GPS respectively.

In tab. 4, "Common Views" with PRARETIME + H-maser onboard a polar-orbiting satellite and with GPS have been opposed for various baselines.

5 Operational Aspects and Potential Applications

As introduced, the space-clock must be linked to a reference time scale such as UTC, or to an individual time lab (k) scale UTC(k) with dedicated conjunction to UTC. At the Institute of Radio Frequency Technology, scenarios for an adequate ground segment layout have been developed which allow to permanently control

two-way time-ranging			two-way time comparison		
G/C - S/C	CV	E-CV	G/C - S/C	CV	E-CV
130 ps	170 ps	190 ps	90 ps	90 ps	120 ps

Tab. 2 Accuracy of PRARETIME+H-Maser two-way time transfer modes (averaging over 10 s).

system	one-way S/C - G/C
PRARETIME + H-Maser (ephemeris error = 1 m, G/C coordinates = 0.3 m, averaging over 10 s)	1 ns
GPS, cf. [9] (broadcasted ephemeris, measured ionosphere, averaging over 780 s, SA-off)	16 ns
GPS, cf. [9] (broadcasted ephemeris, measured ionosphere, averaging over 780 s, SA-on)	49 ns

Tab. 3 Accuracy-comparison between GPS and PRARETIME + H-Maser for S/C - G/C time dissemination.

system	one-way CV baseline < 1000 km	one-way (E-) CV baseline > 5000 km
PRARETIME + H-Maser (ephemeris error = 1 m, G/C coordinates = 0.3 m, averaging over 10 s, 1000 km polar orbit)	1 ns	1.7 ns
GPS, cf. [9] (measured ionosphere, averaging over 780 s)	3.6 ns	8 ns

Tab. 4 Accuracy-comparison between GPS and PRARETIME + H-Maser one-way CV modes.

the onboard clock with the above calculated accuracies and to make the space time scale available to active (= two-way comparison equipment) or passive users (= one-way receiver) [2].

The one-way user must be provided with additional data, e.g. orbit elements. This data as well as the clock model parameters, which are to be regularly updated by the master ground station, can be appropriately coded ("navigation message") and easily integrated in the PRARETIME system data structure.

Comparable to existing GPS time receivers, the user has then immediate access to UTC(k) - but with an accuracy at least one order of magnitude better [2].

To evaluate the described "Common View" ground-clock synchronization scenario, we plan to investigate the possibility to perform validation tests with the help of the existing PRARE instrument on Meteor-3#7 or ERS-2 and with two of the PRARE system master stations (G/S's), which are provided with atomic clocks (fig. 5). It is expected that the achievable synchronization accuracy σ_T is around 2.5 ns, due to the sub-optimum design of the existing PRARE system for time transfer purposes. Nevertheless, these tests should allow to prove the applicability and the improved synchronization accuracy as compared to GPS of this new two-way time transfer system.

Potential applications for a space-clock + PRARETIME system have been identified in the fields of time metrology (connection of clocks contributing to TAI; faster availability of UTC; time scale dissemination), navigation (geometric orbit determination due to parallel receiving capability and full maintenance of PRARE ranging and range-rate features; installation on a GLONASS satellite and correlation of signals), fundamental physics experiments (geoscience; general relativity; "VLBI in space") [2, 10].

6 Conclusion

The proposed PRARETIME equipment is adequate to monitor time and frequency of an ultra-stable H-maser S/C. Time interval measurements around culmination point of the satellite's overflight provide most precise results. Simulations result in an overall synchronization accuracy better than other time transfer modes presently in use, the GPS-CV and MITREX time transfer. To perform ultra-precise time transfer, the one-way signal delay measurement on the G/C is necessary by means of high-quality time interval counters and careful hardware calibration prior to lunch and inflight, and an

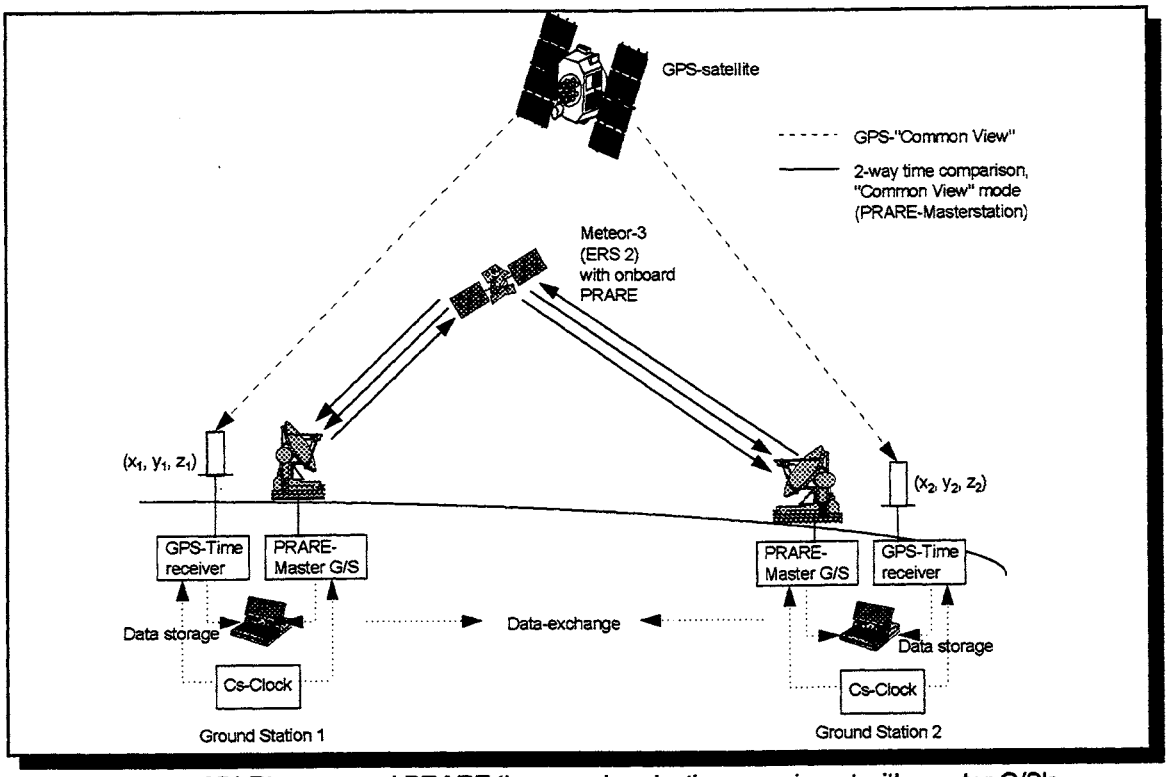


Fig. 5 Scheme of DLR's proposed PRARE time synchronization experiment with master G/S's.

adequate data evaluation software. PRARETIME offers the possibility of optional one-way time dissemination by means of time receivers designed very similiar to GPS time receivers. The excellent performance of PRARETIME for time transfer and navigational measurements emphasizes that this equipment is a powerful candidat as a basic element in a future GNSS II.

References

[1] Sappl, E.
Comparison of a Ground and a Satellite clock by Three-way Time Transfer
to be published, 1995

[2] J. Hahn, S. Bedrich, H. Nau
Study on "H-Maser in Space" as Ultra-Stable Clock for Time Transfer and Time Dissemination, Final Report
DLR Oberpfaffenhofen, January 1995

[3] Bernier, L.-G., G. Busca, P. Rochat, H. Schweda
Hydrogen Maser Frequency Standard
Doc. ON/SHM/PSR/issue 3,
Observatoire de Neuchatel, March 1994

[4] Spilker, Jr., J. J.
Digital Communications by Satellite
Englewood Cliffs: Prentice-Hall, 1977

[5] Lewis, L. L.
An Introduction to Frequency Standards
Proc. IEEE 79, "Special Issue on Time and Frequency", July 1991, 927-935

[6] Schwarze, V. S., E. W. Grafarend
Relativity and Geodesy in the Meteor 3M Satellite Mission
personal communication, 1994

[7] Hartl, Ph.
The Precise Range and Range-rate Equipment (PRARE)
Proc. Workshop on ERS-1 Radar Altimeter Data Products, Frascati, 8-11 May, 1984 (ESA SP-221), 153-160

[8] Hartl, Ph., W. Schäfer, M. Conrad, Ch. Reigber, F. Flechtner
PRARE: First Results of the METEOR-3 Mission (in German)
SPN 2/94, 1994, 51-54

[9] Uhrich, P., C. Thomas
Hydrogen-Maser in Space. Contribution about Global Clock Synchronization
personal communication, 1994

[10] GFZ Potsdam, University of Stuttgart, DLR Oberpfaffenhofen
PRARETIME, Precise Time Transfer and Tracking for Earth Parameter and Orbit Determination
Proposal to ESA in response to AO No. JP/11/9460/PS/nr,1994

CUSTOM, TURNKEY TIME AND FREQUENCY SYSTEMS. A STRUCTURED, EXPANDABLE APPROACH.

DAVID F. WRIGHT

RADIOCODE CLOCKS LTD

ABSTRACT

Radiocode Clocks Ltd have developed a Turnkey Time and Frequency Generation and Distribution System strategy based upon a bus of three, "core" signals from which any Timecode, Pulse rate or Frequency can be produced. The heart of the system is a ruggedised 19 inch, 3U Single Eurocard chassis constructed from machined 10mm aluminium alloy plate and designed to meet stringent Military, Security and Telecommunications specifications. The chassis is fitted with an advanced multilayer backplane with separate ground planes for analogue and digital signals ensuring no degradation of low noise frequency references in the proximity of high speed digital pulse transmissions.

The system has been designed to be used in three possible configurations:-

a) As a stand alone generation and distribution instrument.

b) As a primary distribution unit in a turnkey Time and Frequency system.

c) As a secondary distribution unit at a remote location from the Turnkey Time and Frequency System providing re-generation of core signals and correction for transmission delays.

When configured as a secondary distribution unit the system will continue to provide usable outputs when one, two or even all three of the "core" signals are lost.

The instrument's placement within a system as a possible single point of system failure has required the development of very high reliability translator, synthesizer, phase locked loop and distribution modules together with a comprehensive alarm and monitoring strategy.

INTRODUCTION

The requirements for sources of Precise Time and Frequency have grown substantially in recent years, not just in the number of projects but also in their complexity. An increasing number of users are specifying redundancy, high reliability, expandability and a plethora of Timecode, Pulse rate and Reference frequency outputs in their systems. In dual or triple redundant configurations the switching, monitoring and alarm management of such a myriad of signals can become technically complex, commercially expensive and, in many cases, can reduce system availability. This, coupled with the tightening of international standards with regard to electromagnetic compatibility, electrical safety and product quality, has lead to the development of a new strategy for Time and Frequency distribution.

SYSTEM ARCHITECTURE

Traditionally, when designing a turnkey distribution system, two or three master clocks, frequency standards or timecode generators are fitted with one of each required system output. In some complex cases an atomic frequency standard will provide a stable reference frequency to a pair of low noise quartz oscillators via individual frequency or phase locked loops. These oscillators then produce one of each of the required frequency outputs which are fault detected and switched to a frequency distribution unit, while a GPS or Off-air master clock synchronises two or three timecode generators, each of which produces one of every required time or time interval output. These in turn, are majority voted, fault detected and switched to a timing buffer or distribution unit.

The addition of features such as secondary power supplies, alarms and output monitoring quickly make the system difficult to use, very inflexible, costly to manufacture and support and inefficient in terms of spares ranging.

Our solution to these problem has been to reduce the number of switched signals to an absolute minimum and to use these "core" signals as references with which to generate the required system outputs. Using conventional copper or fibre optic technology the minimum "core" signal count is three, a 1 Pulse per second (1 PPS) epoch marker, a 10MHz reference frequency and a proprietary format RS422 time message.

These "core" signals are generated by one or more Time and Frequency Standards such as GPS or Off-Air receivers, free running atomic or quartz clocks or a combination of the above. The "core" signals, plus alarm and status signals from the master time and frequency source are fed to an intelligent monitor and changeover unit which will fault detect, majority vote and switch between sources. This unit also acts as a system alarm manager. The selected "core" signals are then fed to a high reliability, modular distribution chassis based upon a single Eurocard format and fitted with dual power supplies, an alarm management module, an input/expansion module and a backplane accommodating the "core" signals, power lines, alarm signals and inter-module control signals.

Because there are also smaller applications, where a single distribution chassis is sufficient and the use of an external source of time and frequency is not always necessary, a range of "core" signal generation modules, which could obtain time from an external reference, has been developed for the distribution system. These modules now include GPS, Loran,

WWVB, MSF and DCF receivers, together with free-running master time and frequency sources.

Historically, the distribution of low noise frequencies and digital signals has been undertaken using two separate chassis. As this approach was considered cumbersome, a single chassis architecture was developed for the distribution of all time and frequency signals while maintaining the quality of the outputs. This has been achieved by developing an advanced multilayer backplane with high isolation and separate ground planes for analogue and digital signals allowing high speed pulse trains and low-noise frequencies to be processed in close proximity without any significant degradation in signal quality.

High reliability modules have also been developed which accept the required signals from the backplane and translate, synthesize and distribute virtually any Time and Frequency signal. Most modules are 4E wide and provide five isolated outputs of each signal. The Input module has been developed to provide expansion outputs of the "core" signals so that when the chassis is fully populated, additional modules can be added by simple connection of further chassis allowing virtually infinite expansion capabilities without having to modify or reconfigure all the instruments in the system.

Another important design consideration was the provision of high stability, very low noise reference frequencies within the distribution system. Good long and medium term stability can be achieved by using a Cesium standard or GPS-disciplined Rubidium atomic oscillator but typical specification requirements of $1x10^{-12}$ over 24 hours, $1x10^{-11}$ over 1 second, and phase noise below 110dBc/Hz at 1Hz offset from carrier could only be achieved by the use of a high quality ovened quartz oscillator phase locked to the "core" 10MHz signal.

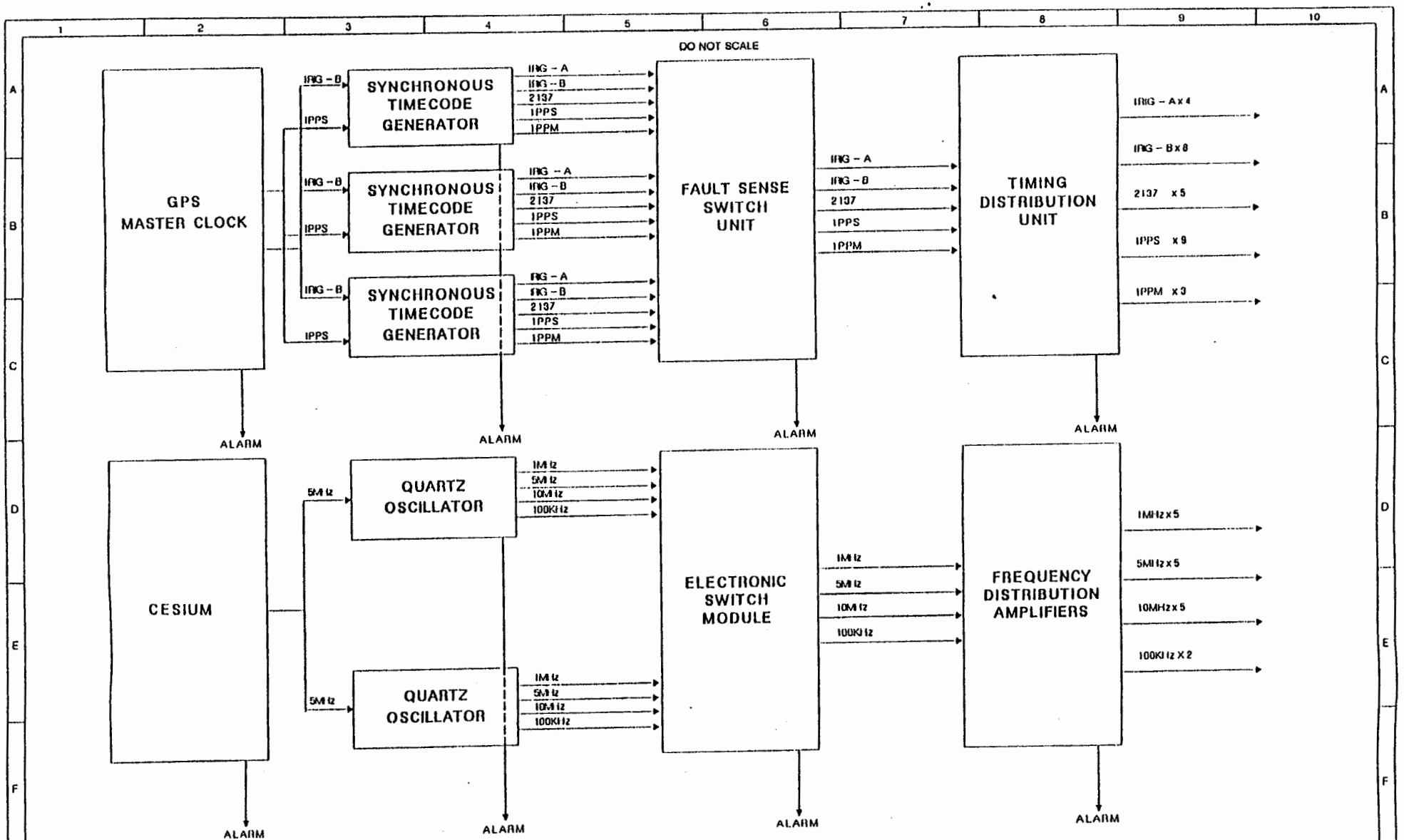
Our standard PLL module uses a number of novel techniques to implement the well proven second order phase locked loop that has the required characteristics. In the short term, up to some tens of seconds, the phase locked loop's ovened oscillator frequency can be more stable than the "core" signal; for this reason the phase locked loop is designed to have a time constant of about one minute. For longer time periods the loop holds the oscillator in phase with the 10MHz reference. Time constants of that order can be realised using analogue techniques but when the requirement for "holdover" operation on loss of reference is considered the use of digital control becomes mandatory. A microprocessor controlled loop provides sophisticated solutions for all of the control problems but has inferior reliability performance. Worse still, the processor generates a broad spectrum of noise signals which are unwelcome in a module whose primary purpose is to provide high quality, low noise analogue signal outputs.

The solution we have developed for this system is part analogue, part digital proportional and integral control. The control loop error signal is formed by direct phase comparison of the 10MHz "core" signal and the local ovened oscillator in a circuit which is effective over more than 320 degrees of phase, almost a whole cycle at 10MHz. The magnitude of the phase difference provides the control signal for a low frequency voltage controlled oscillator which clocks a 16 bit up/down counter. The sign of the phase difference determines the count direction. The counter outputs connect to a 16 bit DAC which provides the integral part of the loop control signal. When the phase difference has been driven to zero in a steady state condition the up/down counter ceases to be clocked and no other clock signals are active in the control system.

The integrator can set the ovened oscillator control voltage anywhere in it's entire range but the largest proportional control contribution required has only a fraction of the value of the control range. This signal is a fraction of the loop error signal already derived from the phase detector; it is fed through an analogue switch and summed with the signal from the DAC to form the ovened oscillator control voltage.

The proportional control is immediately active at a low level in response to a loop error; the integral control works to drive the phase to zero in the longer term. If the "core" 10MHz reference is lost then the integrator has the correct value to hold the oscillator at the last best control value indefinitely. The digital parts of the circuit fit readily into programmable logic and the controller adds no clock noise to the OCXO output.

Development of the Low Noise Phase Locked Loop Oscillator Module has provided us with a new time and frequency distribution sub-system, the TFD8000. This is now considered not only a product but a new strategy providing a structured, expandable approach to turnkey system design which has already proved cost-effective, reliable and easy to maintain in applications within the Defence, Security and Telecommunications Industries.



ISSUE	DESCRIPTION	APPD	DATE
3			
2			
1	FIRST ISSUE		

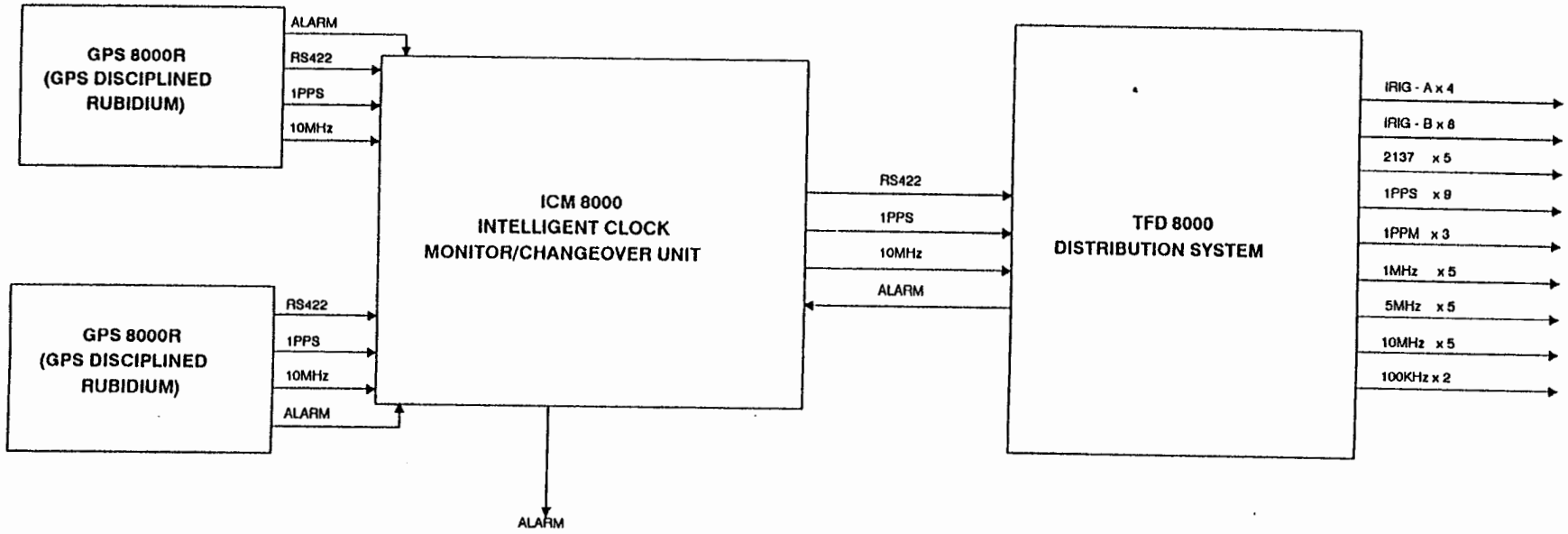
RADIOCODE CLOCKS
© 1995
SHEET 1 OF 1

TITLE
Traditional Time & Frequency System
Block Diagram

DRAWN	TRACED	CHECKED	APPROVED	APP DATE
K Rollason				10/2/95

DRAWING No RCC 3567 BD

DO NOT SCALE



ISSUE	DESCRIPTION	APPD	DATE
3			
2			
1	FIRST ISSUE		

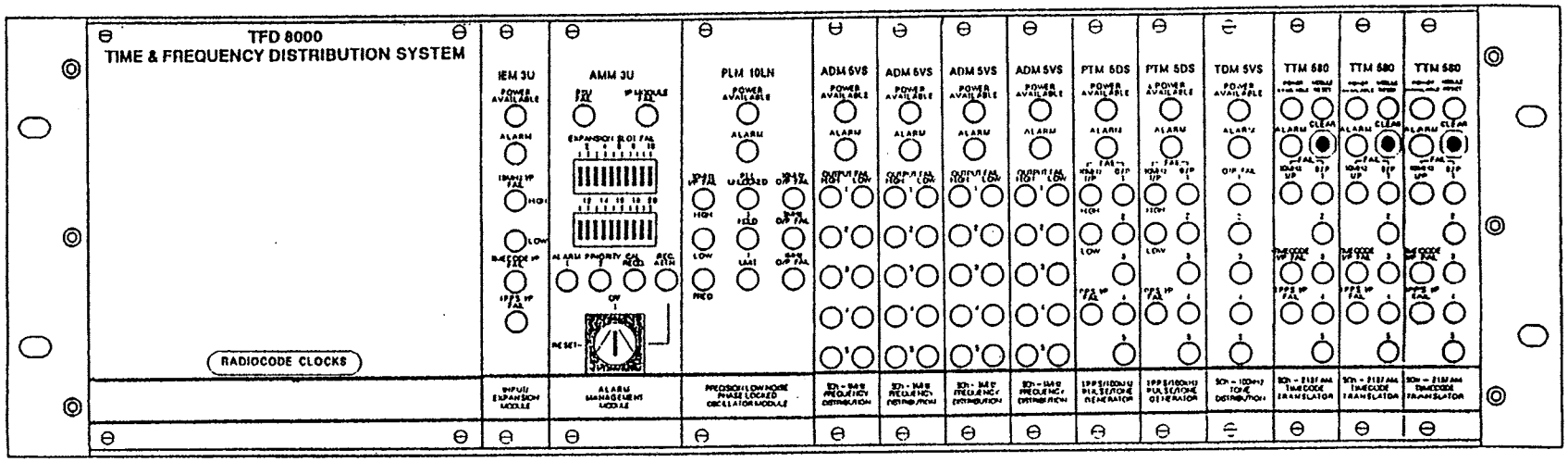
RADIOCODE CLOCKS
© 1995
SHEET 1 OF 1

TITLE
Modern Time & Frequency System
Block Diagram

DRAWN	TRACED	CHECKED	APPROVED	APP DATE
K Rollason				10/2/95

DRAWING No RCC 3567 BD

DO NOT SCALE



ISSUE	DESCRIPTION	APPD	DATE
3			
2			
1	FIRST ISSUE		

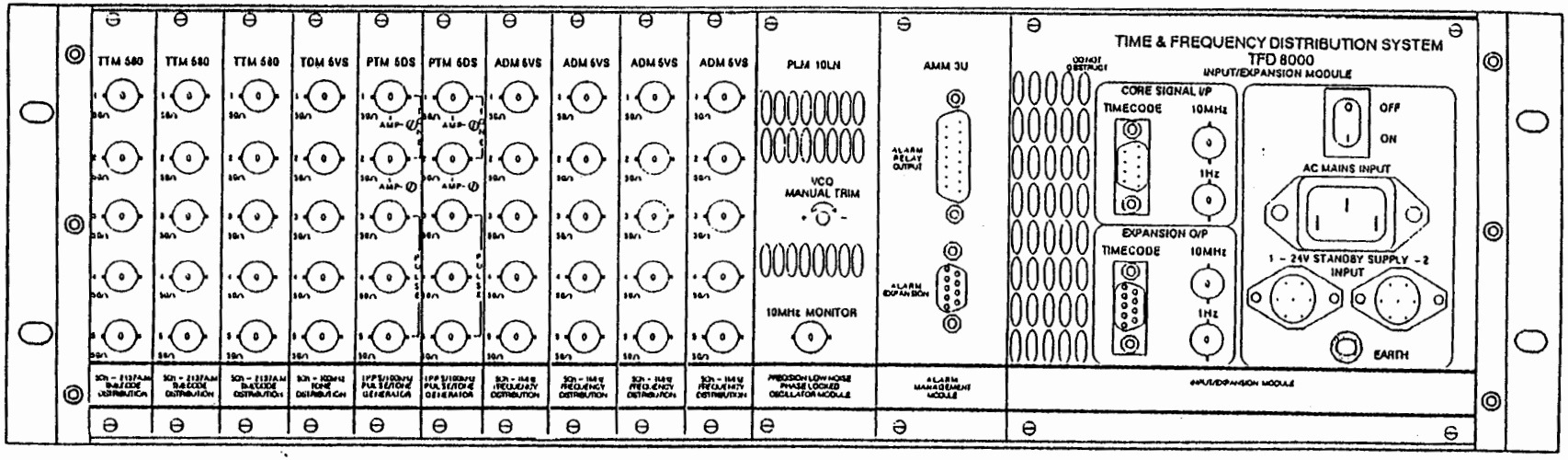
RADIOCODE CLOCKS
© 1995
SHEET 1 OF 1

TITLE
Typical System Configuration
Front Panel Diagram

DRAWN	TRACED	CHECKED	APPROVED	APP DATE
K Rollason				10/12/95

DRAWING No RCC 3610 FD

DO NOT SCALE



ISSUE	DESCRIPTION	APPD	DATE
3			
2			
1	FIRST ISSUE		

RADIOCODE CLOCKS
© 1995
SHEET 1 OF 1

TITLE
Typical System Configuration
Rear Panel Diagram

DRAWN	TRACED	CHECKED	APPROVED	APP DATE
K Rollason				10/2/95

DRAWING No RCC 3611 RD

9TH EUROPEAN FREQUENCY AND TIME FORUM BESANCON 8-10 MARCH 1995

STABILITY OF A CS-CLOCK BASED 2048 KHZ OPTICAL ISDN SYNCHRONIZATION SIGNAL

K. Kalliomäki(1), T.Mansten(2)

1) Univ. of Oulu, Dept. of Electrical Eng., FIN-90570 Oulu, Finland
2) VTT AUTOMATION, Measurement Techn., P.O.Box 1304, FIN-02044, Finland

Abstract

Finnish Telecom has a set of Cs-atomic clocks to control wideband digital telecommunicati on signals. To create an independent atomic time for Finland several methods to transfer the time information from Cs-clocks to National Calibration Laboratory has been evaluated. We have tested the stability of Cs-based 2048 kHz synchronization signal available from VTT telephone exchange, GPS clocks as references. The average frequency difference was 3*10exp-13 during 8 months observation period. An optimal smoothing time seem to 4 days. Due to phase slips the synchronization signal is suitable only to frequency reference.

Introduction

The official time and frequency in Finland has been based on Loran-C and since 1989 on GPS-receivers. The rubidium clocks have been used as "flywheels" to reduce short term noise. Finland has no independent atomic time scale due to lack of a coorninated control of existing Cs-clocks, which are used to control the frequency of wideband telecommunication signals (PDH, SDH). Unfortunately the location of the clocks is nearly 20 km from VTT AUTOMATION, a national calibration laboratory (KML) for most electrical quantities and for time and frequency. Different time transfer methods has been evaluated and tested to obtain an independent atomic-time and -frequency. On the other hand the control of Cs-clock based 2048 kHz signal is important for telecommunication people. Therefore as a first working comparison method a measurement system for 2048 kHz synchronization signal is built, tested and has been in continouous operation since June 1994.

Instrumentation

To synchronize wideband digital telecommunication networks a Telecom Finland has a set of three Cs-clocks (Oscilloquarz) and associated electronics to produce a reliable 2048 kHz synchronization signal. This signal is multiplied to 140 MHz and distributed into capital area via optical connections (Ericssons MD110 telephone exchange) and changed in frequency (eg. 34 MHz and 2048 kHz). After nearly 20 km propagation the reconstucted 2048 kHz signal is available at VTT digital telephone exchange (Fig 1).

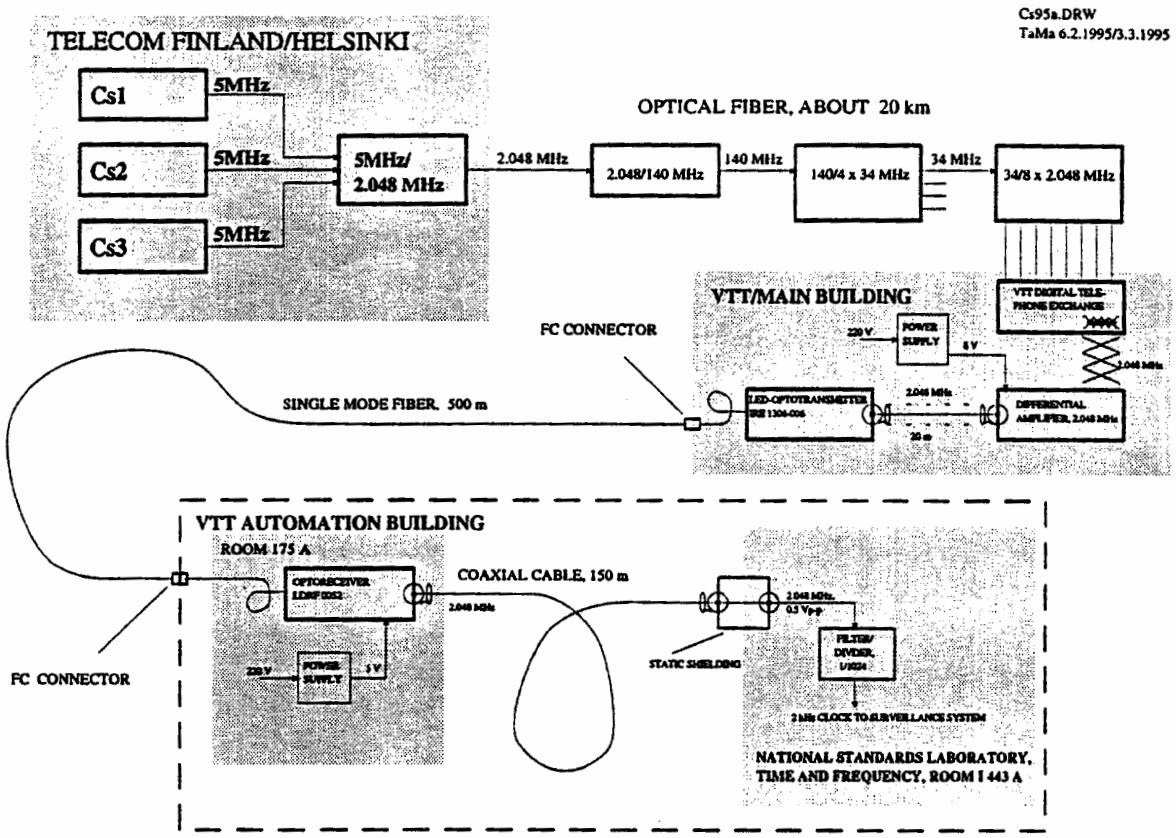


Fig. 1: Path of the synchronization signal

This signal is connected to the National Time & Frequency Laboratory (KML) via an optical cable. The resulting signal is filtered using single resonant circuit (Q value about 10), converted into square wave and divided down to 2 kHz. The phase of this 2 kHz signal is monitored every 10 min. using computer controlled surveillance system of KML. As a short term reference rubidium clock (EFRATOM FRS) is used. Long term (>1 day) stability is referred to GPS (Navstar XR5M) receiver.

Spectrum of 2048 kHz signal

The reconstucted 2048 kHz synchronization signal is not a clean square wave. It contains phase jitter of about 30 nsrms and phase slips. Spectrum analyzer reveals dominating 8 kHz (=2048/256) phase modulation (spectrum peaks about -30 dB down from carrier) and weaker 128 kHz (=2048/16) phase modulation (-50 dB down). In addition weak low frequency (300/600 Hz) modulation is visible, too.

Phase slips

Phase slips occurred randomly. The typical rate was 11 phase slips per month during 5 months period. All phase slips were clearly multiples on synchronization period (488 ns). About half of the slips were one cycle slips but phase jumps up to 200 cycles occurred.

Data handling

The 2 kHz signal (2048 kHz divided by 1024) is referred to GPS-time (Navstar XR5M 6-channel receiver) every 10 min. Phase slips were corrected by adding or substracting multiples of synchronizaton period to obtain a continuous phase record.

Then the data was smoothed (2 h averaging time) to reduce both GPS (typically 30 nsrms)- and synchronization signal noise. After differentiation a raw frequency difference between GPS and Cs-based synchronization signal was obtained. In practice a further smoothing up to several days was needed to reveal possible changes in Cs-based signal.

Results and discussion

The average frequency difference between GPS and synchronization signal was 0.30*10exp-12 and it did not show any statistically significant trend during 230 days follow up period. Specification (CISPR) for frequency error is less than 10*10exp-12, thus the frequency seem to be very well within spesification. Fig. 2 shows the measured frequency difference during 8 months follow up time.

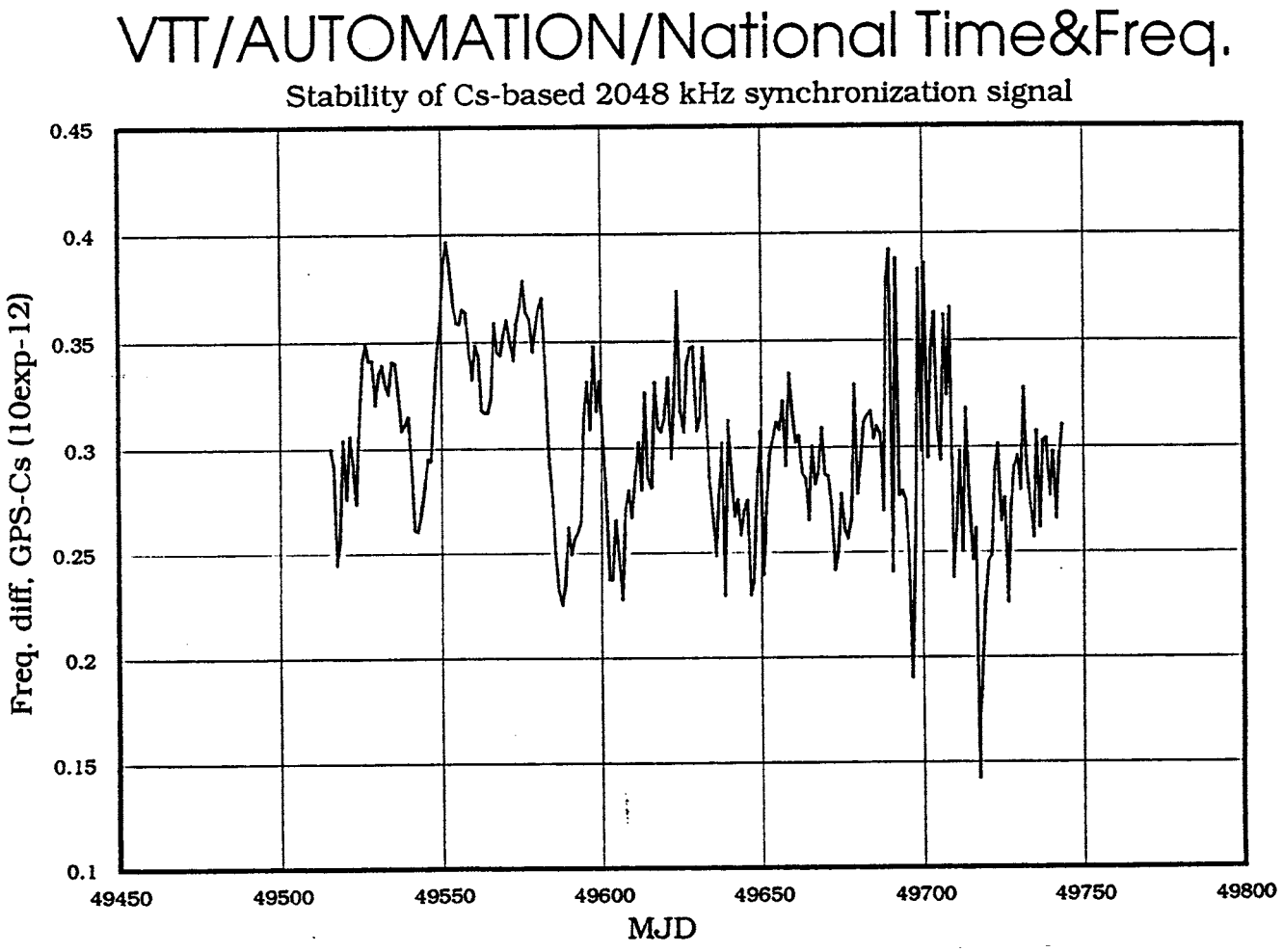


Fig. 2 The measured frequency difference GPS-Synchronization signal

The following table gives frequency difference as a standard deviation of the function of smoothing time.

TABLE 1
Standard Deviation of frequency difference (GPS-Synch.) as a function of smoothing time.

Smoothing time (days)	SD of freq. diff. (10exp-12)
10 min	74
1	0.28
2	0.13
4	0.07
8	0.043
16	0.027

The scatter (SD) of the measurements is caused by short term variation of GPS reference tick (about 30 ns), jitter of the measured synchronization signal (about 30 ns) and possible slow drift of GPS reference or of the synchr. signal. Variation was not constant, it was smallest in July 94 and highest in December 94.
Because the SD is inversely proportional to the smoothing time up to 4 days, short term GPS variation and jitter of synchr. signal dominate below 4 days smoothing time.
Therefore at least four days smoothing time is needed to see possible drift of the monitored Cs-clocks. Fig 1 shows the frequency difference between GPS and Cs-clock based synchronizaton signal, smoothing time is 8 days. Frequency changes are near the flicker floor of a typical Cs-clock and probably reflect the stability of synchronization signal.

A PROTECTED GPS TIME / FREQUENCY REFERENCE SOURCE

Y. BOURRAGUE, V. PELLETIER, E. BENOIST, V. MOURIER

TEKELEC TECHNOLOGIES
29, avenue de la Baltique - 91953 Les ULIS - France

ABSTRACT

This paper describes TEKELEC TECHNOLOGIES Series TE 670, a protected GPS Time / Frequency Reference Source that combines the advantages of a frequency standard, in short term applications, and the advantages of a GPS system, over the long term. The TE 670 was designed to meet the French Air Force's demand for an efficient, economical system to synchronize its bases. The TE 670 is fully cost effective: compared to a cesium standard, it offers longer life and higher reliability, at a lower price.

We present the goals and principles of the Digital Slaving Loops (DSL) between the TE 670 internal oscillator and the 1 pps output signal of the P/(Y) code GPS receiver.

We evaluate an algorithm based on Kalman filtering, and present its ability to combine the most desirable characteristics of the participating frequency / time standard: the short-term stability of an OCXO BVA and the long-term stability of the GPS system.

We also present the test station assembled to prove the system, and the results that we obtained from our test sequence.

1 - INTRODUCTION

Precise control, calibration and measurement of time and frequency are widespread, growing needs in all of the technologies that drive our world. Applications such as radionavigation, radiocommunication and satellite operations continuously demand higher-quality information in these areas, to improve their performances.

The cesium-beam "atomic clock" is today's commonly accepted primary standard for time / frequency operations. This system is based on the precise, stable behavior of the cesium atom. However, cesium-beam technology suffers from high costs, which limits its marketability; and from the complexity of its construction needs, which limits its reliability in rugged environments.

A primary-standard reference signal can also be acquired via the satellites of the worldwide GPS radionavigation network.

GPS time is available to any radio station equipped with a dedicated receiver; but these signals include SA scrambling, which masks their maximum capabilities.

The TE 670 eliminates the shortcomings of referencing by both the primary-standard and GPS methods. The TE 670 can process GPS signals to dilute the SA component, and restore GPS time with a precision of ± 100 ns. The TE 670 provides a frequency stability of 2×10^{-12} within 24 hours. The use of P/(Y) coding in the receiver ensures maximum availability, regardless of circumstances.

2 - FUNCTIONAL DESCRIPTION

The principal functions of the TE 670 are:

- Generate and maintain a reference frequency of 5 MHz.
- Generate and maintain a reference time scale synchronized to UTC. This time scale appears as a 1 pps signal, which is synchronized to UTC, and associated with an external time message.
- Provide dating for an external 1 pps signal to the time scale.

Figure 1 presents the TE 670 functional block diagram.

The TE 670 groups 5 major units:

- The GPS unit acquires the satellite signal; recovers the ephemeris and almanac information components, and the UTC reference; generates a 1 pps time signal, and feeds it to the T/F unit.
- An oscillator, type OCXO BVA, under CPU command. OCXO BVA generates a pilot with very low phase noise, and feeds it to the T/F unit.
- The Time / Frequency (T/F) unit, which is clocked from the GPS unit and piloted from OCXO BVA. This unit includes: the synchronizing unit; 2 digital slaving loops (DSL); signal maintenance and conditioning circuits; and the output interfaces for the signals to be distributed, i.e. 5 MHz, 1 pps, STANAG 4430, and an external dating signal.

- The Central Processing Unit (CPU), which controls all of the major units and centralizes their status reports. For the user located away from the TE 670 site, a remote monitor link provides the needed time and frequency information.
- The Man / Machine Interface (MMI) is the local control facility.

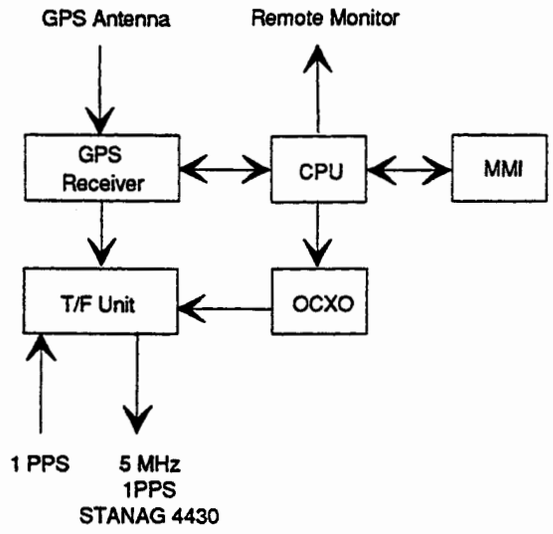


Figure 1.

The time and frequency reference outputs are piloted by the DSL circuits. The timing DSL maintains the 1 pps signal within ± 100 ns, relative to the UTC scale. The frequency DSL ensures that the stability of the 5 MHz output attains better than $2x10^{-12}$ within 24 hours.

A single microprocessor drives both DSL circuits. Each DSL includes peripherals, such as counters, phase detectors, and D/A converters. The microprocessor executes the synchronizing algorithms, and suppresses the effects of correlation between the slaving loops.

3 - TIMING BY DSL

The Timing DSL greatly improves the accuracy of the GPS time scale acquired by the receiver, by diluting its scrambled SA component.

As shown in Figure 2, the GPS unit provides a 1 pps signal that is synchronized to UTC (USNO), which is continuously compared to the 1 pps signal produced by the TE 670 internal oscillator, and measurement yields a resolution of 1 ns. A first filter stage rejects measurements that are obviously incorrect, which result from abnormally great differences between the 1 pps signals.

The second filter, a Kalman type [1], suppresses the effects of SA scrambling and the jitter component in the GPS signal. These effects are estimated from previous phase measurements, and reevaluated with reference to new measurements as they are made.

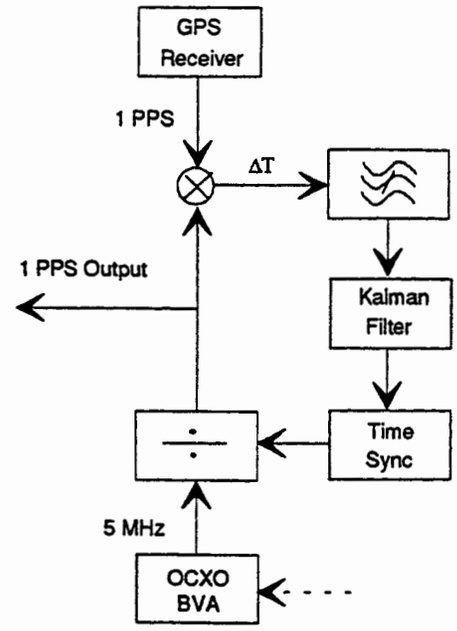


Figure 2.

The microprocessor compares the phase differences against a threshold value. If the phase difference exceeds the programmed tolerance, the Timing DSL corrects the 1 pps signal from the TE 670.

4 - FREQUENCY SLAVING LOOP

The accuracy of the TE 670 output is largely dependent on how accurately and smoothly the 5 MHz pilot can be commanded. As will be explained, Newton's method, especially in the DSL, will be used to meet and correct deviations from the programmed frequency level.

One commonly applied method is to compute and store F_t, the pilot's Time-Frequency transfer function. When F_t is known, the application of frequency corrections, $\Delta f/f$, is given by the equation: $\Delta f/f = F_t(v)$, where (v) is the pilot's command voltage. The resulting curve resembles a straight line with non-linearities on the order of 1/1000.

If the slope of the characteristic curve is resettable non-linear modifications may appear. These will variy directly with functions such as ambient temperature. These are principally linked to the behavior of the electronic units in the Frequency DSL. The system continuously updates the mathematical model of the transfer function, to compensate for these errors without affecting long term stability.

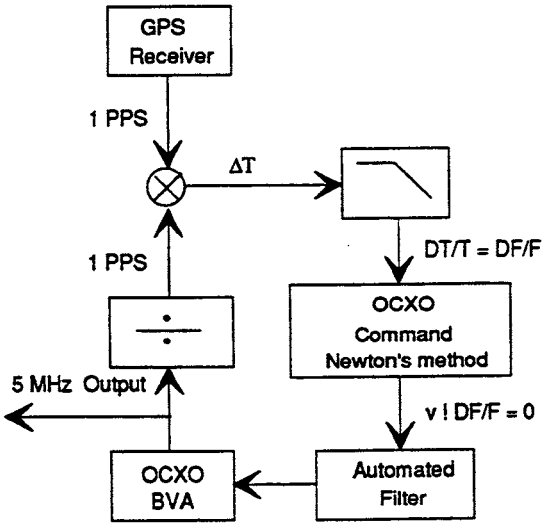


Figure 3.

Figure 3 presents the principle of frequency synchronization. The input circuit of DSL monitors ΔT, the phase difference between the pilot-driven 1 pps signal and the 1 pps signal received from GPS.

A low-pass filter suppresses any rapid fluctuations in the GPS signal. The time constant applied to this function is approximately 1 hour. At the filter's output, the TE 670 obtains the value required to correct the pilot frequency, $\Delta F/F = \Delta T/T$. The OCXO command circuit determines the command voltage, which will be applied to the pilot.

In brief, the method used is as follows [2]:

Take ξ as the root of the equation $f(x) = 0$, where $f(x)$ is the voltage / frequency transfer function of the pilot; $f'(x)$ and $f(x)$ are continuous, and maintain a constant sign throughout the application segment. This characteristic is valid for most oscillators in which voltage is commanded by Varicap diodes. After computing a first approximation of root $x \approx \xi$ we can improve its accuracy by introducing h, a correction factor; in such a case, $\xi = x + h$, where $h = -f(x) / f'(x)$. Finally, the successive approximations of the root are given by the series $x_{n+1} = xn - f(x_n) / f'(x_n)$ (n=0, 1, 2...), as shown in fig. 4.

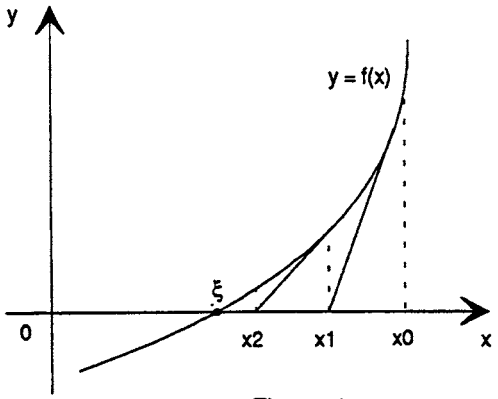


Figure 4.

The application of Newton's method enables fine compensation for:

- non-linearity of the voltage / frequency transfer function of the pilot.
- non-linearity of the pilot converter command.
- shallow gradient temperature variations.

The second filter is adaptive, with its time constant depending on the required frequency correction. At startup of the TE 670, the time constant is small, which enables correction to be quick, but relatively coarse. Under normal operating conditions, fine frequency corrections involve long time constants.

The curves of Figures 5 and 6 are results of the simulation that developed the Frequency DSL. This simulation covered 2 days. Figure 5 presents the time differences at the output of the phase comparator. Figure 6 presents the offset frequency of the TE 670.

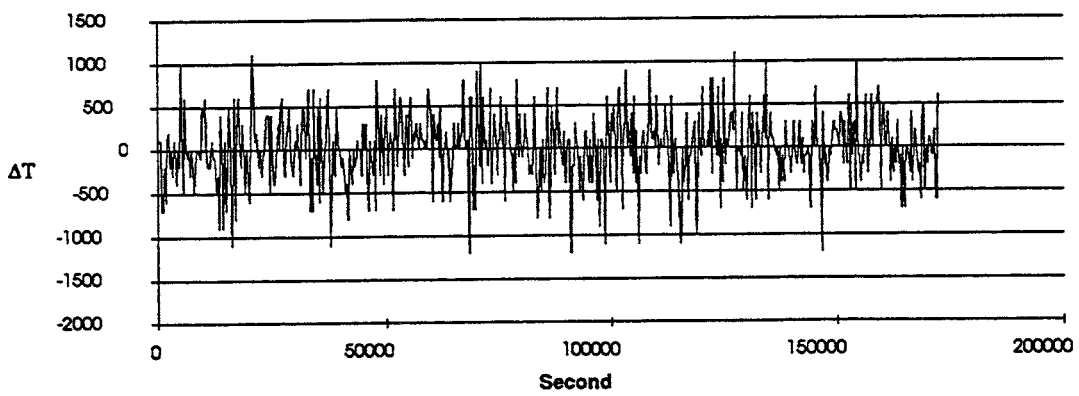


Figure 5.

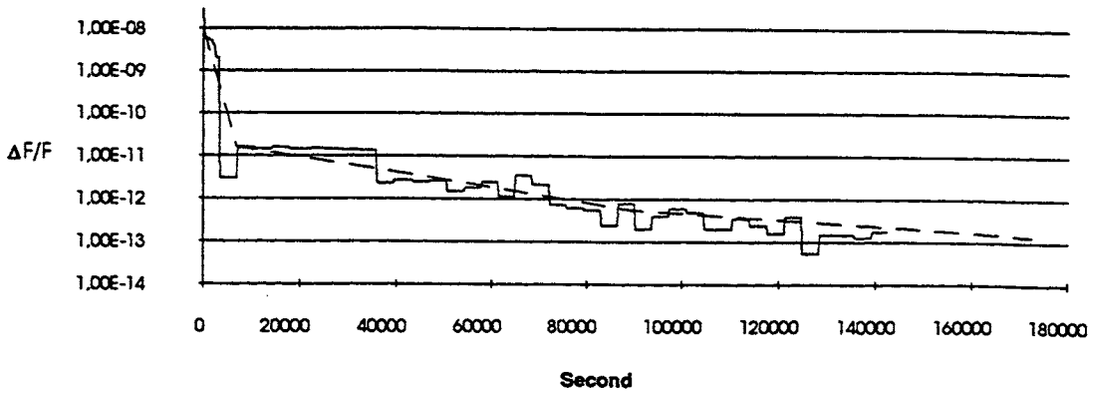


Figure 6.

5 - PRESENTATION

The TE 670 is configured in the 19" format, in a desktop cabinet, 3 standard units high (WxDxH 483 x 370 x 133 mm). Power supply can be 110 / 220 Vac, at 48 to 440 Hz. A separate input port provides for no-break backup battery supply, at 18 to 32 Vdc.

Operations are microprocessor-controlled (68030).

The Man / Machine Interface (MMI) is: a backlighted alphanumeric LCD panel; capacity 16 lines of 40 characters each; associated with 5 programming keys and a digital processing keyboard.

GPS reception, in the L1 band (1.575 MHz) via a 5-channel pc board, including decryption of P/(Y) code (in protected version.)

The T/F unit is fully equipped for all signal processing and digital slaving operations. This unit is piloted by an OCXO BVA, with a low phase-noise characteristic.

In its standard version, the TE 670 delivers the following outputs:

- 5 MHz (1 Vrms / 50 Ω);
- 1PPS (TTL / 50 Ω);
- message, standard time STANAG 4430;
- system status (RS 232 link).

The unit also includes an input port for an external dating signal: 1PPS (TTL / 50 Ω).

6 - TESTING METHODS

The R&D arm of TEKELEC TECHNOLOGIES has developed the PRM Platform, to provide time and frequency standards for testing current and future measuring instruments. See Figure 7 [3].

Three major units make up the PRM:

- a primary time and frequency standard;
- time / frequency measuring instruments;
- modules for validation of time code interfaces.

6.1 Reference station

This unit supplies and distributes frequency and time references, in the form of a UTC time scale UTC (PRM).

The unit comprises 2 frequency sources: active and backup. The active reference source is furnished by a HP 5071A (option 001) Cesium standard. An OSA 3210 source serves as backup. These standards are located in a protected site and kept at: 23 ± 0.5°C. Battery backup is also provided.

The reference frequency is transmitted to the master redundant clock that handles the local UTC time scale, UTC (PRM). The stability of the local time scale is ensured by continuously comparing it with the Paris Observatory time scale (UTC (OP)), via GPS.

The measurement station comprises an Allan Osborne TTR6 GPS receiver, which accepts the reference frequency and the 1 pps time signal from the frequency standard. A HP 9000 workstation monitors 40 distinct common view measurements each day. The characteristics of these measurements (such as number, hour, satellite ID number) are specified at half-year intervals, by the Paris Observatory. The results are sent daily to the Paris Observatory, which analyzes them and reports the intervals between UTC (PRM) and UTC (OP).

The data processing system automatically records the receiver and any exchanges of data with the Paris Observatory. Measurements are processed and results are presented in several forms:

- raw data;
- graphics, either displayed on the screen or plotted on paper;
- creation of log files which are saved to disk.

Statistical handling may be applied to data, such as: averages, standard deviations, maxima, and minima.

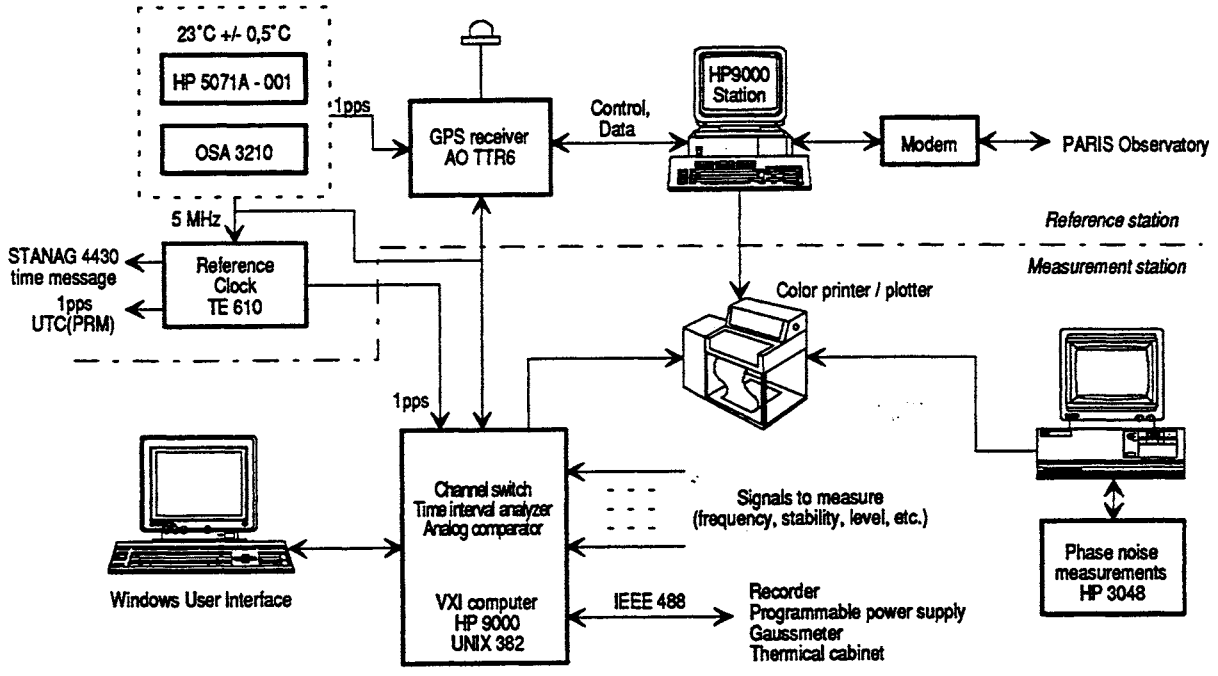


Figure 7.

6.2 MEASUREMENT STATION

This station measures the performance of frequency and time sources, and plots the characteristic curves for new oscillators.

Time and frequency sources are characterized by long and short stability measurements and by measurement of phase noise. Signals from the reference station perform these measurements.

The unit comprises a network of measuring instruments, operated via an IEEE 488 bus:

- a VXI mainframe;
- a phase noise measurement system;
- external measurement devices.

The VXI mainframe groups the following:

- a HP 9000 / V382 processor unit;
- a HP 1472A channel switcher;
- a Racal 2351 time interval analyzer;
- a HP 2308 analog comparator.

The phase noise measurement system, type HP 3048, comprises a phase-locked loop unit, an FFT spectrum analyzer, and a PC with specific data processing software.

All other elements of this unit are also controlled via the IEEE bus, such as: thermal cabinet, precision power supply, magnetometer.

The main characteristics are as follows:

- measurement of phase noise $S\phi(f)$:
  . level: 0 to -169 dB
  . analysis: 1 to 1000 Hz of the carrier wave
- short term stability $\sigma(\tau)$:
  . stability: 10^{-7} to a few 10^{-12}
  . time (τ): 0.1 s to 100 s
- long term stability:
  . stability: 10^{-7} to approx. 10^{-12}
  . period: programmable

6.3 Validation of Time Signals and Codes

In addition to the above mentioned instrumentation, this unit provides time decoding systems: IRIG, parallel BCD. The test of the STANAG 4430 interface is completed by the Tekelec Telecom TE 804 equipment, which processes the following functions: analysis; simulation; transmission error rate; time delays; and time interval between the 1 pps output of the reference clock and the one from the RTT.

7 - PERFORMANCE

Complete data on TE 670 performance can be obtained via the TEKELEC PRM Platform. Because the Platform takes part in UTC time-scale development, TEKELEC has full capability for qualifying a time scale.

A TE 670 time scale provides a degree of precision that is better than ±100 ns relative to UTC (USNO).

Moreover, PRM can measure the accuracy of a frequency source relative to the average frequency standardized by TAI. These results are furnished on a deferred schedule: they must be compared with the Primary Standard Frequencies, which are published in bulletin H of BIPM.

For the TE 670, the frequency accuracy attained over 24 hours is better than 2×10^{-12}. If GPS reception is lost, the resulting frequency drift will be less than 2×10^{-11}/day.

8 - CONCLUSION

We have demonstrated that the TE 670 is fully capable of providing Time and Frequency references that can replace cesium clocks in radionavigation and radiocommunication application, and act in metrology applications to calibrate and synchronize equipment, systems and networks.

This system delivers a time scale better than ± 100 ns, relative to UTC(USNO).

Its long-term frequency performance is the same as the GPS system, yielding a frequency accuracy of 2×10^{-12} over 24 hours.

TEKELEC TE 670 performance is independent of the aging of its internal oscillator, and of the intentional degrading injected into GPS signals.

9 - ACKNOWLEDGEMENT

The authors wish to express their thanks toward the French Air Force for their active, cordial support during the development of this system.

10 - REFERENCES

[1] C. THOMAS,
"Real-time restitution of GPS", Proc. of the 7th EFTF (1993).

[2] B. DEMIDOVITCH et I. MARON,
"Eléments de calcul numérique", Moscow Edition (in French).

[3] Y. BOURRAGUE, P. BERTRAND,
F. MICHAUX, V. MOURIER, J-L. NATTA,
"An innovative time distribution system for remote applications", Proc. of the 8th EFTF (1994).

TELECOM APPLICATIONS

IMPACT OF MEASUREMENT CONFIGURATIONS ON THE BEHAVIOUR OF TIME STABILITY QUANTITIES

Roberto Bonello*, Marco Carbonelli, Domenico De Seta**, Antonio Manzalini*, Daniele Perucchini****

** CSELT, Via G. Reiss Romoli 274, 10148 Turin, Italy*
*** Fondazione Ugo Bordoni, Via B. Castiglione 59, 00142, Rome, Italy*

Abstract

The ongoing introduction in the existing transport networks of transmission systems based on Synchronous Digital Hierarchy has made urgent for network operators the implementation of a synchronization network capable of distributing very high stable timing signals. The evaluation of timing performance of such a network is an hot issue both in the scientific community and in regulatory organisms. In this paper two measurement configurations, namely the *independent clocks configuration*, and the *synchronized clocks configuration*, are thoroughly investigated. Experimental results obtained for the case of a Stand Alone Synchronization Equipment unit and a typical Synchronous Digital Hierarchy equipment in the two different measurement configurations are reported and discussed.

1. Introduction

In modern telecommunication networks, the need of ensuring high quality performance paths and connections for the transport of digital services is more and more stringent. In order to obtain a suitable level of performance, it is necessary to use a synchronization network able to distribute timing signals both to the switching nodes and to all types of synchronous systems deployed in the network. Moreover, the ongoing introduction in the existing transport networks of transmission systems based on SDH (Synchronous Digital Hierarchy) [1] has made urgent for network operators the implementation of a synchronization network capable of distributing very high stable timing signals. The design and the maintenance of a synchronization network based on a widespread deployment of timing units, both within *ad hoc* timing equipment and in SDH equipment, require standard criteria and test configurations for evaluating time and frequency stability performance. At present, hot issues, both in the scientific community and within the regulatory organisms, are the choice of meaningful quantities to characterize frequency stability and the evaluation of the impact of measurement configurations on their behaviour. As far as the choice of stability quantities is concerned, in current regulations [2-4] two quantities play a central role, namely the Maximum Time Interval Error (MTIE) and the Time Deviation (TDEV).

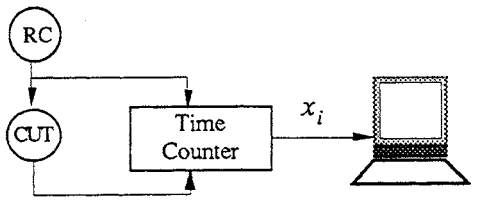
In this paper, measurement configurations adoptable for characterizing both equipment and network time stability performance are thoroughly investigated (sect.2). In particular, two different measurement configurations have been identified: namely, the *independent clocks configuration*, and the *synchronized clocks configuration*. These two configurations can be used to characterize the performance of a single clock and of slaved clock chains. Theoretical performance of the two measurement configurations may be evaluated starting from a fairly general slave clock model (sect.3). Finally, measurement results obtained for the case of a SASE (Stand Alone Synchronization Equipment) unit and an SDH equipment in the two different measurement configurations are reported and discussed. Special attention is payed to the impact of the two configurations on the behaviour of the MTIE (Maximum Time Interval Error) and TDEV (Time Deviation) stability quantities (sect.4).

2. Stability parameters and measurement configurations

Clock stability characterisation for telecommunications purposes is typically based on simple time domain techniques which result in measuring the Time Error TE(t) function [5] between the timing signal generated by the Clock Under Test (CUT), and some reference timing signal, generated by an actual clock considerably better performing than the CUT (RC = Reference Clock). Under this condition, the TE may be regarded as a measure of the error that the CUT incurs in while trying to generate an ideal time scale t. The stability measurements are typically carried out through methods which, using digital counters, allow to extract a sampled version of TE(t): the sequence of N values x_i=TE[t_0+(i-1)τ_0], where t_0 is the initial observation time, τ_0 is the sampling period and i=1, 2,...,N, is used in order to estimate the most important stability quantities. Both in the laboratory and in the field measurement activity, two different measurement set ups are widely used to extract x_i sequences. In the first one, usually referred to as *Synchronized Clocks* (SC) *Configuration* (fig.1), the RC and the CUT are part of a synchronization distribution network and are traceable to the same master clock. In the second one, referred to as *Independent Clocks* (IC) *Configuration* (fig.2), the RC and the CUT operate independently one of the other. Let us note that in the SC configuration the CUT is necessarily a slave clock operating in locked mode, while in the IC configuration the CUT may be either an autonomous clock, a clock slaved to a master clock other than the RC or a slave clock in holdover mode.

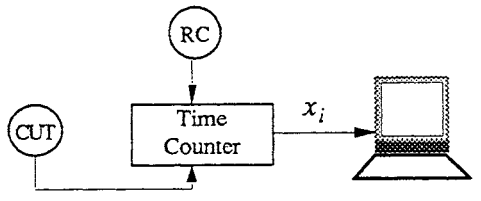
Primary target of a stability characterization activity is to identify which kind of noises affect a timing signal, as well as to evaluate their magnitude. As shown in [6-8], actual clocks may be typically affected by several type of

noise, namely White Phase Modulation (WPM), Flicker PM (FPM), White Frequency Modulation (WFM), Flicker FM (FFM) and Random Walk FM (RWFM) noises. In addition to these noises, an actual timing signal may also show a frequency offset and a frequency drift.



RC =Reference Clock
CUT = Clock Under Test

Fig. 1 - Synchronized Clock Configuration



RC =Reference Clock
CUT = Clock Under Test

Fig. 2 - Independent Clock Configuration.

Both the independent and the synchronized clocks configuration measurement set ups provide useful information on clock performance. As shown in [9], the SC configuration may be adopted when only interested in evaluating the instabilities the CUT adds to the reference timing signal, while the x_i sample sequence obtained by means of the IC configuration reflects also the offset and the drift between the RC and the CUT.

Starting from an x_i sequence, no matter which measurement configuration is adopted, a number of stability quantities can be evaluated: for telecommunications purposes, the preferred quantities are both the Maximum Time Interval Error (MTIE) and the Time Deviation (TDEV) [3][4][9].

The most frequently used estimators for MTIE and TDEV quantities are defined as follows

$$\mathrm{MTIE}(\tau) = \max_{1\le k<N-n}\left[\max_{k\le i<k+n} x_i - \min_{k\le i<k+n} x_i\right],$$
$$n = 1,2,\ldots,N-1$$

$$\mathrm{TDEV}(\tau) = \left\{\frac{1}{6n^2(N-3n+1)}\cdot \sum_{j=1}^{N-3n+1}\left[\sum_{i=0}^{n-1}(x_{i+2n+j} - 2x_{i+n+j} + x_{i+j})\right]^2\right\}^{1/2},$$
$$n = 1,2,\ldots,\left\lfloor\frac{N}{3}\right\rfloor.$$

MTIE is mainly utilized for characterizing the performance of a timing distribution network, as it substantially describes peak to peak differences (errors) between two clocks in assessing time interval duration.

TDEV is currently considered as the best suited time domain estimate for phase noise characterization because its behaviour clearly identifies which noise components are actually dominant, and their magnitudes. Fig.3 shows the different slopes the TDEV exhibits when the most common phase noise components are considered: the actual slopes on an experimental TDEV plot can thus give some indication of relevant phase noise characteristics, whereas simple MTIE results would not. Finally, if drift components are present, TDEV shows a behaviour proportional to τ^2, while a frequency offset does not affect TDEV behaviour.

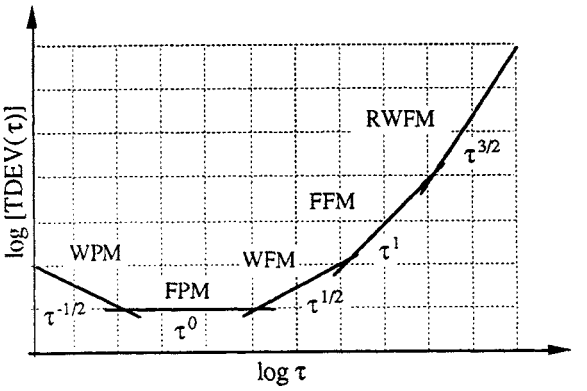


Fig. 3 - TDEV for different phase noise components

3. Theoretical approach

In order to theoretically investigate the impact of measurement configuration on the behaviour of time stability quantities, the well known slave clock noise model shown in fig.4 was adopted. This model was derived from a more general one proposed by Kroupa [14]. In fig.4 the following notation is adopted:

- φ_i [rad] and φ_o [rad] are the additive phase noises on the slave clock input and output signals respectively;
- φ_{VCO} [rad] is the phase noise due to the Voltage Controlled Oscillator (VCO);
- V_{DF} [V] is the voltage noise due to phase Detector and loop Filter (DF);
- K_d [V/rad] and K_0 [rad/(V·s)] are the phase detector and VCO gains respectively;
- $F(s)$ is the loop filter transfer function.

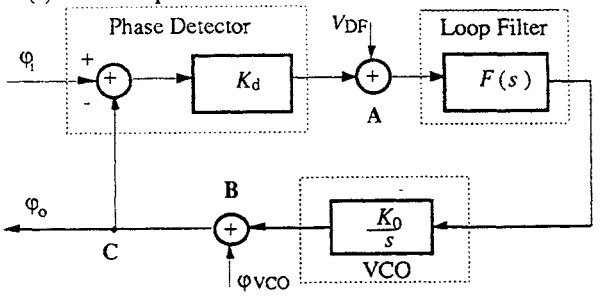


Fig.4 - Slave clock noise model.

Internal noise processes φ_{VCO} and V_{DF} are characterized in terms of PSD, denoted as $S_{VCO}(f)$ and

$S_{DF}(f)$ respectively. These two PSDs may be expressed as follows

$$S_{VCO}(f) = \frac{10^{-7}}{f^3} + \frac{10^{-6.75}}{f^2} + \frac{10^{-10.3}}{f} + 10^{-15.5} \quad [\text{rad}^2/\text{Hz}]$$

$$S_{DF}(f) = \frac{10^{-13}}{f} + 10^{-17} \quad [\text{V}^2/\text{Hz}].$$

The chosen parameter values in above eqns. provide satisfactory models of random fluctuations in electronic devices and in commercial quartz crystal oscillators [9,10].

Referring to fig.4, the transfer functions $H(s)$, $H_A(s)$ and $H_B(s)$ between slave clock input, noise injection points A and B, respectively, and output C, may be expressed by

$$H(s) = \frac{K_0 K_d F(s)}{s + K_0 K_d F(s)}, H_A(s) = \frac{K_0 F(s)}{s + K_0 K_d F(s)} \quad [\text{rad}/\text{V}],$$

$$H_B(s) = \frac{s}{s + K_0 K_d F(s)}$$

Using the superposition theorem, the PSD $S_o(f)$ of total output noise φ_o is simply given by

$$S_o(f) = S_{DF}(f)|H_A(j2\pi f)|^2 + S_{VCO}(f)|H_B(j2\pi f)|^2 \quad [\text{rad}^2/\text{Hz}]$$

under the assumptions of mutually uncorrelated noises, with amplitude small compared to the useful signals, and φ_i=0. In other words, a slave clock in locked mode performs a low-pass filtering on the input, a low-pass filtering on Phase Detector noises and an high-pass filtering on VCO internal noises. Instead, when the slave clock is in holdover mode no filtering action on internal noises is performed, as far as the control loop results open. Analytical calculation of stability quantities from the above model is quite cumbersome. For this reason a computer simulation program using numerical integration algorithms was utilized [11] to simulate an SC measurement set up. The assumption that WPM and FFM noises injected at the output of the VCO are the most important sources in actual slave clocks was made, mainly considering measurement results on slave clock characterization. Based on these considerations and utilizing the slave clock model, the TDEV from 3,000 samples of noise generated at a sampling period of 100 ms was calculated, as reported in fig.5. In this example a 0.02 Hz slave clock bandwidth was chosen. From fig.5, we can easily identify the effects of the noises considered. Special attention should be payed to TDEV behaviour for observation intervals above 200s, where the high-pass filtering action of slave clock on FFM noise shows its major effects.

4. Measurement results

A measurement activity aiming at characterizing the stability of clocks used in telecommunications was recently set up at CSELT laboratories: both SASE units and SDH equipment clocks stability were deeply investigated and characterized [12,13].

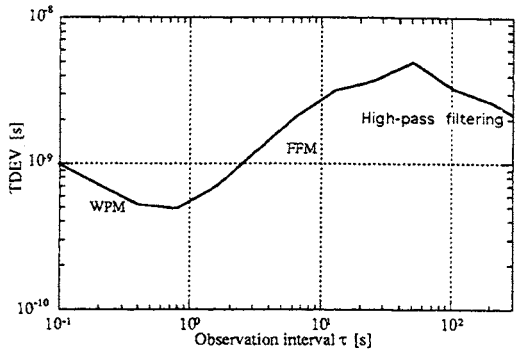


Fig. 5 - TDEV calculated according the model.

Measurements were performed utilizing a Rubidium frequency standard as RC, showing a $5 \cdot 10^{-11}$/day drift.

In figs.6 and 7 the MTIE and TDEV quantities, respectively, of a typical SASE unit are reported for both the SC and IC measurement configurations. The SASE bandwidth was 10mHz. In the IC configuration the SASE unit was in holdover mode, being equipped with an Oven Controlled Crystal Oscillator having a declared stability of $5 \cdot 10^{-9}$/day. The two stability quantities were calculated starting from a sequence of N = 480000 x_i samples with τ_0 = 7.5 ms. From inspection of fig. 7 we can see that:

- for observation intervals up to 3s, WPM noise is dominant;
- for observation intervals from 3s to 1000s, FFM noise is dominant;
- for short observation intervals, the level of noise is of the same order of magnitude of the test bed floor noise.

According to theoretical results presented in the previous section, let us note from fig.7 the effects on synchronized clocks TDEV of the slave clock filtering action, for observation intervals above 20s, where, on the contrary, the TDEV of the SASE in holdover mode persists in the FFM typical behaviour.

In figs.8 and 9 the MTIE and TDEV quantities, respectively, of a typical SDH equipment clock are reported for both the SC and IC measurement configurations. The CUT bandwidth was 0.1 Hz. In the IC configuration the SDH equipment clock was slaved to a Rubidium clock having the same performance as the RC. The two stability quantities were calculated starting from a sequence of N = 3600 x_i samples with τ_0 = 1 s. From fig.7 and 8 we can say that:

- for observation intervals from 1s up to 100s, unfiltered FFM noise is dominant;
- for observation intervals below 1000s, no significative differences can be highlighted between SC and IC configurations;
- for observation intervals greater than 1000s, the MTIE reveals the presence of a frequency offset between the RC and the CUT.

Finally, the loop filtering action on FFM noise is revealed for observation intervals greater than 20s.

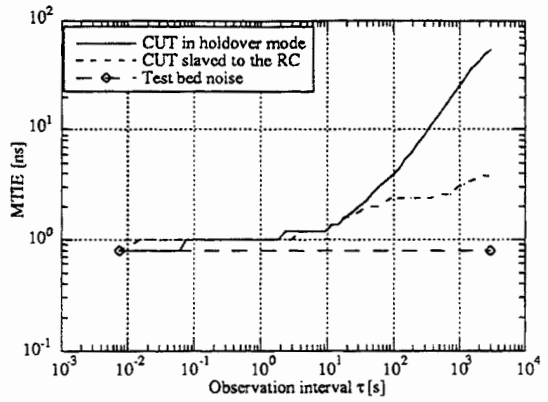


Fig. 6 - MTIE of a typical SASE unit.

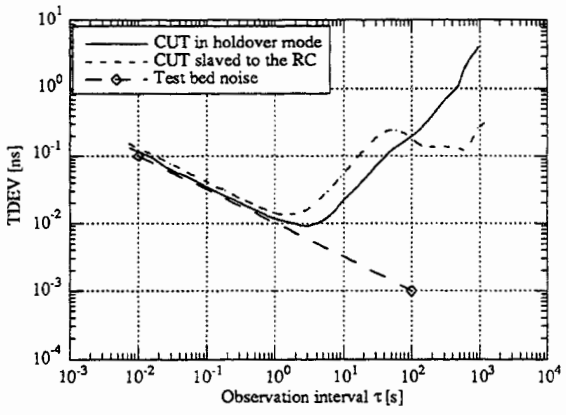


Fig. 7 - TDEV of a typical SASE unit.

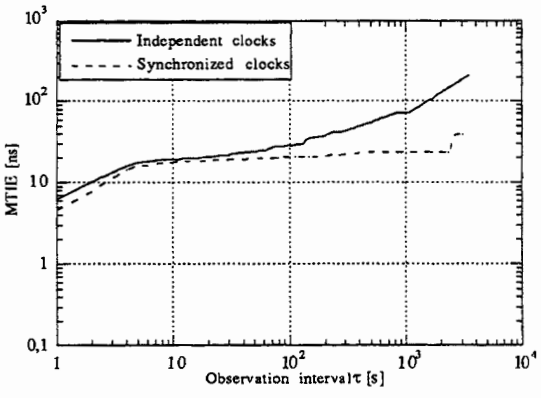


Fig. 8 - MTIE of an SDH equipment clock.

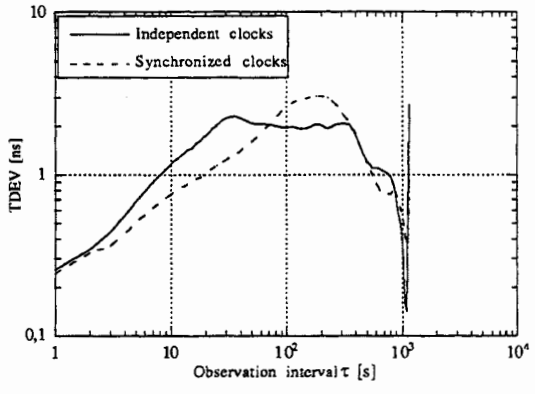


Fig. 9 - TDEV of an SDH equipment clock.

5. Conclusions

In this paper the issue of assessing time stability quantities on slave clocks used in telecommunication networks was addressed. Starting from a simple slave clock model, the most frequently used estimators of MTIE and TDEV were calculated for both SASE and SDH equipment clock both in the synchronized and independent clocks measurement configurations. Also some experimental results were presented, derived from measurement activity carried out in order to characterize the timing performance of typical synchronization units. The effects related to the different configurations used for measurement were discussed, highlighting the particular desirable features displayed by TDEV quantity.

Acknowledgement - The work of M. Carbonelli, D. De Seta and D. Perucchini was carried out in the framework of the agreement between the Italian PTT and the Fondazione Ugo Bordoni.

REFERENCES

[1] CCITT XVIII, Recs 708, 709, Melbourne 1988.
[2] CCITT XV, Rec. G.783, Geneva, 1990.
[3] CCITT XVIII, Rec. G.811, 812, Melbourne, 1988.
[4] CCITT XVIII Draft Rec.G.81s
[5] M. Carbonelli, D. De Seta, D. Perucchini, S. Ruffini, *Synchronization of SDH networks: slave clocks model and stability measures*, Proc IEEE GLOBECOM '92, Orlando, 6-9 December 1992.
[6] S.R. Stein, *Frequency and time-Their measurement and characterization*, in "Precision Frequency Control", vol.2, Academic Press, New York, 1985.
[7] J. Rutman, F. L. Walls, "*Characterization of frequency stability in precision frequency sources*", Proceedings of the IEEE, Vol. 79, n.7, July 1991.
[8] D.W. Allan, M.A. Weiss, J.L. Jespersen, "*A frequency domain view of time domain characterization of clocks and time and frequency distribution systems*", 45° Annual Symposium on Frequency Control, 1991, pp.667-678.
[9] M. Carbonelli, D. De Seta, D. Perucchini, *Clock stability measures for the assessment of synchronization network performance*, Proc. of IEEE SICON ICIE'93, Singapore, Sept. 93, pp.188-192.
[10] M. Carbonelli, D. De Seta, D. Perucchini, "TIErms accumulation along slave clock chains", *Proc. ICC'93*, Geneva, May 1993, pp. 816-820.
[11] S. Bregni, M. Carbonelli, D. De Seta, D. Perucchini, *Impact of slave clock internal noise on Allan variance and Root Mean Square Tima Interval Error measurements*, Proc. of IMTC '94, Hamamatsu, May 1994.
[12] R. Bonello, A. Manzalini, A. Godone, C. Novero, P. Tavella, *Stability characterization of Stand Alone Synchronization Equipment*, Proc of EFTF'95, Besançon, March 1995.
[13] R. Bonello, S. Bregni, M. Carbonelli, D. De Seta, A. Manzalini, A. Mariconda, D. Perucchini, L. Valtriani, *Preliminary experimental results for the development of the new Italian synchronization network*, to be published in Proc. of IEEE ICC'95, Seattle, June 1995.
[14] V.F. Kroupa, "Noise properties of PLL systems", *IEEE Trans. on Comm.*, vol. Com-30, n.10, October 1982, pp. 2244-2252.

ROBUST CONTROL OF G.812 CLOCKS

ANDRZEJ DOBROGOWSKI

Institute of Electronics and Telecommunication
Poznań University of Technology, Poland

ABSTRACT

The need for robust control of G.812 clock is explained. Regression line estimation methods are presented and characterized in terms of the breakdown point of estimator. Some statistical analysis of the raw data for a sample of VCXO's control signal has been done. The robust methods have proved their superiority over nonrobust ones.

INTRODUCTION

The layering concept has been already proved as being the very foundation of digital communication network functional description. A layer supporting the operation of other layers is the synchronization layer comprising synchronization network. The synchronization network should have topology of tree type. It consists of primary reference clock, node cloks and synchronization links. As a node clock serves G.812 clock. The call of the clock comes from designation of ITU-T (former CCITT) recommendation. G.812 clock is a high quality slave clock which receives its synchronizing signal from other clock through synchronization link. As G.812 clock is a high quality clock therefore its control signal must not deteriorate an intrinsic quality of the clock. On the other hand clock's synchronizing signal suffers from interferences during transmission, generation and accumulation of jitter and wander in cascaded regenerators and multiplexes, cascaded G.81s clocks of SDH network elements having relatively poor quality timing signals, phase jumps, corruptions of timing recovery process as well as some imperfections of the receiver.

Signal processing technique (involving averaging and filtering) usually adopted to produce a signal controlling the clock does not eliminate outliers. Obviously, the outliers do not reflect the required difference between the phases of local and remote clocks, therefore they must not be used to control G.812 clock.

In the paper several control algorithms are investigated keeping in mind that we need the control algorithm which does not only delete outliers, but rather produces the appropriate control signal despite of outliers. During normal mode of operation of the clock we are interested in finding the proper control signal at a corresponding instant. G.812 clock possesses very good memory of its previous states thus knowing the linear regression of the control signal helps to work out the proper control signal. Knowledge of linear regression model can also be exploited in holdover mode of operation of the clock in order to produce changing in time control signal of the clock.

SYNCHRONIZATION SUPPLY UNIT - G.812 CLOCK

Synchronization supply unit (SSU) provides its environment with timing signal. Schematic diagram of the SSU featuring the quality of G.812 clock is given in Fig.1. The G.812 slave clock is synchronised to that of the input signals which has the best quality. The phase detector PD compares the phase of the selected reference

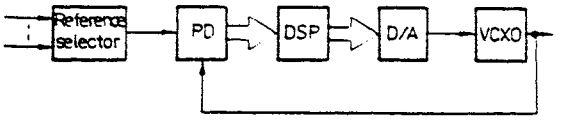


Fig.1. Synchronisation Supply Unit - G.812 clock

with the phase of the local oscillator (VCXO). The output of the phase detector is processed in DSP circuitry of the SSU. The usual processing involves averaging and filtering. As the result we receive the word for D/A converter setting the frequency of the VCXO. Oven control of the voltage-controlled crystal oscillator (VCXO) results in its high stability.

More sophisticated signal processing enables to estimate changing the free-running frequency in course of time (the drift of the free - running frequency). Knowing the drift it is possible to limit its undesirable effect on the oscillator's frequency stability by setting suitable signal at the proper instants during holdover mode of the clock. The estimated drift of the VCO one can use to modify the VCXO's current control signal resulting from last measurements (PD outputs) thus receiving corrected control signal.

From point of view of statistics we are faced with three problems:

- estimation of regression equation (frequency drift of the oscillator),
- estimation of location parameter (current control signal),
- correction of the location parameter (corrected control signal).

In what follows only estimation of regression equation for frequency drift of the oscillator wil be considered.

Because of high inherent stability of the VCXO used in G.812 clock (SSU) on one hand and possible disturbances of its control signal on the other one it is very important to adopt robust statistical procedures for estimation.

REGRESSION LINE ESTIMATION

The relative drift of free-running frequencies of synchronizing and synchronized oscillators is responsible for a trend observed in varying control signal of G.812 clock. Rather very stable master and slave oscillators justify the assumption that the trend observed during rather long period of time is linear. For this reason linear regression was considered as an adequate model for frequency drift of the high quality VCXO.

As we investigate the time dependence of VCXO's control signal the proper choice is the simple regression model

$$y_i = ax_i + b + e_i \quad , \quad i = 1,\ldots,n \qquad (1)$$

where n is the sample size (number of cases), x is the carrier (or explanatory variable), y is the response variable and e is a random error term.

The purpose of linear regression analysis is to fit regression line to observed variables x_i and y_i, for i=1,...,n. It means that we try to estimate the slope a and the intercept b of the regression line.

Knowing the estimates (regression coefficients) $\hat{a}$ and $\hat{b}$ we can find estimated value $\hat{y}_i$ of y_i according to the relation

$$\hat{y}_i = \hat{a}\,x_i + \hat{b} \tag{2}$$

The difference between the observed value y_i and estimated value $\hat{y}_i$ is called the residual r_i of the ith case. Therefore

$$r_i = y_i - \hat{y}_i \tag{3}$$

It should be emphasized that for estimation of VCXO's free-running frequency drift we need estimators which are insensitive to outlying values of phase detector's output signal.

The breakdown point of the estimator characterizes the ability of the estimator to cope with outliers. Rousseeuw and Leroy [1] explain this characteristic as follows. Let us take any specific sample of n data points

$$Z = \{(x_1, y_1), \ldots, (x_n, y_n)\} \tag{4}$$

and let T( ) be a regresssion estimator. This means that applying T( ) to such a sample yields regression coefficients

$$T(Z) = \begin{bmatrix} \hat{a}, & \hat{b} \end{bmatrix} \tag{5}$$

Then consider all possible corrupted samples Z' that are obtained by replacing any m of the original data cases by arbitrary values.
Bias (m; T,Z) that can be caused by corruption is defined as

$$bias(m; T, Z) = \sup_{Z'} \|T(Z') - T(Z)\| \tag{6}$$

where the supremum is over all possible Z'. The breakdown point $\varepsilon^*(T,Z)$ of the estimator T at the sample Z is

$$\varepsilon_n^*(T, Z) = \min\left\{\frac{m}{n}; bias(m; T, Z) \text{ is infinite}\right\} \tag{7}$$

Thus the breakdown point is the smallest fraction of corruption that can cause the estimator T to take on values arbitrarily for from T(Z).

The well-known least squares (LS) method gives regression estimator having the breakdown point equals

$$\varepsilon_n^*(T, Z) = \frac{1}{n} \tag{8}$$

It means that one outlier is enough to carry regression estimates over all bounds. We say that LS has the breakdown point of 0% because for increasing sample size n right side of (8) tends to zero.

The least median of squares (LMS) estimator [1], given by

$$\underset{[\hat{a},\hat{b}]}{\text{Minimize}} \ \text{med}_i \ r_i^2 \tag{9}$$

is very robust with respect to outliers. Its breakdown point reaches the highest possible value of 50%.

Another estimator possessing the high breakdown point results from reweighted least squares (RLS) method [1] defined by

$$\underset{[\hat{a},\hat{b}]}{Minimize} \ \sum_{i=1}^{n} w_i r_i^2 \tag{10}$$

where w_i is the weight and r_i is the residual of case i with respect to the LMS fit. One can use the following weights

$$w_i = \begin{cases} 1 & \text{if } \left|r_i / \hat{\sigma}\right| \le 2.5 \\ 0 & \text{if } \left|r_i / \hat{\sigma}\right| > 2.5 \end{cases} \tag{11}$$

where $\hat{\sigma}$ is a scale estimate given by

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} w_i r_i^2}{\sum_{i=1}^{n} w_i - 2}} \tag{12}$$

RAW DATA

Raw data for a sample of VCXO's control signal are given in Table I. In the experiment [2] the VCXO was synchronized to the reference oscillator. The experiment lasted 135. Every day, at midnight, the reading of the current VCXO controlling signal was carried out. During the experiment four times we were forced to suspend the reading schedule for several days. The lacking data were supplemented artificially (in Table I they are marked with asterisk). They are mainly outliers. A scatterplot is given in Fig.2.

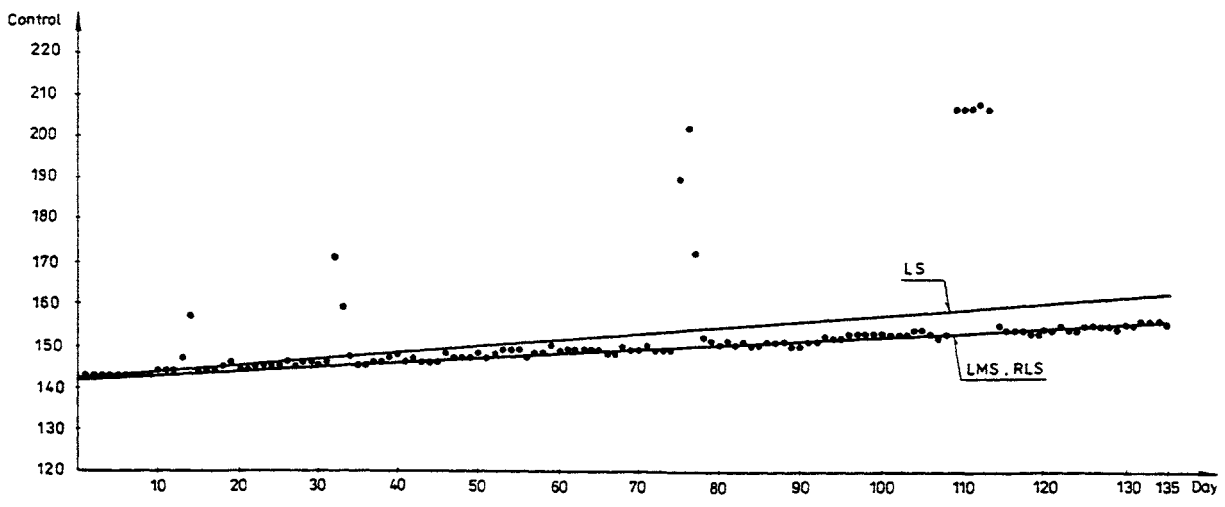


Fig.2. Scatterplot with regression lines for VCXO's control signal

TABLE I

Raw data for a sample of VCXO's control signal

Day	Control	Day	Control	Day	Control	Day	Control
1	143	35	159	69	149	103	153
2	143	36	147	70	149	104	154
3	143	37	145	71	150	105	154
4	143	38	145	72	149	106	153
5	143	39	146	73	149	107	152
6	143	40	146	74	149	108	153
7	143	41	147	75	190*	109	207*
8	143	42	148	76	202*	110	207*
9	143	43	146	77	172*	111	207*
10	144	44	147	78	152	112	208*
11	144	45	146	79	151	113	207*
12	144	46	146*	80	150	114	155
13	148*	47	146	81	151	115	154
14	157*	48	148	82	150	116	154
15	144	49	147	83	150	117	154
16	145	50	147	84	151	118	153
17	144	51	147	85	150	119	153
18	144	52	148	86	151	120	154
19	144	53	149	87	151	121	154
20	145	54	149	88	151	122	155
21	146	55	149	89	150	123	154
22	144	56	147	90	150	124	154
23	144	57	148	91	151	125	155
24	145	58	148	92	151	126	155
25	145	59	150	93	153	127	155
26	145	60	149	94	152	128	155
27	145	61	149	95	152	129	154
28	146	62	149	96	153	130	155
29	145	63	149	97	153	131	155
30	146	64	149	98	153	132	156
31	146	65	149	99	153	133	156
32	145	66	148	100	153	134	156
33	146	67	148	101	153	135	155
34	171*	68	150	102	153		

REGRESSION ANALYSIS

A regression analysis was carried out using the program which is called PROGRESS. The program performs a robust regression analysis based on the least median of squares (LMS) method. Description of PROGRESS is given in the book [1].

Main results of regression analysis for nonrobust LS as well as robust LMS and RLS are contained in Table II, Table III and Table IV respectively.

TABLE II

Least squares regression

Variable	Coefficient	Stand. error	T - value
Day	0.14755	0.02595	5.68508
Constant	142.433	2.03416	70.0205

Sum of squares = 18368.02000
Degrees of freedom = 133
Scale estimate = 11.75183

TABLE III

Least median of squares regression

Variable	Coefficient
Day	0.098360
Constant	142.5656

Final scale estimate = 0.71090
Coefficient of determination = 0.98563

TABLE III

Reweighted least squares based on the LMS

Variable	Coefficient	Stand. error	T - value
Day	0.09760	0.00164	59.54109
Constant	142.584	0.12779	

Weighted sum squares = 60.90740
Degrees of freedom = 121
Scale estimate = 0.70948
Coefficient of determination (R^2) = 0.96700
There are 123 points with non-zero weight
Average weight = 0.91111

Reegression lines are shown in Fig.2.

CONCLUSION

A robust regression analysis exploiting MLS or RLS methods enables to establish an actual trend of VCXO's frequency drift. It is because these methods cope with outliers very effectively. Regression lines resulting from MLS or RLS are better models of drift than LS regression line (e.g., compare relevant coefficients of determination or the course of regression lines in relation to scatterplot).

The relative gain of the VCXO which was used in the experiment is equal 2.8×10^{-9}. Therefore LS method estimated the drift of the VCXO to be equal 4.13×10^{-10}/day, but MLS and RLS gave the drift equal to 2.75×10^{-10}/day and 2.73×10^{-10}/day respectively.

ACKNOWLEDGMENTS

I would like to express my gratitude to Professor Peter J.Rousseeuw for providing me with program PROGRESS.

REFERENCES

[1] P.Rousseeuw and A.Leroy, Robust Regression and Outlier Detection, Wiley, New York, 1987.

[2] K.Borowiak, Predykcja Częstotliwości Drgań Własnych Generatora, D.Sc.Thesis, Poznań University of Technology, Poznań, 1988.

ABSOLUTE FREQUENCY MEASUREMENT OF THE 3-0 R(21) TRANSITION OF CO AT 1.5605 μm FOR OPTICAL COMMUNICATION STANDARDS

by D. J. E. Knight,
DK Research, 110 Strawberry Vale, Twickenham, Middx. TW1 4SH, UK,

and K. I. Pharaoh and D. A. Humphreys,
Division of Electrical Science, National Physical Laboratory, Teddington, Middx. TW11 0LW, UK.

ABSTRACT

An absolute frequency measurement has been made on the 3-0 R(21) line of CO at 1.560500 μm by frequency doubling and mixing with a Rb-stabilized laser at 780 nm [2]. The frequency of a diode laser locked to the Doppler- and pressure-broadened peak at 40 torr, 94°C was 192 113 068.1 ± 3.5 MHz. The pressure shift between 40 and 250 torr was -(0.25 ± 0.03) MHz/torr, and the zero-pressure intercept was 192 113 078.1 ± 4.0 MHz. Frequency differences of (-45 025(34) MHz and +41 757(37) MHz to the adjacent R(20) and R(22) lines respectively were measured by lambda meter.

1. INTRODUCTION

The frequency measurement of the 3-0 R(21) CO line at 1.560500 μm has been carried out to meet part of a requirement by the National Measurement System for absolute frequency measurements on two frequency standards in the 1.5 μm optical communication band. Investigations on the CO line with a diode laser spectrometer and vacuum lambda meter preparatory to this work have been described [1].

The frequency reference used for the measurement was a Rb-stabilised laser at 780 nm, of which the frequency was measured by reference to the National Physical Laboratory 633-nm He-Ne/I_2 frequency standard [2]. It is of interest to compare the results with the values determined from the molecular constants, which inter-alia make use of absolute frequency measurements up to the 2.3μm region [3,4,5].

2. APPARATUS AND METHOD

The general scheme of the experiment is shown in Figure 1. The frequency measurements were made by frequency doubling the amplified signal of an extended-cavity 1.56 μm diode laser in a KTP crystal [6] and mixing with the Rb-stabilized laser at 780 nm [2]. This gave a beat frequency near 2 GHz with the 1.56 μm laser at the centre of the CO line. The CO absorption

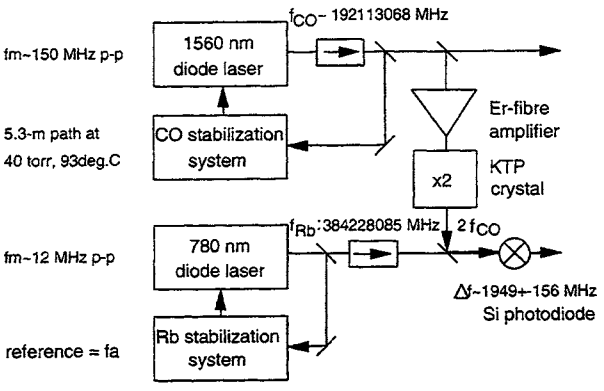
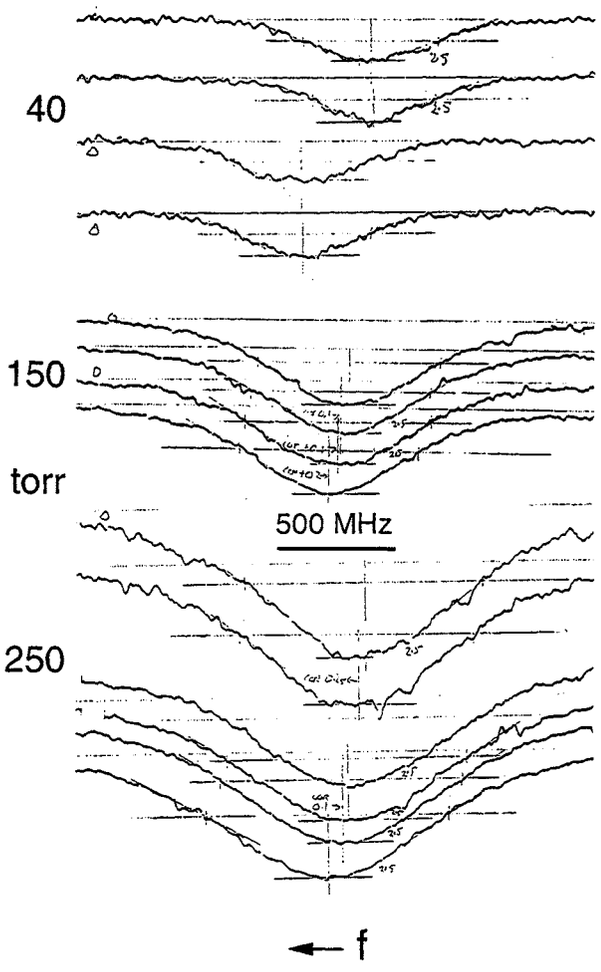


Figure 1: scheme of CO absolute frequency measurement

was observed in transmission after two passes of a 2.66-m absorption cell, at pressures of 40, 150 and 250 torr. Typical traces are shown in Figure 2. The size of the dip varied between 7.4(4)% and 17.5(4)%, and the pressure-broadened widths at full width half maximun (FWHM) varied between 0.6 and 1 GHz, as in Figure 3. The zero-pressure extrapolation of the width indicates a residual, mainly Doppler, width of about 0.5 GHz, Figure 3. The cell wall temperature was raised to 94°C to populate better the high-J states [4].

The Rb-stabilised laser system included a diode laser locked to an external invar-spaced optical spectrum analyser cavity for spectral narrowing. For the measurements described here, this laser was then set, by tuning the cavity, to the Doppler-limited f_a peak of ^{87}Rb. The f_a-peak frequency was measured, and later checked, in separate experiments against the d/f saturated-absorption peak on the same line [2], using heterodyne techniques to measure the frequency difference. The value for the d/f peak measured for the Rb-stabilised system and gas cell used in this experiment was $f_{d/f}$ = 384 227 981.7 ± 0.4 MHz [2]. The value measured and used for the "a" peak was f_a = 384 228 085.2 ± 2.3 MHz. The uncertainty is a



the 40-torr relative vertical scale is 70%

Figure 2: typical plots of the CO 3-0 R(21) absorption dip at 192.1130 THz

1-sigma equivalent value combining both statistical and estimated components.

Measurements were made over several days at each pressure and by two different methods: (i) by measurement of graphical plots of the absorption line obtained by scanning the laser across the line, as in Figure 2, calibrating the frequency from the heterodyne beat, a few MHz wide, by briefly lifting the recorder pen, and
(ii) by direct heterodyne measurements on the laser locked by the first-derivative method to the line peak, using frequency modulation of about 150 MHz peak-to-peak. After frequency doubling to 780 nm this produced a beat signal about 300 MHz wide on the spectrum analyser. A single measurement comprised using the peak hold facility on the spectrum analyser, and reading the edges at 10 and 45 seconds to interpolate the centre frequency.

3. RESULTS

The results obtained by method (i), the graphical method, are shown as a function of pressure in Figure 4, and those obtained by method (ii), the locked-laser method, are likewise shown in Figure 5. The lines shown are linear regression fits to the data at the three pressures used, with dotted lines drawn through the ends of the 1-sigma error bars to obtain the uncertainty spread at the zero pressure intercept. The uncertainties in Figures 4 & 5 represent the statistical spread of the results.

Details of the results derived from Figures 4 and 5 are shown in Table 1, which also gives the number of measurements and the number of days over which they were made. The graphical method refers to the line centre at the half-height of the dip, and the locked laser method obtains an average position for the peak of the line.

It can be seen in Table 1 that the two methods agree as to the zero-pressure intercept, and as to the pressure shift visible in Figures 4 and 5, but that the results obtained

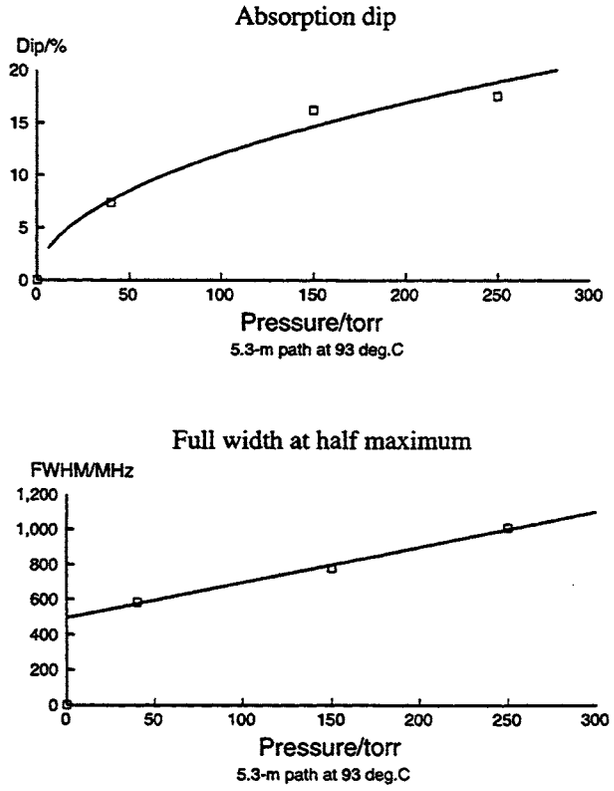


Figure 3: depth and width of CO absorption vs pressure

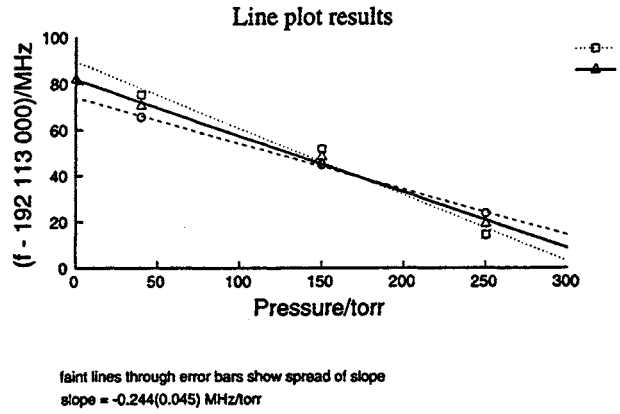


Figure 4: CO 3-0 R(21) line measured by graphical method

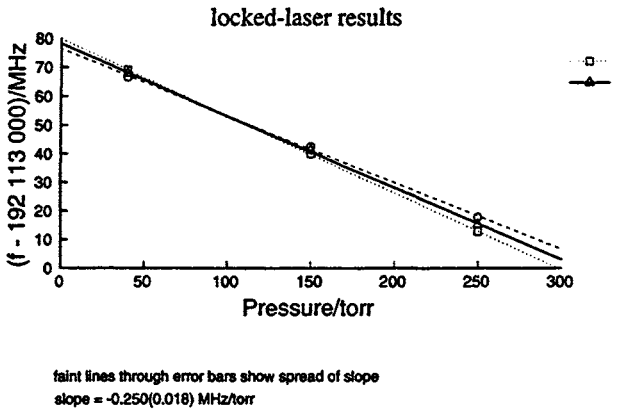


Figure 5: CO R(21) line measured with locked laser

by the locked-laser method are more precise. The uncertainties given include estimated components for e.g. background-slope offset of the locked laser, and accuracy of the pressure measurement, in the case of the pressure shift, and are 1-sigma equivalent values. The preferred results are shown in bold type.

In order to check for alternative lock points for the CO-stabilised laser, it was frequency scanned to the adjacent 3-0 R(22) and R(20) lines. Measurements were made using a Burleigh vacuum lambda meter with interpolation of the last digit by repeated measurement [1]. Measurements were alternated between the R(21) and the other lines to obtain frequency ratios and differences against the 3-0 R(21) line. In this way the R(21) line becomes a calibration reference and the internal He-Ne 633 nm laser acts only as a short-term reference. These results are shown in Table 2. The zero-pressure extrapolation of the frequencies is also shown, and compared with the values derived from the CO constants and used for calibration by Sasada [7]. The R(22) and R(21) measurements are in agreement, but the R(20) line values differ by about one sigma of the present measurements.

The following is an overall summary of the results of these measurements. The frequency of a diode laser locked to the Doppler- and pressure-broadened peak of the 3-0 R(21) line of CO at 40 torr, 94°C, was 192 113 068.1 ± 3.5 MHz. The pressure shift between 40 and 250 torr was -(0.25 ± 0.03) MHz/torr, and the zero-pressure intercept was 192 113 078.1 ± 4.0 MHz. Frequency differences of (-45 025(34) MHz and +41 757(37) MHz to the adjacent R(20) and R(22) lines respectively were measured by lambda meter.

4. CONCLUSION

This measurement of the 3-0 R(21) CO line is the first direct frequency measurement of a CO line in the 1.5 μm band. The zero-pressure extrapolation agrees with predictions from the molecular constants, which are of comparable accuracy, and the pressure shift is weak and similar to that found for the 2-0 R(24) line in absolute frequency measurements at 2.3 μm [3].

Table 1: CO v=0-3 R(21) line measurements vs Rb d/f line via Rb 'a' peak

Quantity	f/MHz	df/MHz	n	notes: method
P=0 intercept	192 113 081.8	7.9	17+21+18# 4 2 2	graph: half max. days
P=0intercept	**192 113 078.1***	**4.0**	51+18+30 3 2 2	locked-laser days
pressure shift	-0.244	0.051		graph
(MHz/torr)	**-0.250**	**0.031**		locked-laser
working standard				
P=40 torr,94C	192 113 070.5	5.3	17	graph: half max.
P=40 torr,94C	**192 113 068.1**	**3.5**	51	locked laser

Notes:-

#: numbers of measurements at pressures of 40, 150 & 250 torr respectively; the numbers of days on which the measurements were taken are given below;

*: cf. the CO energy-level prediction (P=0): 192 113 077.4 ± 1.4 MHz [5].

Table 2: measurements of CO adjacent-line spacings by lambda meter

Quantity	f/MHz	df/MHz	n	notes: method
R(22)-R(21)	**41 757**	**37#**	2	at 40 torr, 92°C
R(21)-R(20)	**45 025**	**34#**	2	at 40 torr, 92°C
P=O values, assuming same pressure shift as for R(21);				cf.spectroscopy [7]
3-0 R(22)	192 154 835	37	2	192 154 833(1+)
3-0 R(21)	**192 113 078.1**	**4.0**		192 113 080.6(1+)
3-0 R(20)	192 068 053	34	2	192 068 083(1+)

Note:-

#: includes an estimated uncertainty contribution of ± 30 MHz.

The result also compares frequency measurements made by reference to the NPL 633-nm He-Ne/I_2 frequency standard, which is traceable to Cs via the BIPM *Mise en pratique* [8], and those made on CO by laser frequency chain from Cs to the 2.3 μm region [3]. The good agreement found is thus of interest, in spite of the present Doppler-limited accuracy of about ± 2 parts in 10^8.

For optical communications, this CO measurement provides the most accurate absolute frequency standard to date in the Er-fibre amplifier band, and implies the availability of the yet more accurate half-frequency of the 384 THz Rb standard [2]. The accuracy is of the same order of that of the ^{20}Ne Lamb-dip stabilised laser at 1.523 μm [1,9,10], which however lies outside the Er-fibre band.

ACKNOWLEDGEMENTS

This work was performed under the UK National Measurement System programme for Optoelectronics. We are grateful to G. Barwood for providing the Rb-stabilised laser system and for help in its use, to W. R. C. Rowley for advice, and to M. Milton for loan of a 1.56 μm diode laser system.

REFERENCES

1. 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 μm based on ^{20}Ne and at 1.56 μm derived from the 0.780 μm 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).

2. G. P. Barwood, P. Gill and W. R. C. Rowley, "Optically Narrowed Rb-stabilised GaAlAs diode laser frequency standards with 1.5 x 10^{-10} absolute accuracy." In Y. C. Chung, Ed., *Frequency-stabilized lasers and their applications, Boston, Massachusetts, 16-18 November, 1992*, Proc. SPIE, **1837**, 262-270 (1993).

3. C. R. Pollock, F. R. Petersen, D. A. Jennings and J. S. Wells, "Absolute frequency measurements of the 2-0 band of CO at 2.3 μm; calibration standard frequencies from high resolution colour centre laser spectroscopy." J. Mol. Spectrosc., **99** 357-368 (1983).

4. A. G. Maki and J. S. Wells, *Wavenumber calibration tables from heterodyne frequency measurement.* Washington D.C.: US Government Printing Office, NIST Special Publication 821, 1991.

5. J. S. Wells and A. G. Maki, private communication, October 1992.

6. W. Wang, K. Nakamura, Y. Toda and M. Ohtsu, "1.5 μm diode-laser based nonlinear frequency conversions by using potassium titanyl phosphate." *Appl. Phys. Lett.*, **61**(16) 1886-1888 October 19 (1992).

7. H. Sasada and K. Yamada, "The calibration lines of HCN in the 1.5 μm region." *Appl. Optics* **29**(24), 3535-3547, 20 August (1990).

8. BIPM "Documents concerning the new definition of the metre." Metrologia **19**, 163-177 (1984). Note: at the 1992 Consultative Committee on the Definition of the Metre, much improved frequency values were adopted for the *mise en pratique*, notably for the 633 nm I_2-stabilised HeNe laser. For details see: - T. J. Quinn, "*Mise en pratique* of the definition of the metre 1992." *Metrologia,* **30**, 523-541 (1993/94).

9. D. J. E. Knight, K. I. Pharaoh, S. L. T. Drysdale, J. Meldau, G. Duxbury and C. J. Hodges, "Progress towards laboratory frequency standards for the optical communication bands." In *Proc. 6th European Frequency and Time Forum, Noordwijk, NL, 17-19 March 1992.* Noordwijk: ESA SP-340, June 1992. Pp. 531-536.

10. H. Sasada and O. Kubota, "Lamb dip-stabilized 1.52 μm He-Ne lasers." *Appl. Phys.*, **B55**(2), 186-188 August (1992). Note: the accurate lambda-meter measurements are described in:- J. Ishikawa and H. Watanabe, "A fringe-counting wavemeter for infrared laser diodes." *IEEE Trans. Instrum. Meas.*, **42**(2), 423-425 April (1993).

TIME APPLICATIONS

PHASE STABILITY OF XR5 GPS-ECEIVERS AT VTT AND UNIVERSITY OF OULU

J. Mannermaa, K. Kalliomäki, T. Mansten

University of Oulu, Department of Electrical Engineering,
FIN-90570 Oulu, Finland

Abstract

Simultaneous measurements of the phase of three Navstar GPS receivers (XR4 and XR5) were carried out at Technical Research Centre of Finland (VTT) and University of Oulu. Several 35 minutes long data packets were collected during one day according to common schedule. Carefully controlled rubidium clocks were used as referencies. Power spectral densities and Allan variances were calculated. The phase and frequency spectra for every data packet was computed and the simultaneous results of similar receivers were compared to each other.
The results reveal clearly the effect of SA (Selective availability) on power spectra. Phase spectra seem to be white and high up to 0.2 cpm (cycles per minute) and then the density decreases following random walk process.

Introduction

The official time and frequency in Finland is based on GPS-receivers. The rubidium clocks are used to smooth short term noise. The first GPS-satellites (block 1) transferred accurate time and position but the performance of new satellites (block 2) was deteriorated intentionally with SA (Selective Availability) to prevent too accurate navigation of civilian users.
To improve time/frequency reception, one has to know the properties on SA noise. The shape of the spectrum and the upper and lower limiting frequency of SA noise are important. When the properties of SA noise are known, effective filtering algoritms can be developed to improve the accuracy of time.
Simultaneous measurements of GPS second ticks were carried out at two laboratories 600 km apart by using rubidium clocks as short term referencies.

Instrumentation

Measurements of the phases of second tick of three Navstar-GPS receivers were carried out at VTT (Technical Research Centre of Finland, Helsinki) and University of Oulu. At both places similar computer controlled surveillancy system were used.
The studied receivers were one two-channel XR4 (Oulu) and two six-channel XR5:s (Oulu and VTT). The measurement resolution was 1 ns and the interval 1 s when the data for the Allan variances was recorded and 2 s for other cases. Carefully controlled rubidium clocks were used as referencies at both places.

Data handling

First order linear regression model was applied to phase readings of every data packet to remove the frequency error of rubidium reference clock. Six 35 minutes long data packets of phase values were collected during one day according to common schedule. The data was processed both in time and frequency domain. At first, the frequency was calculated and then the regression analysis was applied to phases to know the noise power of signals. Moreover, the Allan variances (TVAR) of measurements were computed to find out the stability of phase. In frequency domain, the power spectral densities of phases and filtered frequencies were calculated by using the algorithm of the Maximum Entropy Method (MEM). That is due to the ability of the MEM to smooth the final results by a suitable way whereupon the essential details are easily seen.

Results and discussion

As a measure of the stability of the studied XR5s both the regression analysis and the Allan variance of the phase was used.
At first, the regression analysis was applied to the data set of each measurement and the slope and residuals were calculated. The results are presented in Table 1.

Table 1. The phase residuals and slopes of GPS sessions at VTT and Oulu.

Session N:o	VTT / Automation		University of Oulu	
	Residuals µs	Slope exp-12	Residuals µs	Slope exp-12
01	0.035	-8.2	0.114	-63.0
02	0.027	22.0	0.121	29.0
03	0.071	16.0	0.072	9.9
04	0.049	-1.8	0.108	28.0
05	0.052	-16.0	0.102	-5.3
06	0.045	-11.0	0.069	1.9

Neither the residuals nor the slopes between results obtained at VTT and Oulu seem to correlate. Probably the reason is the different operating mode.

Secondly, the Allan variances of the phases are presented in Fig. 1.

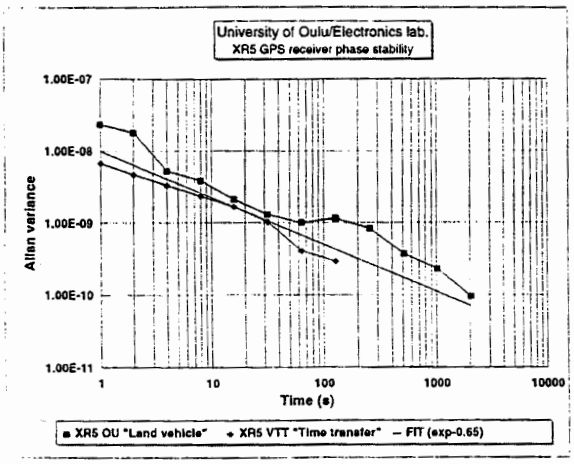


Fig. 1. The Allan variances of the phases of three GPS receivers.

A fitted line describes the average slope of the variance curves being about -0.65. This means that phase stability improves by about two decade per three decades of smoothing. As expected, in the time transfer mode the noise power of a GPS receiver is less than in the "land vehicle" mode when a receiver concentrates to coordinates.

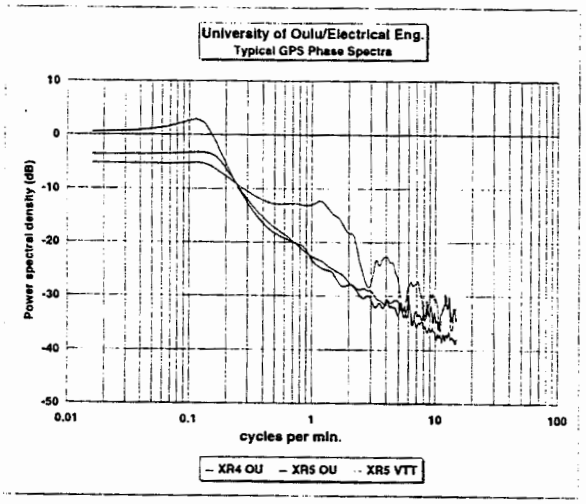


Fig. 2. The power spectra of all phase values.

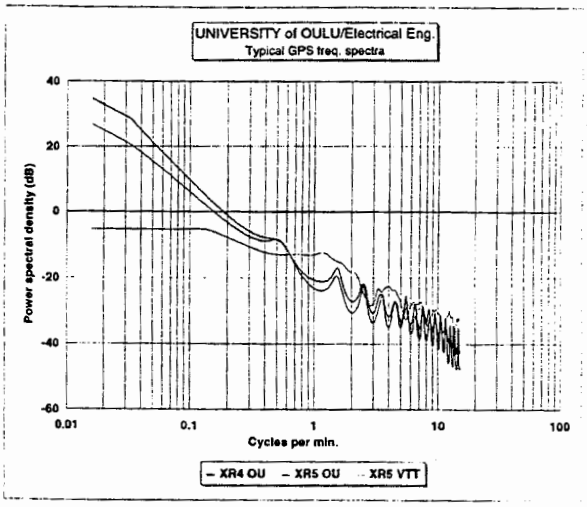


Fig. 3. The power spectra of all frequency values.

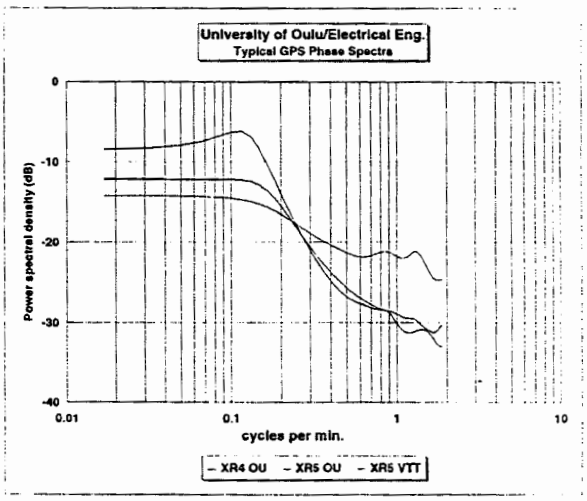


Fig. 4. The power spectra calculated from phase values taken in every eight second.

Thirdly, the power spectral densities were calculated and the average spectra of phases and frequencies are presented in Fig. 2 and 3 (computed by all values), respectively and in Fig. 4 and 5 (computed by every fourth value), respectively.

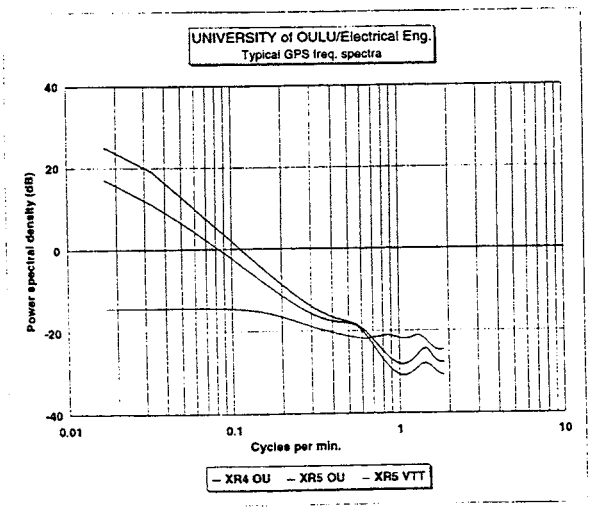


Fig. 5. The power spectra calculated from frequency values taken in every eight second.

Phase noise spectra are flat (white noise like) and the noise level are high up 0.2 cycles per min. Then a sharp corner occurs in all cases. This corner is not visible in frequency noise spectra. Then the spectrum decreases proportional to 1/f^2 (random walk process) to about 1 cpm. After that frequency the slope is gentle. The spectrum can be explained by slowly variated and finite changes (jumps) in the position data send to each satellite.

There seem to significant differences between noise levels of tested receivers. This can be explained by different operating modes of operation when XR5 VTT works on "time transfer" mode whereas XR5 OU works on "land vehicle" mode.

The correlation control method of GPS-receiver seems to dominate in phase and frequency spectra as a first approximation. 50 ns phase correction jumps and loosely coupled frequency control loop is typical for XR5. The control method is different in XR4. It seems, that frequency synthesized of XR4 jumps between two fixed values and the duty cycle is varied. In addition, the crystal frequency is slowly controlled.

The development of a prototypic time synchronized speaking clock

C.C. LIN, C. S. LIAO, J. L. SHI, C. M. SHEU, and H. D. LIN

Telecommunication Laboratories, MOTC
No. 12, Lane 551, Section 3, Mintsu Rd., Yang-Mei, Tao-Yuan, Taiwan 326, ROC
TEL : +886-3-424-4441 FAX : +886-3-424-5474

Abstract

The time signal accuracy of the current speaking clock in Taiwan, ROC was insufficient due to indirectly and manually traced to our national standard time. An industrial computer based prototypic time synchronized speaking clock(P-TSSC) was developed by Telecommunication Laboratories(TL) at ROC for which used 117 public voice time service station that could provide voice signal of accurate time to generic users. The P-TSSC traced to the national time standard UTC(TL) by using the Inter Range Instrumentation Group(IRIG) code via a dedicate line. The P-TSSCcould provide a substitute source of time and overcome the limitations of high frequency(HF) radio receiver. The P-TSSC software programming, system architecture, and hardware configuration including time code synchronized generator unit, voice data library unit, voices generate unit and industrial computer controller unit were discussed

Introduction

The conventional public voice time service is managed by the local telephone company, which gets the time signal through our BSF high frequency time broadcasting station and distributes to the public through the PSTN. A telephone number "117" is designated for this service. The timing accuracy limitations of 117 station are based on the high frequency radio time signal reception and some human factors. A computer based prototypic time synchronized speaking clock(P-TSSC) is developed for enhancing the time announcement accuracy and is continued to provide a good service for the generic users. The P-TSSC is operated in synchronism with UTC(TL) which is provided by the National Standard Frequency and Time Section in Transmission Technology Laboratory, Telecommunication Laboratories (TL), Directorate General of Communications (DGT), Ministry of Transportation and Communications (MOTC), TAIWAN, ROC. The UTC(TL), coordinated universal time kept by TL, has been traced to Bureau International des Poids et Mesures(BIPM) since 1972. This P-TSSC provided the general functions of a speaking clock, the synchronized time display and status monitor. The propagation delay compensation is also reserved for future need. Time synchronizing method using an IRIG-B time code signal transmitted from TL to 117 station will be introduced. The P-TSSC hardware is constituted of an IRIG-B time code synchronized generator unit, a voice data library unit, a voices generate unit and an industrial computer controller unit. A user-friendly software which includes system polling among controller and interfaces, analog local time display, analog world time display of six cities, digital local time display, voice announcement toggle switch, and propagation delay compensation is also provided.

Synchronizing signal from TL to 117 station

The time signal of P-TSSC was traced and synchronized to UTC(TL) via a prevail line by using an IRIG code. The distance between TL and 117 station is about 40 km. The propagation delay was measured precisely(by reflecting the signal and dividing by two) and compensated. The measured time correlation between TL and 117 station was better than 10 microseconds after compensation. The test data was shown in figure 2.

P-TSSC hardware design

The P-TSSC hardware consisted of IRIG-B time code synchronized generator unit, voice data library unit, voices generate unit and industrial computer controller unit as shown in figure 3. The block diagram of IRIG-B time code synchronized generator unit was illustrated in figure 4. The SYNC detectors checked the phase difference between the standard IRIG-B code carrier and internal oscillator. A microprocessor accepted this phase difference and adjusted oven controlled oscillator via the digital to analog (D/A) converter to synchronize the frequency of the internal oscillator with external signal from TL. The synchronizing process was complete within few minutes. The propagation delay from TL to 117 station was compensated by scalar. The resolution of this scalar was about 1 microsecond. The code generator was for generation of new IRIG-B time code. Other time information which applied for secondary users was also created.

The block diagram of voice data library unit was shown in figure 5. This unit consisted of a computer

interface, data & address buses, SRAM(Static Random Access Memory) disk, battery, power indicator, address decoder, and address switch. SRAM disk was used for providing high reliability when the poor environment condition. The on board battery could be utilized for more than ten year's. The voices generate unit as shown in figure 6, consisted of a speeches processing element, a speech compressor, an A/D converter, a D/A converter, and a telephone interface. The digital voice data was compressed by the compressor after the A/D converter and to speech processor element forms the digital data for computer. Announcing voices, the process direction reversed as from speech processor element to compressor and sequently to D/A converter as the voice announcement to telephone interface. A 80386 industrial personal computer was necessary for the controller unit to handshake with other units and to process voice's combination. A watch-dog timer prevented the computer in trouble last for 1.5 seconds and restarts the system

P-TSSC software design

The P-TSSC software was complied by Borland C++ 3.0 and was a user-friendly design. The software structure was the top-down modularized structure design. Figure 7 was the working status of the program. When the P-TSSC was started, all the initial default values were set and function explanations were shown on the screen. Users can select the desired function using keyboard and then the screen change to the setting values. If users did not select any item, the P-TSSC would decode the external standard IRIG-B time code, calibrated the system clock and prepared the voice announcement service.

The P-TSSC was traced to UTC(TL) automatically when power was connected. The screen display divided into 3-areas as shown in figure 8. Those were status area, time display area, and command area. The status area shows the external standard IRIG-B signal status, synchronization status, and voice announcement status. If an anomaly condition happen, the status alarm would be highlighted. The analog time display, digital time display, and international city time display were shown in the time display area. The command area was located at lower part of the screen. There were 6 function keys shown in the command area. The first one was for the analog time display. Whenever this button be pressed, figure 9 would appear on the screen The second one was the international 6 cities' time display button. Whenever this button be pressed, the screen showed as figure 10. The third one was the digital time display and the screen displayed as figure 11. The forth was the toggle switch for voice announcement. The announcement repeated every ten second. The fifth bottom was for the delay compensation. Users could enter the propagation delay by pressing this bottom and follow the dialog indication. The last one was the "Quit" button for leaving the program back to DOS system.

Conclusion

The P-TSSC was an accurate speaking clock that automatically synchronizes the external standard time signal and whose time could be captured, read, and fed to other required system. It enhanced the time announcement accuracy, provided a good service, and avoided the traditional mechanical operation and human errors. It was the first step of our telecommunication time dissemination program. We would continue to provide other time services for the public and reach the goal to synchronize the national time of our country.

Reference

[1] "Time Code Instrumentation" Systron-Donner, Inc..

[2] David C. Robinson, "Using Bus Level IRIG Time Code Translators to Time Tag Data and Synchronize Multiple Processing Stations", Bancomm division of Datum Inc..

[3] "Industrial PC Products User's Manual", ADVAN TECH Technology Inc., 1993.

[4] Shelby Bass, Roger Olson and Jack Mcnabb, "Disciplined Range Time Code Translators with Sub-Microsecond Accuracy," TRAK Microwave Corp.,1988.

[5] "PCD-892 Flash/RAM/ROM Disk Card User's Manual", ADVANTECH Technology Inc., 1992.

[6] "V-Link PC Voice Board User's Manual", Eletech Electronics Inc., 1992.

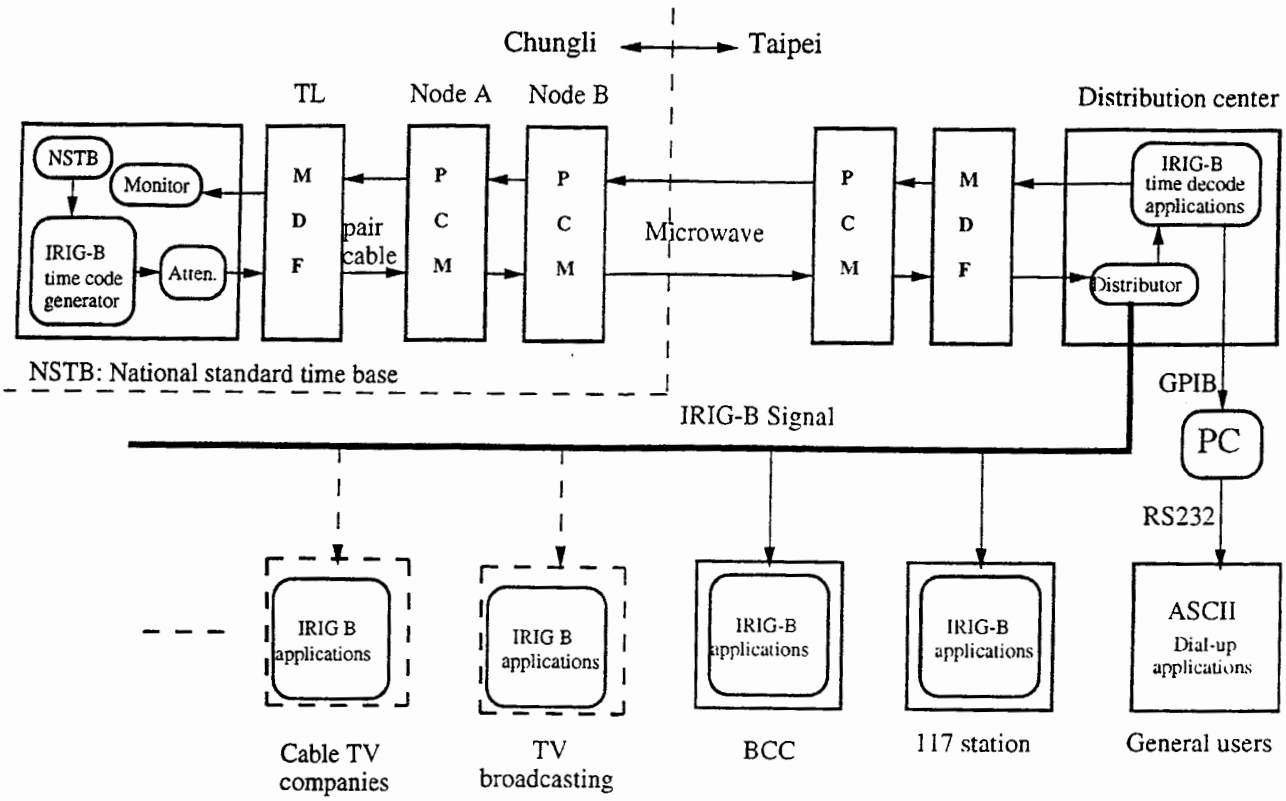


Figure 1 Transimition route of P-TSSC

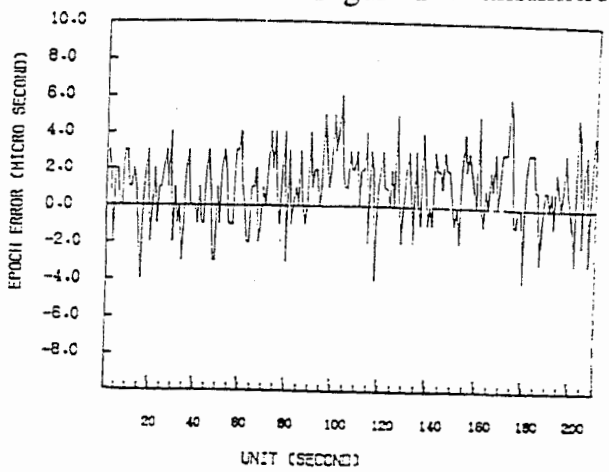


Figure 2 Measured time correlation

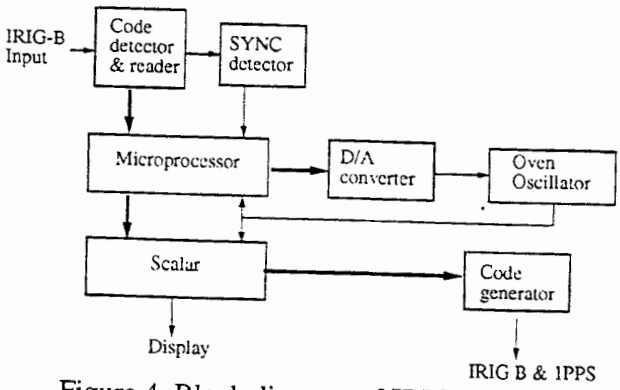


Figure 4 Block diagram of IRIG-B time code synchronization generator unit

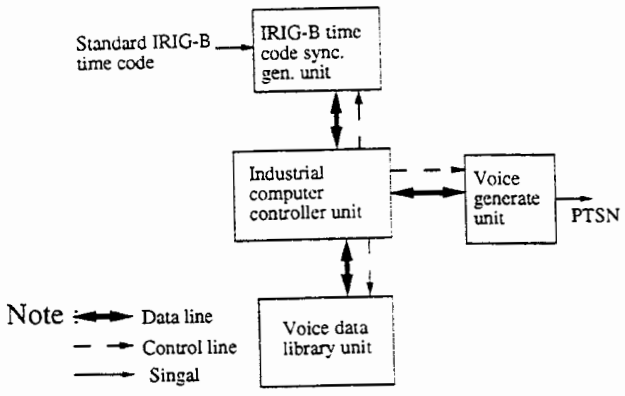


Figure 3 Block diagram of P-TSSC hardware

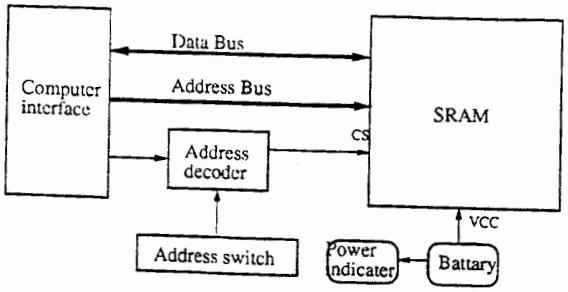


Figure 5 Block diagram of voice data library unit

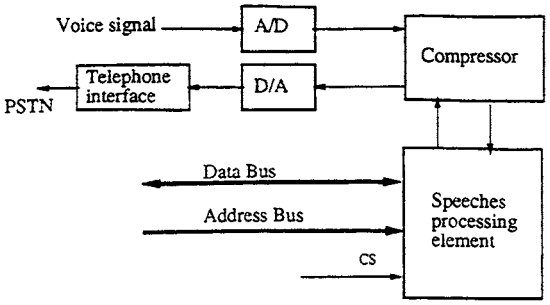


Figure 6 Block diagram of voices generate unit

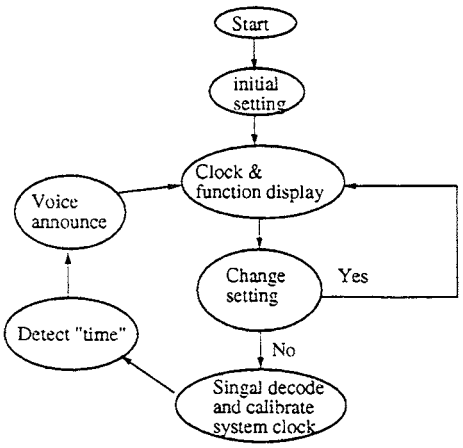


Figure 7 working status of the program

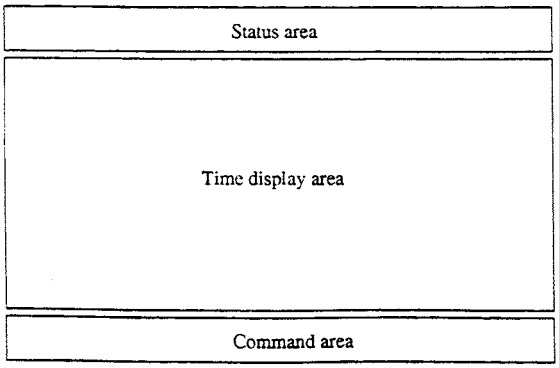


Figure 8 The screen display

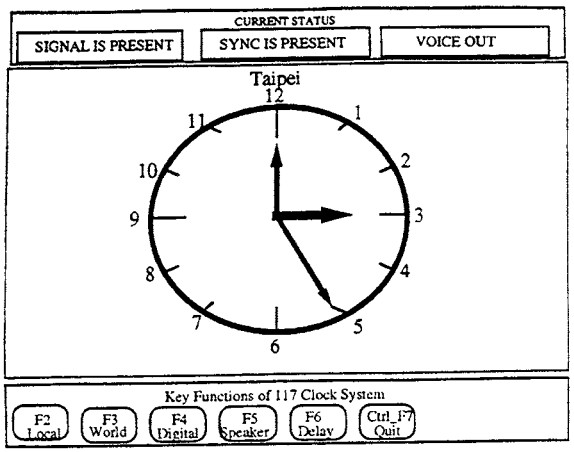


Figure 9 Anolog time display

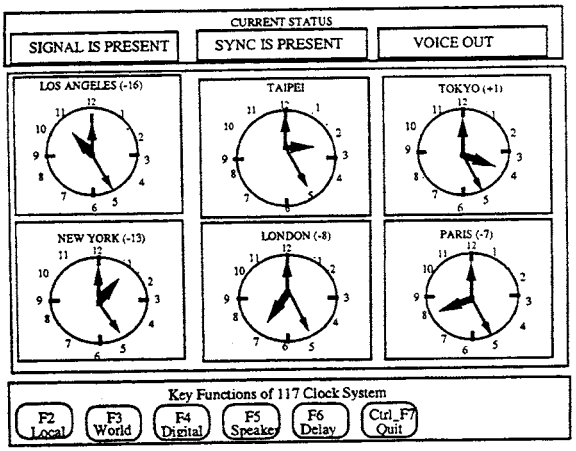


Figure 10 International 6 cities' time display

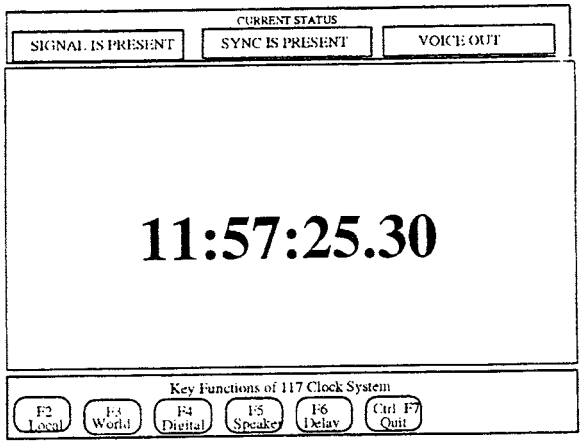


Figure 11 Digital time display

DIGITAL CORRELATOR FOR TRACKING PSEUDO-RANDOM CODED SIGNALS

KHALIFA Nabil, BAZZI Oussama, TORGUET Roger

IEMN Department OAE (UMR CNRS 9929), University of Valenciennes, Le Mont Houy
59304 Valenciennes

ABSTRACT

A digital system for tracking a pseudo-random signal is presented. It is based on an improved sequential search technique coupled to a time compression hardware.

The system proposed uses three digital correlators which were implemented using a field programmable gate array circuit.

The technique was evaluated by tracking GPS signals. It may well be extended to general spread spectrum applications.

I - INTRODUCTION

Recently, telecommunications have greatly advanced because of the development of new techniques like spread spectrum signaling.

A brief overview of spread spectrum transmission techniques is first given. The architecture of a digital correlator based receiver is presented and applied to the GPS "Global Positionning System" application. One of the main drawbacks of the spectrum spreading is the signal to noise ratio degradation at the receiver input, due to the fact that noise (supposed to be white) is proportionnal to the transmission channel band width. This requires a matched filter in the receiver path as well as a signal delay tracking to overcome the Doppler phenomenon [9].

Our proposed system uses a digital correlator allowing signal detection. It comprises a delay locked loop unit which tracks the phase shift of the received code relative to the reference one (synchronization process).

The system is realized with three parallel correlators performing correlations with respectively the reference code and one left and one right half a bit shifted versions of this code.

Phase alignement of later received reference codes is realized using a sequencer with an NCO (Numerically controlled oscillator) this allows to track continuously the received code phase in the presence of Doppler.

Results obtained with our system realized using an FPGA (Xilinx^{TM} circuit XC 3090 p c 84) are presented.

II - SPREAD SPECTRUM SIGNALING TECHNIQUE

In conventional telecommunication frequency filtering of the carrier is needed in order to select a given emission with a receiver.

However in spread spectrum signaling [1] telecommunication, this selection is done by correlation of a pseudo-random sequence « PN », which is modulated by useful data transmitted , [2]with a local replica one (fig. 1).

The frequency of the pseudo random sequence is chosen much greater than the data one, for example in the global positioning system « GPS » data is at 50 Hz while PN frequency is at 1.023 MHZ yielding a spectrum spreaded by a factor of 20,000.

The main advantages of this technique [3] are :

* protection against intentional jamming or accidental interference
* to have a digital transmission
* multiple access without synchronization by using different PN codes
* privacy
* multipath mitigation
* accurate ranging.

Pseudo Random Sequence features :

A pseudonoise sequence can be generated from a n-stage shift register in which the output of particular stages are fed back to the input register via a modulo-two adder [8]. This generator delivers a periodic binary sequence with a maximal length or pseudonoise (PN) sequence of $N = 2^n - 1$ which is composed by 2^{n-1} ones and $2^{n-1} - 1$ zeros.

The name pseudonoise comes from the fact that the correlation property [7] of PN sequences resembles to the white noise one. If we define the ± 1 valued sequence $C'_n = 1 - 2\,C_n$, $C_n = 0,1$ then the autocorrelation function

$$Rc'(\tau) = \frac{1}{N} \sum_{K=1}^{N} c'_K\, c'_{K+\tau}$$ is given by

$$Rc'(\tau) \begin{cases} 1, & \tau = 0,\ N,\ 2N \\ \dfrac{-1}{N} & \text{otherwise} \end{cases}$$

where $N = 2^n - 1$ (fig. 3). There is an isomorphism between ({0,1}, +) and ({1,-1}, x) [6].

Therefore $C_k + C_{k+z} \leftrightarrow C'_k C'_{k+z}$, digital correlation becomes easy to perform

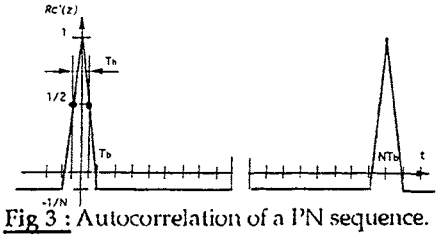


Fig 3 : Autocorrelation of a PN sequence.

IV - PROPOSED SYNCHRONIZATION PROCESS

Synchronization plays a major role in the spread spectrum system receiver, so this process must be accurate.

To achieve synchronization between the spreading code and despreading one we need two steps : code acquisition and tracking (fig 2).

1 - Acquisition process

Considerable research efforts have been devoted to the acquisition problem particularly, where low signal to noise ratio environments in presence of strong interfering signals or deep noise are involved. There are three main techniques for acquisition :

* The serial search technique needs a complex algorithm to choose the search strategy [4], [5],the parallel search technique which requires a more complex hardware, finally sequential techniques [6] have the merit to be easy to implement, however acquisition time is longer than that needed for the parallel search technique. This later technique was chosen in our proposed system and was coupled to a time compression hardware in order to overcome the drawbacks concerning the acquisition time .

* Time compression uses memory banks. It consists in parallel write-read cycles with a read clock much faster than the write one corresponding to the in-coming signal bit rate. This allows a faster processing by a factor equal to the time compression one.

2 - Tracking process

Numerous sources contribute to disturb the acquisition process :

a - fading channels and interference,
b - the unknown incoming carrier phase in the noncoherent receiver,
c - additive white gaussian noise,
d - Doppler phenomena due to the relative motion between transmitter and receiver.

Once acquisition is achieved, the Doppler effects perturbs synchronization of the pseudorandom code modulated with data (at base band level) in two possible ways :

* first, presence of Doppler causes bit sliding (fig. 4a)

* second, Doppler modulates the received code so this later will be reversed at positions fixed by the unknown Doppler frequency)(fig 4.b).

To overcome the first problem we propose a tracking system based on three parallel correlators performing respectively operations between received code and

a) the reference code
b) one earlier half a bit replica of reference code
c) one later half a bit replica of reference code.

The system will compare simultaneously the three correlation results in order to follow the relative phase sliding direction and deduces whether to accelerate or to decelerate the reference code by one chip at the end of the correlation cycle.

This was applied to a GPS [9] receiver system where slow tracking is needed due to the fact that the maximal Doppler shift of the signal base band center frequency (1.023 MHz) is around 6 Hz due to the satellite motion (if we neglect the receiver speed) . This leads to a sliding of one chip per 200 ms (200 code cycles).

Now let us observe what happens when central signal correlation starts to fall down. As shown in (fig 6.a) Rc' (0) is decreasing and only one among Rc' (-Tb) and Rc' (+Tb) is increasing (Rc'(-Tb) if it's a left sliding or Rc'(+Tb) if it is a right sliding). When Rc'(0) = Rc'(-Tb) = 1/2 or Rc'(0) = Rc' (+Tb) = 1/2 the state machine centers the reference code by way the centring mechanism shown in (fig. 5) centers the reference code. Once this is done Rc'(0) will rise again (relative to the Doppler) to reach the maximum level and so on. This fluctuation causes a 6 dB maximum attenuation which may cause desynchronization when the signal to noise ratio after despreading is low (such as GPS signal at setting or rising satellites).

To improve this processing we should reduce the fuctuation. One solution consists to do the same operations applied to half a bit offset (instead of one bit) between the three correlations. So : Rc'(Tb/2) Rc'(0) ; Rc' (+Tb/2). The maximum attenuation is then less than 3 dB as shown in (fig. 6b). The ideal solution for this system would be to operate on Tb/n where n >>1, but in this case, realization become very hard.

Once this is achieved the Doppler envelope which affects the correlation levels will be easily removed using a classic RF phase locked loop oscillator driven by the filtered correlation signal.

V - RESULTS

GPS signals which are received burried in noise with signals to noise ratio ranging from -20 to -10 db were tracked with our proposed system. The central correlator results are shown in figure 8 for different GPS Doppler shifts were they are fully rectified..

VI - CONCLUSION

An acquisition and tracking digital system for direct spread-spectrum applications is proposed. Among advantages are simple implementation and flexibility as well as evolutive features allowing real-time processing using a time compression technique. Results obtained with GPS signals are finally reported.

REFRENCES

[1] Cook and Marsh « An introduction to spread spectrum », IEEE Commun. Mag, pp. 8-16, Mar. 1983.

[2] Pichholtz, Schilling and Milstein, « Theory of spread spectrum communications - A tutorial IEEE Trans. Commun., pp. 855- 883, May 1982.

[3] Simon, et al., Spread Spectrum Communication. Rockville MD : Comput. Sci. Press, 1985.

[4] G.F. Sage., « Serial synchronization of pseudo noise systems » IEEE Trans. Commun Technol, Vol. COM-12, pp. 123-127, Dec. 1964

[5] A. Weinberb, « search strategy effects on PN acquisition performance », in Proc. Nat. Telecommun. Conf., New Orleans, LA, Dec.1981, pp. F1.5.1 - F1.5.5.

[6] R.B. Ward and K.P. Yiu, « Acquisition of pseudonoise signals by recursion-aided sequential estimation », « IEEE Trans. Commun., vol. COM-25, pp. 784-794, Aug. 1977.

[7] L.V. Sarwate and M.B. Pursley, « Cross-correlation properties of pseudo-noise and related sequences », Proc. IEEE ; vol 68 PP. 597-619, May 1980.

[8] O. Bazzi, « Processeur acousto-optique adapté au décodage des signaux GPS de télélocalisation», Thèse de Doctorat, Univ. de Valenciennes, 1992.

[9] SPILKER-GPS Signal Structure and Performance Navigation-US Institute of Navigation Vol 25 Summer 1978.

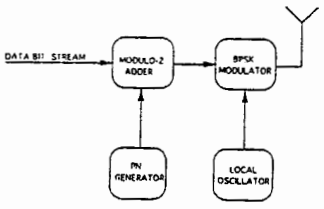


Fig 1.a

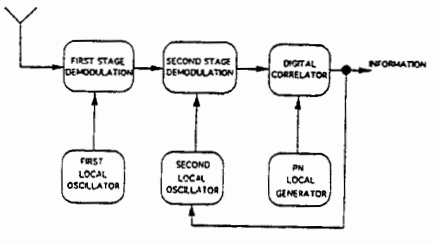


Fig 1.b

Fig 1: A Spread Spectrum Communication system
a) Tranceiver b) Receiver

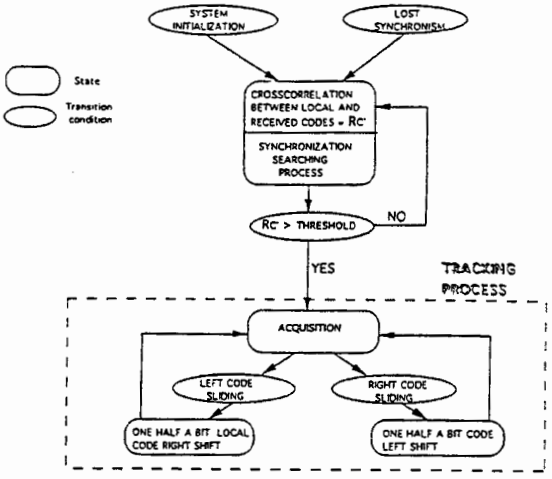


fig 2 : State diagram for acquisition and tracking process

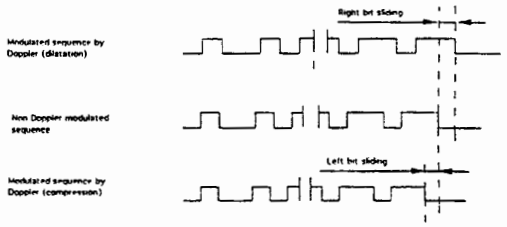


Fig 4.a: Bit sequence sliding due to the Doppler

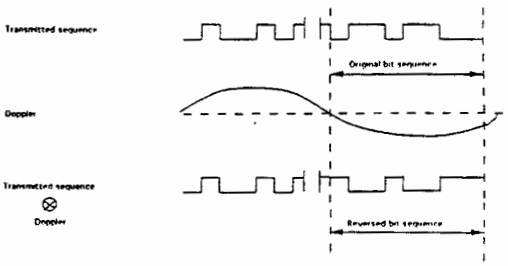


Fig 4.b: Modulated sequence due to the Doppler

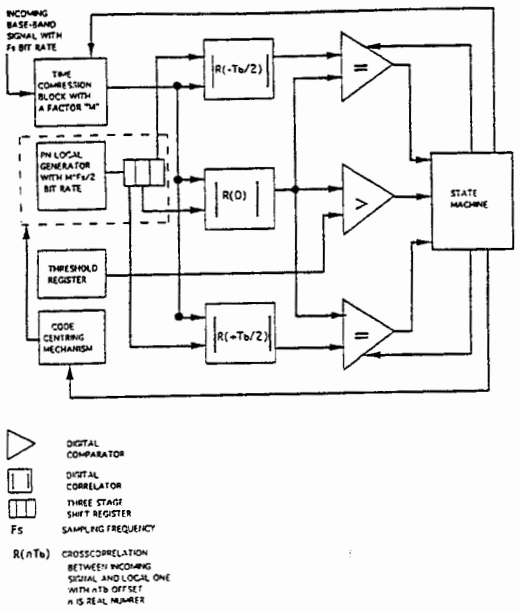


Fig 5 : Digital Correlator architecture

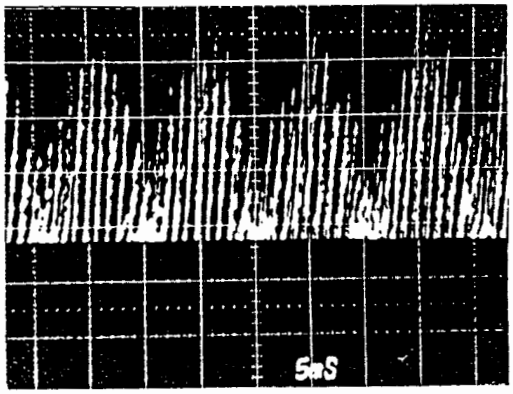
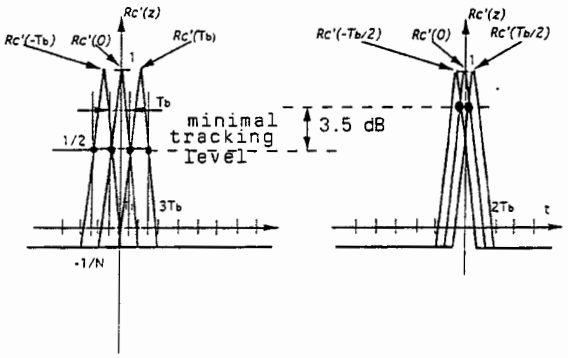


Fig 7 : Received GPS signals using the proposed correlator



a) b)

Fig 6 : Correlation results for different shifts of reference code

a - With one-bit shifts

b - With one half a bit shifts allowing 3.5 dB enhancement in the tracking level

THE ACCURACY OF LORAN-C BASED TIME SCALES IN FINLAND DURING 1991-1994

K.Kalliomäki [1], T.Mansten [2]

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

ABSTRACT

The European Loran-C network has been transferred to European hands at the beginning of 1995. The time accuracy of the transmissions will be enhanced, which may add interest in the maintenance of time scales via Loran-C transmissions in Europe when the new NELS organization restarts the transmissions. VTT has maintained secondary time scales for several years by using Norvegian Sea Sylt transmitter in "common view" measurements and international time information letters from PTB, SNT and USNO. The results show that a Loran-C based time scale could be maintained at ± 100 ns/10 day uncertainty level via PTB and at ± 150 ns/10 day uncertainty level via SNT. USNO reports of master (Ejde) contain some element unknown to us which decreases their usefulness.

1. INTRODUCTION

The Norwegian Sea Loran-C network has been transferred to European hands at the beginning of 1995. At the moment the network is not in operation. The master-slave stations are to be reorganized and their GRI's changed. The time accuracy of the transmissions will be enhanced, which may add interest in the maintenance of time scales via Loran-C transmissions in Europe when the new NELS organization restarts the transmissions. Sylt transmitter will be a master in the new organization, therefore our measurements from the past few years may be of interest to Loran-C time users.

The phase of the Norwegian Sea Loran-C Sylt transmitter has been monitored and reported by several time and frequency laboratories, for instance PTB and SNT. The master station, Ejde, is situated on Faeroe Islands and has been monitored and reported by USNO.

The VTT Automation situates near Helsinki. It contains, amongst others, the National Standards Laboratory of Time and Frequency (VTT). VTT has no Cs clocks. The time and frequency of VTT is based on USNO time via two GPS receivers. The time and frequency are traceable to to BIPM via atomic clock transports with SU and SNT and via time reports of BIPM. VTT monitored Loran-C Sylt transmitter as a secondary means of traceability and compared the results with those of SNT and PTB. The Ejde measurement results of USNO were used, too.

Swedish National Time and Frequency Laboratory (SNT) maintains the time and frequency of Sweden. The reference time scale at SNT is based on one of three Cs clocks which contribute to to UTC. SNT has monitored the phase of Loran-C Sylt and distributed the results as time information letters to VTT.

The positions of transmitters and receivers concerned are in the map, Fig. 1. The corresponding distances are in Table 1.

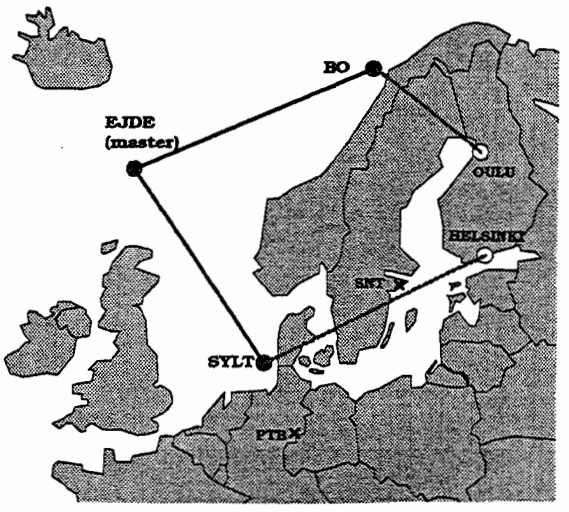


Fig. 1: Norwegian Sea chain Loran-C transmitters and receivers

Table 1: Transmitter-receiver distances

Transmitter	Receiver	Distance, km
Bo	Oulu Univ.	600
Sylt	PTB	330
Sylt	SNT	770
Sylt	VTT	1155

2. MEASUREMENTS, TIME INFORMATION LETTERS

VTT measures the time difference Δ_1 between Finnish second tick UTC(VTT) and the tick produced by our Loran-C receiver

$$\Delta_1 = UTC(VTT) - UTC(Rx_{VTT}Sylt) \quad (1)$$

PTB measures the time difference Δ_2 between UTC(PTB) and the tick from PTB Loran-C receiver

$$\Delta_2 = UTC(PTB) - UTC(Rx_{PTB}Sylt) \quad (2)$$

SNT measures the time difference Δ_3 between UTC(SNT) and the tick from SNT Loran-C receiver

$$\Delta_3 = UTC(SNT) - UTC(Rx_{SNT}Sylt) \quad (3)$$

USNO measures the time difference Δ_4 between UTC(USNO) and the tick from SNT Loran-C receiver

$$\Delta_4 = UTC(USNO) - UTC(Rx_{USNO}Ejde) \quad (4)$$

From Eqs. (1)-(4) we get the three differences

$$\Delta_2 - \Delta_1 = UTC(PTB) - UTC(VTT) - [UTC(Rx_{PTB}Sylt) - UTC(Rx_{VTT}Sylt)] \quad (5)$$

$$\Delta_3 - \Delta_1 = UTC(SNT) - UTC(VTT) - [UTC(Rx_{SNT}Sylt) - UTC(Rx_{VTT}Sylt)] \quad (6)$$

$$\Delta_4 - \Delta_1 = UTC(USNO) - UTC(VTT) - [UTC(Rx_{USNO}Ejde) - UTC(Rx_{VTT}Sylt)] \quad (7)$$

The receiver tick differences $UTC(Rx_{VTT}Sylt) - UTC(Rx_i)$ contain some propagation delay variations and receiver adjustments, but can be treated as constants c1, c2, c3.

$$\begin{aligned} c_1 &= UTC(Rx_{PTB}Sylt) - UTC(Rx_{VTT}Sylt) \\ c_2 &= UTC(Rx_{SNT}Sylt) - UTC(Rx_{VTT}Sylt) \\ c_3 &= UTC(Rx_{USNO}Ejde) - UTC(Rx_{VTT}Sylt) \end{aligned} \quad (8)$$

The differences (5), (6) and (7) are averaged in partly overlapping 11 day groups around 9-ending MJD's.

BIPM publishes "Circular T" time information letters, from which the time differences between UTC and different laboratories are obtained for 9-ending MJD's.

$$\begin{aligned} \Delta_5 &= UTC - UTC(PTB) \\ \Delta_6 &= UTC - UTC(SNT) \\ \Delta_7 &= UTC - UTC(USNO) \end{aligned} \quad (9)$$

By combining these results with Eqs. (5), (6), (7) and (8) we get three time scales for UTC-UTC(VTT) for 9-ending MJD's

$$\begin{aligned} [UTC - UTC(VTT)]_1 &= \Delta_5 + (\Delta_2 - \Delta_1) + k_1 \\ [UTC - UTC(VTT)]_2 &= \Delta_6 + (\Delta_3 - \Delta_1) + k_2 \\ [UTC - UTC(VTT)]_3 &= \Delta_7 + (\Delta_4 - \Delta_1) + k_3 \end{aligned} \quad (10)$$

The three arbitrary constants k_1, k_2, k_3 contain the constants c_1, c_2 and c_3. They were chosen and fixed at the beginning of the comparison period so that the average of the final results is near zero. The used values were $k_1 = -1.4$ µs, $k_2 = 0.15$ µs, $k_3 = 0.75$ µs.

3. RESULTS

The Sylt measurement data of PTB, SNT, USNO and VTT from Sept. 5, 1991 to the end of 1994, Eqs. (1), (2), (3) and (4), is in Fig. 2.

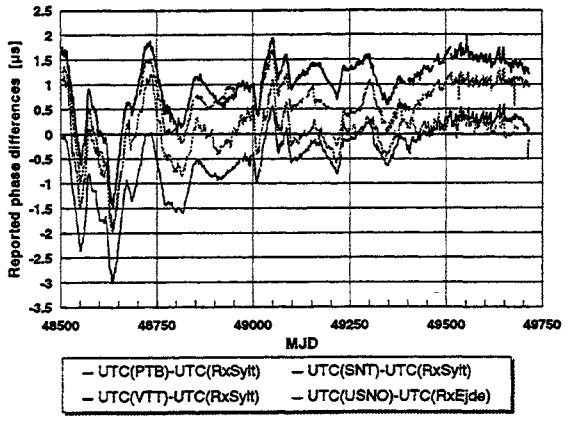


Fig. 2: Sylt/Ejde measurement results

In Fig. 3 are the three differences, Eqs. (5), (6), (7).

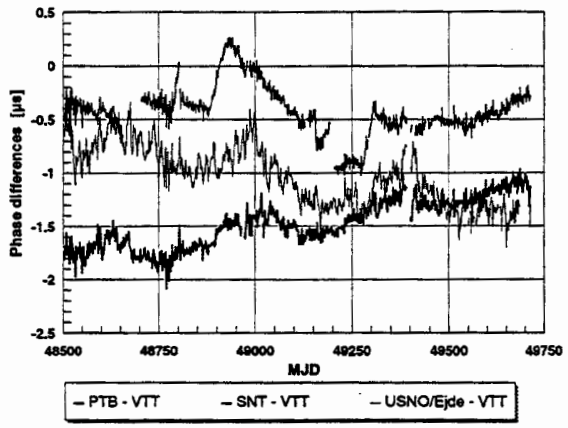


Fig. 3: The differences between Loran-C measurement results

Fig. 4 represents the differences of UTC and laboratory time scales obtained from BIPM "Circular T" reports, Eq. (9).

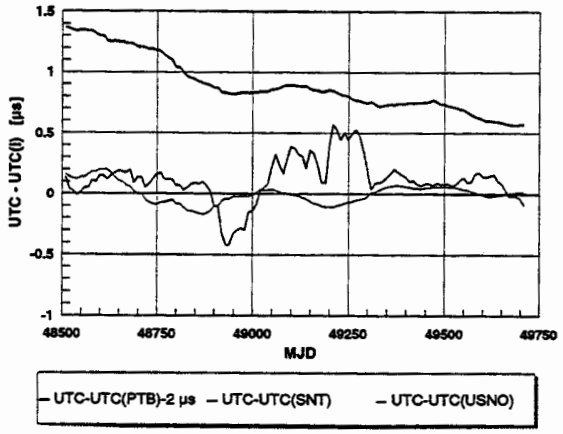


Fig. 4: UTC - UTC(i) from BIPM "Circular T" time information letters

In Figs. 5, 6 and 7 are the three time scales for VTT, Eq. (10), obtained from the common view measurements. The results obtained via PTB provide the best time scale. The time scale obtained via SNT, Sweden is fair, but contains one jump because of the SNT receiver repair and another jump, probably because of some receiver band adjustment. The third time scale obtained via USNO measurements of Ejde is not usable because of eminent drift and high noise.

The means and standard deviations of the results are in Table 2. The receiver jumps were removed from the time scale via SNT for the SD of time scale 2.

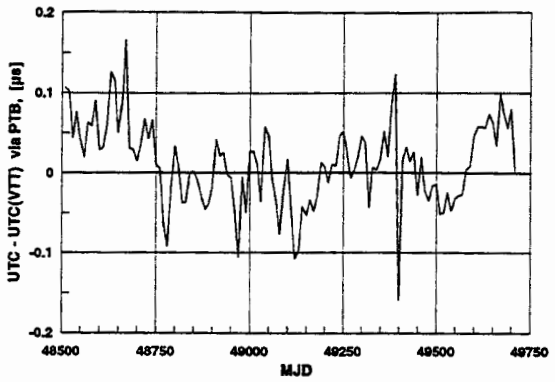


Fig. 5: UTC - UTC(VTT) via PTB

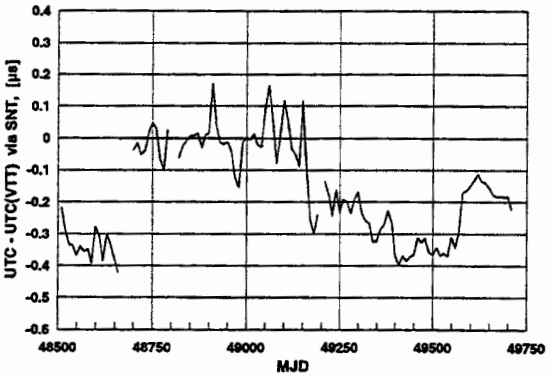


Fig. 6: UTC - UTC(VTT) via SNT

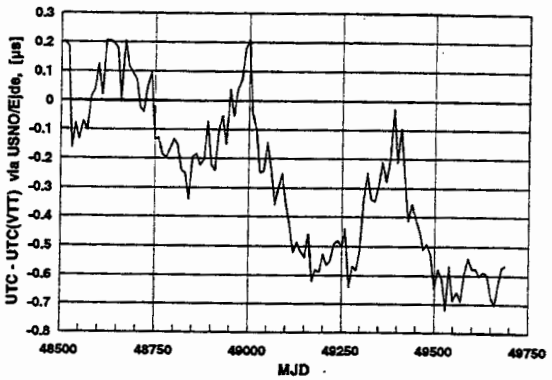


Fig. 7: UTC - UTC(VTT) via USNO/Ejde

Table 2: Common view Loran-C time scale averages and SDs

Reference	Average, ns	SD, ns
1 via PTB	12	53
2 via SNT	-127	84 (corrected)
3 via USNO	280	270

The University of Oulu has measured the Loran-C Bo slave transmitter by using the same GPS receiver (Navstar XR5M, 6 channels) as VTT. In Fig. 8 is represented the phase difference of the two measurements for 200 days in the latter half of the year 1994. The results show a slight trend, probably due to seasonal variations. SD of the sample is 65 ns.

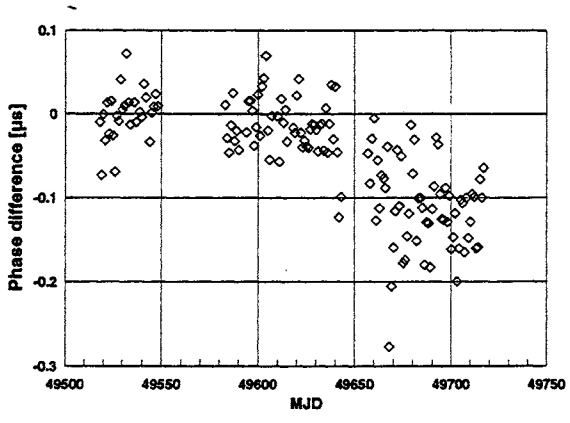


Fig. 8: The difference via GPS receivers, $UTC(Rx_{OULU}Bo) - UTC(Rx_{VTT}Sylt)$

4. CONCLUSIONS

Loran-C time scales via PTB and SNT have provided a low noise, low trend secondary path for time traceability. PTB data allows for ±100ns uncertainty/10 days. SNT data allows for ±150ns uncertainty/10 days provided that the possible receiver reparations and adjustments are known. USNO reports of Ejde (master) have a considerable trend and high scatter compared to Sylt (slave). By utilizing USNO/Ejde results a time scale can be maintained at ±1µs level/10 days only.
Sylt was phase-locked to Ejde master station, so one would not expect the observed long-time deviation trends between the two stations.

The constants k_1, k_2 and k_3 can be corrected by portable clock visits. the last H-maser visit from VNIIFTRI to VTT in January 1992 caused a 300 ns correction to UTC(VTT).

Reference: Tapio Mansten, Kalevi Kalliomäki and Bengt Selin; The accuracy of Sylt- and RNS-based time scales at MTI and SNT during 1990-1991, EFTF 92, Noordwijk

THEORETICAL TREATMENT OF CLOSING ERRORS OCCURRING IN TWO-WAY SATELLITE TIME TRANSFERS

J A Davis

Division of Electrical Science,
National Physical Laboratory,
Queens Road, Teddington, Middlesex, United Kingdom, TW11 OLW.

ABSTRACT

To obtain high accuracy absolute time transfers from a Two-Way Satellite Time and Frequency Transfer (TWSTFT) system the delay asymmetries of the earth station instrumentation must be calibrated. This has been successfully demonstrated by using a portable (flyaway) earth station which is co-located next to the permanent stations. Results from TWSTFT experiments performed between three or more stations using the INTELSAT (VA-F13) satellite at 307°E have revealed closing errors of up to several nanoseconds. These closing errors may limit the accuracy obtainable from portable earth station calibrations.

In this paper a theoretical treatment of the absolute delay calibration of a pair of satellite earth stations is presented. At present, portable earth station calibrations are calculated from 'side by side' measurements. This may give rise to closing errors occurring in the final calibration result. In the calibration method described here time transfers are made between the portable, local and remote stations during the same measurement session. An absolute calibration of each two-way time transfer is then obtained which is free from closing errors.

keywords
Time transfer, satellite communications, microwave measurement, delay measurement.

1) INTRODUCTION

Two-Way Satellite Time and Frequency Transfer (TWSTFT) is a high precision, easy to use method, which ultimately offers the possibility of sub-nanosecond accuracy time transfer [1] [2]. However, systematic closing errors of several nanoseconds have been observed in the time transfers made between a group of standards laboratories [3] [4]. The presence of these closing errors may undermine the possibility of obtaining sub-nanosecond accuracy time transfers, in particular the validity of the portable earth station calibrations which underpin the two-way work. Closing errors demonstrate an inconsistency in the theoretical description of the TWSTFT method.

In this paper the equations describing the TWSTFT method are re-examined. A re-formulation is presented which accounts for the closing errors. This re-formulation enables calculations of time transfers to be made which are not invalidated by the systematic delay changes which are the cause of closing errors.

2) RECENT EXPERIMENTAL RESULTS

During 1994 an international TWSTFT experiment was performed between six European and two North American laboratories using the INTELSAT satellite at 307°E. Closing errors were observed during time transfer experiments within Europe [3] and also during experiments between Europe and North America [4] [5].

In order to calibrate absolutely the delay asymmetries occurring between the different TWSTFT earth stations a portable FlyAway Satellite Terminal (FAST) earth station has been used to calibrate each station in turn, acting as a portable transfer standard [6], [7]. The co-location of the permanent and FAST stations enabled side-by-side common view measurements to be made between the receive chains of the two TWSTFT earth stations [5]. These common view measurements demonstrated that at the nanosecond level, the instrumentation delays within the receive chain of the TWSTFT earth stations were dependent on the characteristics of the transmitting earth station. This experimental observation has forced the re-evaluation of the TWSTFT theory.

3) LIMITATIONS OF THE EXISTING THEORY

An overview of the existing calculation method for a two-way time transfer, including the calibration of the earth stations using a portable station, is given in [7]. Using the same conventions the time difference $T_1 - T_2$ between the two clocks at stations 1 and 2 are given by equation (1) of [7] as:

$$T_1 - T_2 = 0.5(\Delta T_1 - \Delta T_2) + 0.5\{(\tau_1^u + \tau_2^d) - (\tau_2^u + \tau_1^d)\} + 0.5(\tau_{12} - \tau_{21}) + \Delta\tau_R + 0.5\{(\tau_1^{TX} - \tau_1^{RX}) - (\tau_2^{TX} - \tau_2^{RX})\} \quad (1)$$

where $0.5(\Delta T_1 - \Delta T_2)$ is the the half difference of the counter timer readings, $0.5\{(\tau_1^u + \tau_2^d) - (\tau_2^u + \tau_1^d)\}$ is the path asymmetry, $0.5(\tau_{12} - \tau_{21})$ is the transponder asymmetry, $\Delta\tau_R$ is the 'sagnac' term due to the earths rotation and $0.5\{(\tau_1^{TX} - \tau_1^{RX}) - (\tau_2^{TX} - \tau_2^{RX})\}$ is the earth station delay asymmetry.

The above equation may be simplified. The TWSTFT earth stations are operating using Ku band with transmit and receive frequencies in the 14.0GHz-14.5GHz and 12.5GHz-12.75GHz bands respectively. At the nanosecond level the effects of delay non reciprocality of the earth's ionosphere and troposphere may be neglected due to transmissions being made at relatively high frequencies.

For time transfers within Europe using the INTELSAT satellite at 307°E, the same frequency allocation of a single transponder is used to carry all of the transmissions. Under such conditions it has been assumed that the transponder delays are reciprocal

and $\tau_{12} = \tau_{21}$ [7]. Since the earth station down-link delays has been found to be dependent on the characteristics of the up-link station's signal, it is possible that the satellite transponder delays may demonstrate similar characteristics. It is very unlikely that the up-link instrumentation delay will depend upon which earth station is receiving the transmission. However this may happen if, for example the transmit code or EIRP levels of the up-link station are deliberately changed between two transmissions. For completeness, the reciprocality of the satellite transponder delays and constancy of the up-link instrumentation delays are not assumed.

4) EARTH STATION INSTRUMENTATION DELAYS

The dependence of the earth station down-link delays on the characteristics of the transmitting earth station's signal has been clearly demonstrated. This result leads to a requirement to specify the partner stations when determining the earth station instrumentation delays. The following conventions are used in this specification. Consider the delay τ_{AB}^{Rx}, the first subscript (A) denotes the transmitting station and the second (B) the receiving station, the superscript denotes either Tx or Rx operation. The term $\Delta\tau_{R12}$ represents the sagnac correction which must be added to the two way measurements to calculate the time transfer $(T_1 - T_2)$. τ_{12}^{S} is the satellite transponder delay when a signal is being transmitted by station 1 and received by station 2.

5) ORIGIN OF CLOSING ERRORS

Consider the case where time transfers are made in rapid succession between distant stations 1,2 and 3. Clock and instrumentation delay changes occurring between the individual measurement periods are neglected. The following equations may be constructed from equation (1):

$$T_1 - T_2 = 0.5(\Delta T_{12}^1 - \Delta T_{21}^2) + 0.5(\tau_{12}^S - \tau_{21}^S) + \Delta\tau_{R12} + 0.5\{(\tau_{12}^{Tx} - \tau_{21}^{Rx}) - (\tau_{21}^{Tx} - \tau_{12}^{Rx})\} \quad (2a)$$

$$T_1 - T_3 = 0.5(\Delta T_{13}^1 - \Delta T_{31}^3) + 0.5(\tau_{13}^S - \tau_{31}^S) + \Delta\tau_{R13} + 0.5\{(\tau_{13}^{Tx} - \tau_{31}^{Rx}) - (\tau_{31}^{Tx} - \tau_{13}^{Rx})\} \quad (2b)$$

$$T_2 - T_3 = 0.5(\Delta T_{23}^2 - \Delta T_{32}^3) + 0.5(\tau_{23}^S - \tau_{32}^S) + \Delta\tau_{R23} + 0.5\{(\tau_{23}^{Tx} - \tau_{32}^{Rx}) - (\tau_{32}^{Tx} - \tau_{23}^{Rx})\} \quad (2c)$$

The three time transfers $(T_1 - T_2)$, $(T_2 - T_3)$ and $(T_3 - T_1)$ may be added in order to construct a time transfer closure, the sum of which is by definition zero. From equations (2a-2c):

$$(T_1 - T_2) + (T_2 - T_3) + (T_3 - T_1) = 0.5(\Delta T_{12}^1 - \Delta T_{21}^2) + 0.5(\Delta T_{23}^2 - \Delta T_{32}^3) + 0.5(\Delta T_{31}^3 - \Delta T_{13}^1) + \Delta\tau_{R12} + \Delta\tau_{R23} + \Delta\tau_{R31} + 0.5(\tau_{12}^S - \tau_{21}^S) + 0.5(\tau_{23}^S - \tau_{32}^S) + 0.5(\tau_{31}^S - \tau_{13}^S) + 0.5\{(\tau_{12}^{Tx} - \tau_{21}^{Rx}) - (\tau_{21}^{Tx} - \tau_{12}^{Rx})\} + 0.5\{(\tau_{23}^{Tx} - \tau_{32}^{Rx}) - (\tau_{32}^{Tx} - \tau_{23}^{Rx})\} + 0.5\{(\tau_{31}^{Tx} - \tau_{13}^{Rx}) - (\tau_{13}^{Tx} - \tau_{31}^{Rx})\} = 0 \quad (3)$$

The sum of the sagnac terms $\Delta\tau_{R12} + \Delta\tau_{R23} + \Delta\tau_{R31}$ is zero.

It is possible using equation (3) to demonstrate how the closing errors described in [3] and [4] have originated. Consider the existing theory describing the calculation of a two-way time transfer, for example the derivations of equations (2) and (3) of [1]. In this analysis it has been incorrectly assumed that both the earth station instrumentation delays and the satellite transponder delays remain constant, irrespective of the configuration of transmitting and receiving stations. The following delays are (incorrectly) assumed to be equal.

$$\tau_{12}^{Tx} = \tau_{13}^{Tx}, \tau_{21}^{Tx} = \tau_{23}^{Tx}, \tau_{32}^{Tx} = \tau_{31}^{Tx}, \tau_{21}^{Rx} = \tau_{31}^{Rx}, \tau_{12}^{Rx} = \tau_{32}^{Rx}, \tau_{13}^{Rx} = \tau_{23}^{Rx}, \tau_{12}^{S} = \tau_{13}^{S}, \tau_{21}^{S} = \tau_{23}^{S}, \tau_{31}^{S} = \tau_{32}^{S} \quad (4)$$

The equalities given in equations (4) have enabled equation (3) to be (incorrectly) simplified to give:

$$(T_1 - T_2) + (T_2 - T_3) + (T_3 - T_1) = 0.5(\Delta T_{12} - \Delta T_{21}) + 0.5(\Delta T_{23} - \Delta T_{32}) + 0.5(\Delta T_{31} - \Delta T_{13}) = 0 \quad (5)$$

Because the assumptions made in equations (4) are inadequate, the combination of the counter timer readings used to calculate a time transfer closure in equation (5) are not zero. The discrepancy has been defined as the closing error (C_{123}) where [3]:

$$C_{123} = 0.5(\Delta T_{12} - \Delta T_{21}) + 0.5(\Delta T_{23} - \Delta T_{32}) + 0.5(\Delta T_{31} - \Delta T_{13}) \quad (6)$$

However from equation (3) it may be seen that the 'closing error' is in reality given by:

$$C_{123} = 0.5\{(\tau_{21}^{Tx} + \tau_{21}^{S} + \tau_{21}^{Rx}) - (\tau_{12}^{Tx} + \tau_{12}^{S} + \tau_{12}^{Rx})\} + 0.5\{(\tau_{32}^{Tx} + \tau_{32}^{S} + \tau_{32}^{Rx}) - (\tau_{23}^{Tx} + \tau_{23}^{S} + \tau_{23}^{Rx})\} + 0.5\{(\tau_{13}^{Tx} + \tau_{13}^{S} + \tau_{13}^{Rx}) - (\tau_{31}^{Tx} + \tau_{31}^{S} + \tau_{31}^{Rx})\} \quad (7)$$

The presence of a non zero closing error is a direct consequence of the earth station instrumentation delays being a function of the characteristics of the partner station's signal.

6) CLOSING ERRORS DETERMINED USING CO-LOCATED EARTH STATIONS

The conclusive evidence for the instrumentation delay changes which give rise to closing errors, was obtained from a common view experiment using a portable (FAST) earth station. This FAST station was co-located with each of fixed earth stations in turn, during which time common view experiments were performed.

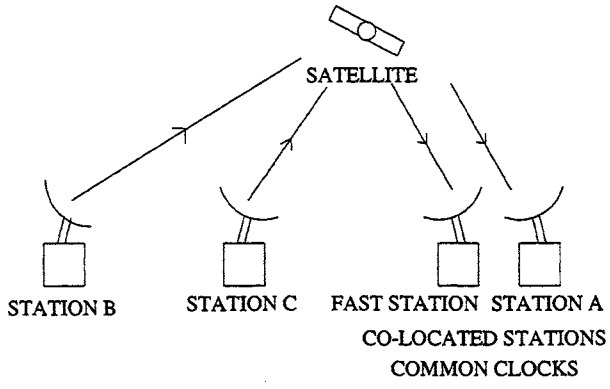


Figure 1 Co-located common view experiment

Let F and A represent the FAST and permanent earth stations and B and C the distant earth stations as shown in Figure 1. Common view measurements made between the co-located FAST and fixed earth station give:

$$D_{FA}^{B} = (\Delta T_{FB} - \Delta T_{AB}) = (\tau_{BF}^{Tx} + \tau_{BF}^{S} + \tau_{BF}^{Rx}) - (\tau_{BA}^{Tx} + \tau_{BA}^{S} + \tau_{BA}^{Rx}) \quad (8a)$$

$$D_{FA}^{C} = (\Delta T_{FC} - \Delta T_{AC}) = (\tau_{CF}^{Tx} + \tau_{CF}^{S} + \tau_{CF}^{Rx}) - (\tau_{CA}^{Tx} + \tau_{CA}^{S} + \tau_{CA}^{Rx}) \quad (8b)$$

where D^B_{FA} are the difference between the counter timer readings of the FAST station and station A when both stations are receiving signals from station B. Similar equations may be obtained for co-located measurements between the FAST station and the fixed stations B and C.

Experimental results have demonstrated that the two common view measurements D^B_{FA} and D^C_{FA} are not equal [4]. The common view measurements made between the FAST station and station A were found to have changed after switching between transmitting stations. It follows from equations (8a) and (8b) that at least one of the equalities given in equation (4) must be invalid. The most plausible explanation is that $(\tau^{Rx}_{BF} - \tau^{Rx}_{BA})$ or $(\tau^{Rx}_{CF} - \tau^{Rx}_{CA})$ are not constant, so that the receive delays of the earth station instrumentation are dependent on the choice of up-link signal. In principle, the satellite transponder delays may possess a similar dependence on the choice of up-link signals.

It is important to determine that the delay changes observed in these common view experiments [4] are actually responsible for the closing errors observed in [3]. The common view differences may be combined to form the appropriate closing error C_{123} where:

$$C_{ABC} = 0.5(D^A_{FB} - D^A_{FC} + D^B_{FC} - D^B_{FA} + D^C_{FA} - D^C_{FB}) =$$
$$0.5\{(\tau^{Tx}_{BA} + \tau^{S}_{BA} + \tau^{Rx}_{BA}) - (\tau^{Tx}_{AB} + \tau^{S}_{AB} + \tau^{Rx}_{AB})\} + 0.5\{(\tau^{Tx}_{CB} + \tau^{S}_{CB} + \tau^{Rx}_{CB}) -$$
$$(\tau^{Tx}_{BC} + \tau^{S}_{BC} + \tau^{Rx}_{BC})\} + 0.5\{(\tau^{Tx}_{AC} + \tau^{S}_{AC} + \tau^{Rx}_{AC}) - (\tau^{Tx}_{CA} + \tau^{S}_{CA} + \tau^{Rx}_{CA})\} \quad (9)$$

Experimental results show that the closing errors determined from these common view measurements are in good agreement with the closing errors determined from direct addition of time transfers [4]. This demonstrates that the delay changes observed during the common view measurements are responsible for closing errors.

7) PORTABLE EARTH STATION CALIBRATIONS

7.1 Introduction

The existence of closing errors will introduce significant errors into the existing portable earth station calibration method. An outline of this calibration method is given in [7]. In order to calibrate two TWSTFT earth stations A and B with respect to each other using a portable (FAST) station F, the relative delay asymmetry within the earth station's instrumentation must be determined. This is given in [7] as S_{AB} where:

$$S_{AB} = 0.5\{(\tau^{Tx}_{AB} - \tau^{Rx}_{BA}) - (\tau^{Tx}_{BA} - \tau^{Rx}_{AB})\} \quad (10)$$

Including the possible asymmetry of the satellite transponder, equation (10) becomes:

$$S_{AB} = 0.5\{(\tau^{Tx}_{AB} + \tau^{S}_{AB} + \tau^{Rx}_{AB}) - (\tau^{Tx}_{BA} + \tau^{S}_{BA} + \tau^{Rx}_{BA})\} \quad (11)$$

Following the method described in [7] the following equations are obtained for co-located FAST calibration measurements at stations A and B. Both the FAST and local stations are operating using common clocks:

$$0.5\{(\tau^{Tx}_{AF} + \tau^{S}_{AF} + \tau^{Rx}_{AF}) - (\tau^{Tx}_{FA} + \tau^{S}_{FA} + \tau^{Rx}_{FA})\} = 0.5(\Delta T_{FA} - \Delta T_{AF}) \quad (12a)$$

$$0.5\{(\tau^{Tx}_{BF} + \tau^{S}_{BF} + \tau^{Rx}_{BF}) - (\tau^{Tx}_{FB} + \tau^{S}_{FB} + \tau^{Rx}_{FB})\} = 0.5(\Delta T_{FB} - \Delta T_{BF}) \quad (12b)$$

In the existing analysis [7] the equalities given in equation (4) are erroneously assumed. Without making these assumptions S is calculated by subtracting equation (12b) from equation (12a):

$$S = 0.5(\Delta T_{FA} - \Delta T_{AF}) - 0.5(\Delta T_{FB} - \Delta T_{BF}) = 0.5\{(\tau^{Tx}_{AF} + \tau^{S}_{AF} + \tau^{Rx}_{AF}) -$$
$$(\tau^{Tx}_{FA} + \tau^{S}_{FA} + \tau^{Rx}_{FA})\} - 0.5\{(\tau^{Tx}_{BF} + \tau^{S}_{BF} + \tau^{Rx}_{BF}) - (\tau^{Tx}_{FB} + \tau^{S}_{FB} + \tau^{Rx}_{FB})\} \quad (13)$$

S is not equivalent to the required delay asymmetry S_{12} given in equation (11), demonstrating that the existing method for calibrating portable earth stations is in error.

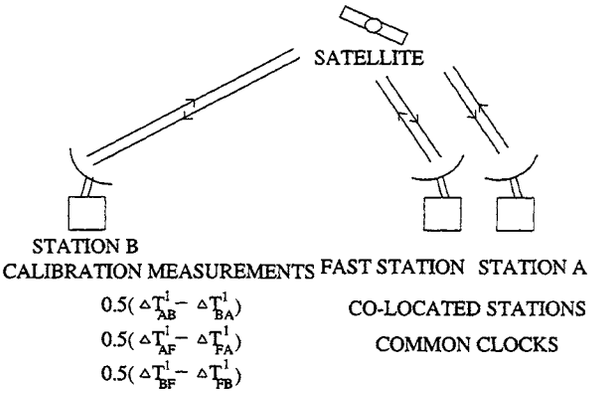


Figure 2a Fast calibration at station A (session 1)

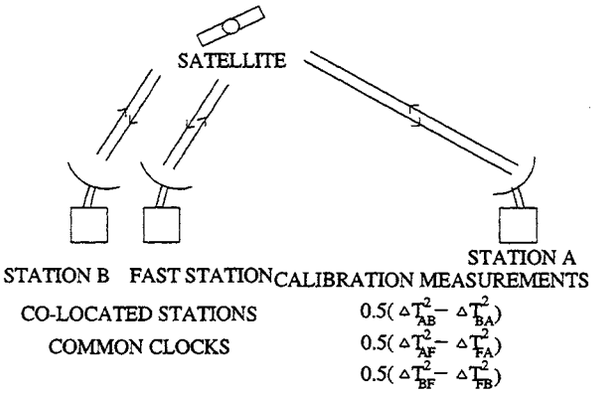


Figure 2b Fast calibration at station B (session 2)

7.2) Improved portable earth station calibration

An alternative method is required to calculate the earth station delay asymmetries using a portable earth station. Additional information is required from the calibration which is not provided by simple side by side measurements. A measurement scheme is developed here, which enables the earth stations to be correctly calibrated using a portable station. In addition to the side by side measurements made using the portable earth station, a series of time transfer measurements are made between the FAST, local, and remote stations as shown in Figures (2a,b). These measurements are made at each location. The additional information obtained from these time transfers is then sufficient to enable the accurate calibration of the required delay asymmetry given in equation (11).

Consider the case of two fixed earth stations A and B which are calibrated by using a portable (FAST) station F at both locations. The delay asymmetries between stations A, B and F are given by the following expressions:

$$S_{AB} = 0.5\{(\tau^{Tx}_{AB} + \tau^{S}_{AB} + \tau^{Rx}_{AB}) - (\tau^{Tx}_{BA} + \tau^{S}_{BA} + \tau^{Rx}_{BA})\} \quad (14a)$$

$$S_{AF} = 0.5\{(\tau^{Tx}_{AF} + \tau^{S}_{AF} + \tau^{Rx}_{AF}) - (\tau^{Tx}_{FA} + \tau^{S}_{FA} + \tau^{Rx}_{FA})\} \quad (14b)$$

$$S_{BF} = 0.5\{(\tau^{Tx}_{BF} + \tau^{S}_{BF} + \tau^{Rx}_{BF}) - (\tau^{Tx}_{FB} + \tau^{S}_{FB} + \tau^{Rx}_{FB})\} \quad (14c)$$

where S_{AB} is the delay asymmetry to be calibrated.

Two measurement sessions are performed. In measurement session 1, co-located common clock measurements $0.5(\Delta T^1_{AF} - \Delta T^1_{FA})$ are made between station A and the FAST station. Time transfers measurements $0.5(\Delta T^1_{AB} - \Delta T^1_{BA})$ and $0.5(\Delta T^1_{BF} - \Delta T^1_{FB})$ are also made (Figure 2a). In the following equations the superscripts denote the measurement session, the simplification $T^1_A = T^1_F$ is made due to the use of a common clock.

$$T^1_A - T^1_B = 0.5(\Delta T^1_{AB} - \Delta T^1_{BA}) + S_{AB} + \Delta\tau_{RAB} \qquad (15a)$$

$$0 = 0.5(\Delta T^1_{AF} - \Delta T^1_{FA}) + S_{AF} \qquad (15b)$$

$$T^1_B - T^1_A = 0.5(\Delta T^1_{BF} - \Delta T^1_{FB}) + S_{BF} + \Delta\tau_{RBA} \qquad (15c)$$

Measurement session 2 is performed after moving the FAST station from A to B. Co-located common clock measurements $0.5(\Delta T^2_{BF} - \Delta T^2_{FB})$ are made between station B and the FAST station. Time transfers $0.5(\Delta T^2_{AB} - \Delta T^2_{BA})$ and $0.5(\Delta T^2_{AF} - \Delta T^2_{FA})$ are also made (Figure 2b). The simplification $T^2_B = T^2_F$ is made due to the use of a common clock.

$$T^2_A - T^2_B = 0.5(\Delta T^2_{AB} - \Delta T^2_{BA}) + S_{AB} + \Delta\tau_{RAB} \qquad (16a)$$

$$T^2_A - T^2_B = 0.5(\Delta T^2_{AF} - \Delta T^2_{FA}) + S_{AF} + \Delta\tau_{RAB} \qquad (16b)$$

$$0 = 0.5(\Delta T^{2B}_{BF} - \Delta T^{2B}_{FB}) + S_{BF} \qquad (16c)$$

It is assumed that the earth station instrumentation delays and transponder delays are not a function of the earth station's location. In particular, moving the portable earth station from A to B is assumed not to change its instrumentation delay asymmetry or the transponder delays of its transmitted or received signals. The method would not successfully calibrate a time transfer link in which different transponders are used to carry the two separate signals which form the time transfer. The transponder delay asymmetry observed by the portable station would then be different at each location.

The required asymmetry is given by S_{AB} where:

$$S_{AB} = 0.5(\Delta T^1_{BA} - \Delta T^1_{AB}) + 0.5(\Delta T^1_{FB} - \Delta T^1_{BF}) + 0.5(\Delta T^2_{BF} - \Delta T^2_{FB}) \qquad (17a)$$

$$S_{AB} = 0.5(\Delta T^2_{BA} - \Delta T^2_{AB}) + 0.5(\Delta T^2_{AF} - \Delta T^2_{FA}) + 0.5(\Delta T^1_{FA} - \Delta T^1_{AF}) \qquad (17b)$$

The two equations given above may be used to give independent measurements of the relative delay asymmetry between stations A and B. Any discrepancies may be due to clock noise occurring between the individual measurement periods and delay instabilities and location dependent delay changes may also occur in the period between performing the measurements at both sites. The two measurements have been shown to agree well [4].

7.3) Discrepancies between the two methods

The delay asymmetries between station A and station B may be determined either by the method developed here or by the existing method. Calculating the difference between equations (17) and (13) will determine the error present in the existing method. This error is given by E where:

$$E = 0.5(\Delta T^1_{BA} - \Delta T^1_{AB}) + 0.5(\Delta T^1_{AF} - \Delta T^1_{FA}) + 0.5(\Delta T^1_{FB} - \Delta T^1_{BF}) \qquad (18a)$$

$$E = 0.5(\Delta T^2_{BA} - \Delta T^2_{AB}) + 0.5(\Delta T^2_{AF} - \Delta T^2_{FA}) + 0.5(\Delta T^2_{FB} - \Delta T^2_{BF}) \qquad (18b)$$

These discrepancies are numerically equivalent to the closing errors calculated between the stations A, B and the FAST station during each of the calibration sessions. By calculating values of these closing errors is possible to determine the error in the determination of the earth station delay asymmetry using the conventional side by side method. These errors have been shown to amount to several nanoseconds [4].

8) CONCLUSION

The equations underlying the TWSTFT method have been re-examined, to take account of experimentally observed closing errors. The existing method of earth station calibration using a portable station has been shown to be in error. An improved measurement method has been formulated which correctly calibrates the earth station delays in the presence of the systematic delay changes which are the cause of closing errors.

ACKNOWLEDGEMENTS

This work has been supported under the UK National Measurement System Program for Time and Frequency.

REFERENCES

[1] D.A.Howe, 'High accuracy time transfer via geostationary satellites: Preliminary results', IEEE Trans. Instrum. Meas. vol.37, pp.418-423, 1988.

[2] D.A.Howe 'Progress towards one-nanosecond two way time transfer accuracy using Ku-band geostationary satellites', IEEE Trans. Instrum. Meas. vol. UFFC-34, No6, pp639-646, 1987.

[3] 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', (To be published April 1995).

[4] J A Davis, P R Pearce, P Hetzel, D Kirchner, H Ressler, W Powell, A McKinley, J A DeYoung, W Klepczynski, G De Jong, A Söring, F Baumont, C Hackman, and L Veenstra, 'Closing errors observed during two-way time transfer experiments' Proc. EFTF95, (these proceedings).

[5] J A DeYoung, W J Klepczynski, A Davis, W Powell, P Mai, P Hetzel, A Bauch, J A Davis, P R Pearce, C Hackman, D Kirchner, G de Jong, A Soering, P Grudler, F Baumont, H Ressler and L Veenstra, 'Preliminary results of the International two way time and frequency transfer experiment', Proc PTTI94.

[6] W Klepczynski, J A Davis, D Kirchner, G De Jong, F Baumont, P Hetzel, A Soering, C Hackman, M Granveaud and W Lewandowski, 'Comparisons of two-way satellite time transfer and GPS time transfer during the Intelsat field trial' Proc PTTI94.

[7] D. Kirchner 'Two way time transfer via communication satellites' IEEE Proc. No.79, pp.983-990, 1991.

TIME REFERENCE SYSTEM OF THE ESO VERY LARGE TELESCOPE.

Werner R. Lange(1), Martin Ravensbergen(2).

(1) Lange-Electronic GmbH, Ganghoferstr. 29, 82216 Gernlinden, Germany.
Tel. ++49 8142 1041. Fax ++49 8142 13056.
(2) European Southern Observatory, Karl Schwarzschild Str. 2, 85748 Garching bei Muenchen. Germany.
Tel. ++49 89 32006409. Fax. ++49 89 3202362. e-mail: mravensb@eso.org

ABSTRACT

The European Southern Observatory is constructing a new observatory in Chile for the Very Large Telescope. The VLT consists of 4 telescopes. Each of them has a primary mirror with a diameter of 8 meters. The control architecture is based on workstations and VME computers. The VME computers are distributed over the telescopes and have a real time operating system.
Since the availability of the Global Positioning System (GPS), the generation of an accurate time signal on isolated telescope sites does not create a problem any more.
However, distribution a time signal to all VME computers with a high accuracy is an issue. The IRIG-B code, which is commonly used, is not adequate as accuracy requirements are in the 10 microsecond range. Also, synchronisation of software processes in one VME computer and in different VME computers must be solved.

This paper presents the design that is adopted for the VLT.
An overview of the requirements of the Time Reference System (TRS) is given. These requirements have been defined on the basis of experiences with the timing system of the ESO NTT telescope.
The hardware units are described. These are: a Central Time Standard, a Time Bus Distribution system and a VME Time Interface Module.
The distribution is based on fiber optic transmission, using a simple digital modulation that outperforms the analog IRIG-B modulation.
The Time Interface Module in the VME computer does not only perform the timing signal decoding, but contains also user-programmable timers that are synchronously clocked with the time source.

At the moment, all units of the TRS have been tested. Series production of the units for the distribution and the VME Time Interface Modules is in progress.

1. THE VLT OBSERVATORY

The 4 VLT telescopes can operate individually and in an interferometric mode. In the latter case, the light from the celestial object is combined via optical delay lines in the combined Coudé focus.
The control of all the functions is done with Local Control Units (LCU's). These are VME computers with the real-time operating system VxWorks. The LCU's are connected to Local Area Networks. Operator control is done from UNIX workstations, which are connected to these networks.
The VLT observatory will have about 150 LCU's.
The size of the observatory is about 300 meter as its biggest diameter.

2. TIME REFERENCE SYSTEM

The individual telescopes with the distributed real time LCU's highlight that there is a requirement for a synchronous timing between LCU's. There are fast control loops which are physically separated between 2 or more LCU's. This is already the case for the individual telescopes, but the interferometric mode has even stronger requirements in this respect. This is accomplished with the Time Reference System (TRS).
The TRS will:

- contain the central clock for the observatory,
- provide synchronisation facilities for the workstations,
- distribute a timing signal to the LCU's,
- provide accurate timing facilities in the LCU's.

2.1. TIME BASIS

Astronomical coordinates are based on Sidereal Time. This time is however local and can therefore not be accepted as the time basis for the VLT observatory. International time is UTC and ESO has decided to use UTC as the time basis for the TRS. All calculations of coordinates will be done in the UTC time frame. Sidereal Time is calculated with UTC as source.

2.2. DIFFERENCES BETWEEN UT1 AND UTC.

The difficulty that must be solved with this approach is the difference between UT1 and UTC. UT1 is based on the duration of a 'mean solar day' while UTC is based on a constant interval for the second. The UT1 is needed as an intermediate step in order to calculate the telescope coordinates. The difference between UT1 and UTC is a variable. It is estimated by the International Earth Rotation Society (IERS) one month in advance and is available as a table with the values listed per day. These values must be entered into the VLT computer system.
If the value approaches 1 second, a leap second is inserted into the UTC.

2.3. LEAP SECONDS

Leap seconds in the UTC are not (yet) handled by commercial available equipment. Some suggestions about this point are described in chapter 4.1.
It is necessary for the VLT to enter the UT1-UTC into the computer system. This list contains (obviously) also the information about leap seconds. A 'warning' about an upcoming leap second can then be made available in the VLT software system.
On the basis of these 2 arguments, ESO decided to put no requirements about leap seconds. It is however certainly an aspect that deserves proper handling by the UTC reception and distribution systems so that it is 'on-line' available for the users.

2.4. MODIFIED JULIAN DAY

The LCU's must count the time synchronously, also at the change of the day. They must therefore receive the full time information, that is also Modified Julian Day (MJD) or day of the year and year.
MJD has the preference as this allows simple calculation in the LCU's.
It should be noted that the MJD in this context is an integer. In the normal representation, it is a real (that is, it includes the fraction of the day).
Another aspect is that MJD is based also on 'mean solar days' and relates therefore to UT1, *not* to UTC. The difference between these 2 is always less than 1 second.
Considering the existence of time codes, ESO chose to specify the addition of MJD (as integer) in the time code format. This MJD increments at midnight UTC. One could define this a Modified Julian 'Date'.

2.5. ACCURACY

The tracking of a celestial object requires an accuracy of the time at the LCU of better than 100 µs. If this value is reached, the effect of this is neglectable in the tracking.
Time tagging of events which are related to some types of astronomical observation require an accuracy of better than 10 µs.
The requirements from the interferometer on the accuracy are not known in detail yet, but this comes to the same class.
ESO decided therefore to put an accuracy requirement of 10 µs.

2.6. WORKSTATION TIMING

The timing must be unique and guaranteed in the whole observatory. LCU's need accurate timing. Workstations need synchronisation as well, although not with the same accuracy requirements as they are not real time computers. ESO has put the accuracy requirement for the workstations to 10 ms.

2.7. SYNCHRONISATION

Time reception with the GPS system is very reliable nowadays. Still, we consider a fault in the GPS transmission or reception a 'likely' failure. Also, the repair of a fault the GPS is not under direct control from ESO, but it depends on a third party.
We have put more emphasis on the *synchronisation* of the time in the VLT observatory than on highest *accuracy*.
Therefore, ESO specified one central clock from which all timing is derived.
Synchronous timing is then obtained, even if there is a fault in the GPS reception

2.8. STABILITY

It is an argument, particularly for an astronomical observatory, to minimise the repair effort during night time. ESO specified therefore a stability of the 'flywheel' that is sufficient to maintain the accuracy over a time interval of 10 hours. This is equivalent to 3.10^{-10}.
This allows to delay the repair activity to daytime.

2.9. RELIABILITY

The central part of the reception of the TRS is vital for the proper operation of the VLT observatory. The site provides an Uninterruptable Power Supply (UPS) of 230V AC. All workstations and LCU's will be powered from that supply and it will power the TRS as well.
ESO has bad experiences with local UPS systems: they have proven to be the weakest part of a system.
We made the choice to specify a double central part of the system, without local battery back-up.

2.10. JITTER AND DISTRIBUTION CODING

Noise, which enters the time distribution system, causes time jitter in the detection mechanism at the receiver end.
The ESO observatory in La Silla uses the commonly used coding, the IRIG-B code. This is an analog signal: the carrier is a 1 kHz sine wave and the code is modulated as amplitude modulation. Each bit has a duration of 10 ms. This is not adequate for modern timing systems for the following reasons:

- The 1 kHz carrier does not allow easy re-generation of the UTC with high accuracy: In view of the 10 µs accuracy requirement, one needs a clock frequency of 1 MHz at the receiver end. To phase lock this 1 MHz with a 1 kHz incoming signal will result in unacceptable jitter. This can only be avoided with special components like Voltage Controlled Crystal Oscillators and using very low bandwidth in the PLL.
  This effect is reduced when a higher carrier frequency is chosen. Although codes with higher carrier frequencies are defined by IRIG, they are not common.
- The TRS must operate in an Electro Magnetic environment that can be compared with 'light industry', certainly not with a laboratory environment. Modern motor control is made with variable frequency inverters which generate electrical spikes. These spikes interfere with an electronic time bus distribution signal.

The second aspect has been clearly observed at the NTT telescope at La Silla and could only be solved by a redesign of the input circuitry at the receiver. Still, the accuracy of the timing at the receiver can not be guaranteed.

ESO specified therefore to use for the VLT observatory:

- Optical fibers for the distribution. Fibers are already common for LAN's.
- A digital coding scheme with a 1 MHz carrier.

2.11. DECODING IN THE LCU'S

The functional requirements that were put on the LCU board were:

- decoding of the time distribution signal.
- a quartz oscillator for test and backup.
- UTC register, accessible from the LCU computer bus.
- Six user-programmable timers, synchronous with the time distribution signal.
- Programmable VME interrupts from the UTC register and the timers.

3. DESIGN AND TEST RESULTS

Based on these requirements, the Time Reference System (TRS) has been designed with:

- a Central Time Standard (CTS).
- a Time Bus distribution system.
- a Time Interface Module in every LCU that requires accurate timing.

This is schematically shown in fig. 1.

3.1. CENTRAL TIME STANDARD

The CTS has a GPS receiver, which disciplines a Rubidium Oscillator. A Rb oscillator is sufficient to obtain a low jitter in the time and to keep the accuracy within the specifications in case of a fault in the GPS.
The CTS has a Time Server unit. This unit is connected to the Local Area Network. The workstations will synchronise their clocks with the CTS clock via the Network Time Protocol (NTP) [1].
The CTS has also a coder that codes the UTC into a serial distribution signal.
Physically, the CTS consists of 2 units:

- a DATUM GPS Time/Frequency System [2], incorporating the GPS receiver, Rb clock and serial coder.
- a DATUM Network Time Server [3], incorporating the Time Server with the NTP protocol.

Modifications were implemented by DATUM and LANGE in the Datum GPS unit in order to comply with the requirements:

- The Modified Julian Date was made visible on the CTS display.
- Modified Julian Date was inserted in the serial code.
- The modulation method for the serial code was changed, see chapter 3.2.

The GPS unit provides an accuracy of 400 nanoseconds compared with the UTC, including SA effects.

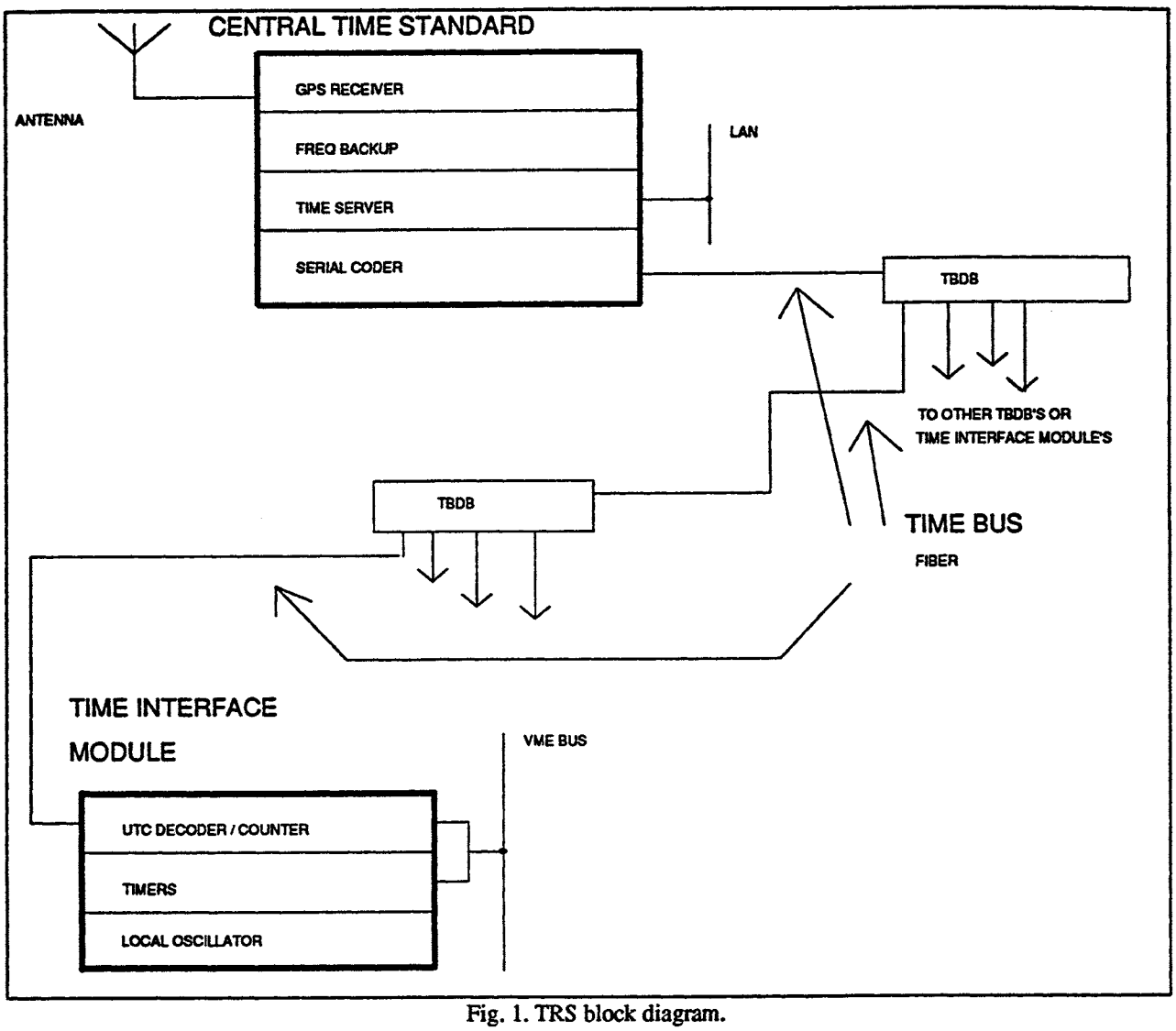


Fig. 1. TRS block diagram.

3.2. TIME BUS CODING

One can see on the basis of the requirements that the problems arise at the point where the distribution starts. The IRIG-B coding *scheme* as such can be used, but the modulation technique needs to be modified.

We decided to change the modulation method into a digital pulse width modulated signal with a carrier frequency of 1 MHz. If the code bit is '1', the duty cycle is 75%. If the bit is '0', it is 25%. This modified time code format can be translated by a standard IRIG-B time code reader if this reader has an appropriate low pass filter in its input.

3.3. DISTRIBUTION

This fiber optic Time Bus needs then to be distributed to the LCU's. This is done with a tree-like structure: The serial code signal is transformed into an optical signal at the CTS and is distributed to all LCU's via Time Bus Distribution Boxes (TBDB). These TBDB's act like active optical splitters.

The size of the VLT observatory allows to accept the differences in time delay in the fibers due to propagation.

Special attention had to be given to the design of the TBDB's: Optical transmitters and receivers have a different propagation delay for the rising and the falling edge of the digital signal. This is known as pulse distorsion. The effect is a function of the optical power, temperature, etc. It can easily be 50 ns per TBDB. This effect is cumulative when the signal passes through more than 1 TBDB. The VLT design needs a maximum of 5 TBDB's in the Time Bus: this would result in an unacceptable pulse distorsion. Selecting fiber optic components with a bandwidth of 25 MHz allowed to have a TBDB with the following specifications:

- Propagation delay: < 70 nanoseconds.
- Pulse width distorsion: < 25 nanoseconds.

The TBDB is constructed as a receiver board and a transmitter board with 4 outputs. If more outputs are needed on a given TBDB, more output boards can be added.

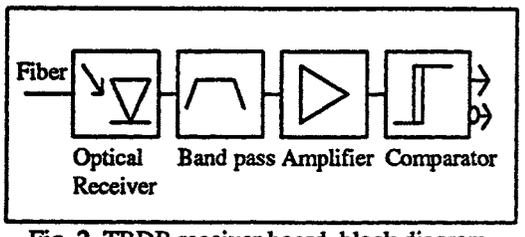


Fig. 2. TBDB receiver board, block diagram.

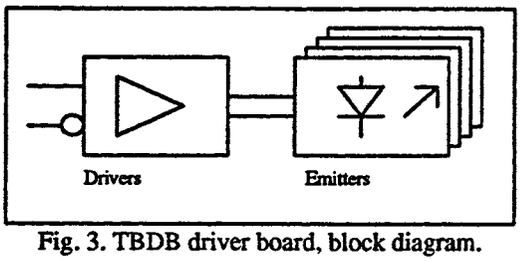


Fig. 3. TBDB driver board, block diagram.

The fiber which is used for the distribution is a 62.5/125 µm fiber. The same type of fiber is used for the LAN's.

3.4. TIME INTERFACE MODULE

3.4.1. TIME BUS DECODING

The decoding of the time bus signal in the TIM is relatively straight forward: the 1 MHz PWM signal is fed to a Phase Lock Loop. The other PLL input is the VCO output: a 1 MHz block wave with approximately 50 % duty cycle. That signal is used as a synchronous clock throughout the TIM.
As the PLL does not have a digital divider, the jitter between the Time Bus carrier and the 1 MHz block wave is neglectable.
As both PLL input signals are digital, the phase shift between the 2 signals is neglectable as well.
These 2 arguments allow to use a simple PLL: there are not high requirements on the stability of the VCO. The PLL does not need to have a low bandwidth either.

The falling edges of the block wave are used to clock a flip-flop which D-input is fed by the time bus signal. The envelope of the Time Bus code signal is then available on the output of the flip-flop.
The 'lock indicator' of the PLL is used as a 'Time Bus OK' logic signal. If false, the TIM switches to a 1 MHz quartz as clock source and loading of the UTC counters with the Time Bus data is inhibited.

The quartz oscillator has the common available tolerance of 100 ppm. It is intended to allow:

- the TIM to continue functionally in case of a Time Bus fault, and
- the TIM to operate in a stand alone mode for testing of the subsystem.

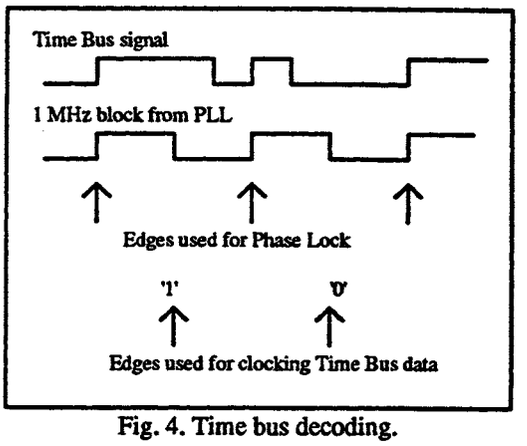


Fig. 4. Time bus decoding.

The UTC counter on the TIM is synchronously clocked by the 1 MHz clock. Intermediate results of this process are synchronous clocks of 1 Hz to 1 MHz in decimal steps. Any of these can be selected as clock for any of the user programmable timers.

The propagation delay from Time Bus input to VME interrupt is less than 350 nanoseconds.

3.4.2. SYNCHRONOUS TIMING

The major design driver for the Time Interface Module (TIM) was to have all the LCU timing synchronous with the UTC:

- User-programmable timers should generate (repetitive) interrupts to the VME bus that are synchronous with the UTC.
- These interrupts should start on a programmable UTC.

These features, combined with the fact that there are 6 user programmable timers, allow synchronisation of different software processes in the same LCU and even in different LCU's.

The set-up of the timers is done via the LAN. At a defined UTC, both timers start to count by hardware. The synchronicity between 2 interrupts in the same LCU and also in two different LCU's is then guaranteed.
These interrupts can be single or repetitive.
It is possible to start and to stop this on a defined UTC, thus allowing to generate the same amount of timer interrupts in the 2 processes.

More details of the TIM hardware and driver software design are described in [4].

4. IMPROVEMENTS

The design of the TRS is finished and tests have been succesfully passed. Nevertheless, we want to raise some aspects which could improve the overall system performance.

4.1. LEAP SECOND INFORMATION

The connection between the GPS unit and the Network Time Server is made with a short cable. It utilises the classical IRIG-B coding, which does not have a defined bit field for leap second information. The NTP protocol does have a field for it. It is however not used.
The TIM uses the IRIG-B code scheme and misses therefore also information about leap seconds.
The CTS has a serial port which provides access to this information. This needs then special software.

It would be good that the leap second information from the GPS is coded into the IRIG-B code scheme on an agreed field.

4.2. TIME FORMAT

The present TIM uses the hours, minutes and seconds bit fields to decode the time of the day. The CTS provides also the time of day in straight binary seconds.
Although HH:MM:SS is convenient of hardware displays, TIM hardware and driver software will be more straightforward if the binary seconds would be used.

4.3. AUTOMATIC UT1-UTC

The difference UT1-UTC needs to be entered 'by hand' into the VLT software system.
It would be good to have an automatic access to this information.

4.4. ACCURACY

The accuracy is mainly defined by propagation delay. This could be easily compensated in the TIM.

4.5. TIME BUS CODING

The presented PWM coding works well. Coding and decoding is simple. It contains however a DC component which caused some difficulty in the design of the TBDB. From that point of view, a phase encoding would be better.

5. CONCLUSION

From the allocated accuracy budget of 10 µs, the presented system uses:

- 400 ns for the CTS,
- Approx. 100 ns per TBDB, this means 500 ns in total,
- 350 ns for the TIM
- 1500 ns propagation delay through the fibers.

This results in approx. 2.5 µs, most of it is propagation delay which could be compensated in the TIM.

6. ACRONYMS

CTS	Central Time Standard.
ESO	European Southern Observatory.
GPS	Global Positioning System.
IERS	International Earth Rotation Society.
IRIG	Inter Range Instrumentation Group
LAN	Local Area Network.
LCU	Local Control Unit.
MJD	Modified Julian 'Date'.
NTP	Network Time Protocol.
PLL	Phase Locked Loop.
PWM	Pulse Width Modulation.
TBDB	Time Bus Distribution Box.
TIM	Time Interface Module.
TRS	Time Reference System.
UTC	Universal Time Coordinated.
VLT	Very Large Telescope.

7. REFERENCES

[1] Network Time Protocol.
E.g. D. Mills: "Internet Time Synchronization: The Network Time Protocol", IEEE Transactions on Communications, Vol. 39, No 10, Oct. 1991.

[2] GPS Time Frequency system,
Model 9390-5500. (This is the generic number; model number that implements modifications is 9390-55199).
DATUM Inc.

[3] Network Time Server,
model Tymserve 2000.
DATUM Inc.

[4] M. Ravensbergen, K. Wirenstrand,
'Time Reference System of the VLT',
To be published in: 'Telescope Control Systems", SPIE Vol. 2479, April 1995, Orlando Florida USA.

-o-

EXPOSANTS - EXHIBITORS

POSEIDON SCIENTIFIC INSTRUMENTS	14 Mouat street	AUS-6160 FREMANTLE
JTELEC	52 Parc des Essarts	F-78690 LES ESSARTS LE ROI
KONTRON ELEKTRONIK	6 rue des Frères Caudron	F-78143 VELIZY
C-MAC Quartz Crystals Ltd	Edinburgh Way	UK-CM20 2DE HARLOW, Essex
CQE Composants Quartz et Electronique	Rue R. Keller	F-10150 PONT STE MARIE
CEFOR Ingénierie	Square des Champs-Elysées	F-91026 EVRY
RODITI INT. Corp. Ltd	130 Regent Street	UK-W1R 6BR LONDON
3S NAVIGATION	23141 Plaza Pointe Drive	USA-LAGUNA HILLS, CA 92653
ALLEN OSBORNE Ass. Inc.	756-J Lakefield Road	USA-WEST LAKE VILLAGE, CA 91361
TRAKMICROWAVE Corp.	4796 Eisenhower Blvd	USA-TAMPA FLORIDA 33634
GILLAM-SATEL	Route de Demigny	F-71100 CHALON/SAONE
TIMETECH GmbH	Nobelstrasse 15	D-70569 STUTTGART
QUARTZLOCK Instruments	Plymouth Road - "Gothic"	UK-TQ9 5LH TOTNES
HEWLETT-PACKARD	Startbaan 16	NL-1180A2 AMSTELVEEN
Frequency Electronics Inc.	55 C. Lindbergh Blvd	USA-MITCHEL FIELD, NY 11553

PRESIDENTS DE SEANCE - SESSION CHAIRMEN

	Session
L. MALEKI - Séance d'ouverture / Opening session Jet Propulsion Laboratory - 4800 Oak Grove Drive - USA-Pasadena, CA 91109	
M.J. UNDERHILL - Résonateurs I / Resonators I The Institution of Electrical Engineer - Savoy Place - UK-London WC2ROBL	1
M. BRUNET - Applications spatiales / Space Applications CNES - Dept. AE/TT/TF - Av. Edouard Belin - F-31055 Toulouse Cedex	2
J.P. AUBRY - Matériaux I / Materials I Thomson-CEPE - B.P. 165 - F-95105 Argenteuil	3
G. PETIT - Transmission de l'heure / Time transfer BIPM - Time section - Pavillon de Breteuil - F-92312 Sèvres Cedex	4
B. NEUBIG - Oscillateurs / Oscillators Tele Quarz GmbH - Landstrasse 2 - D-74924 Neckarbischofsheim	5
L. CUTLER - Standards atomiques I / Atomic Standards I Hewlett Packard Lab. - Deer Creek Road - USA-Palo Alto, CA 94303	6
J. ZELENKA - Résonateurs II / Resonators II Technical University of Liberec - Halkova 6 - 46117 Liberec - Czech Republic	7
C. AUDOIN - Standards de fréquence atomiques : génération du signal et analyse / Atomic Frequency Standards : Signal Generation and Analysis Laboratoire de l'Horloge Atomique - CNRS - Bât. 221 - F-91405 Orsay	8
V. PLESSKY - Ondes de surface I / SAW I ASCOM Microsystems SA - Case Postale 415 - CH-2022 Bevaix	9
K. HILTI - Applications aux Télécommunications / Telecom applications Swiss Telecom PTT - Ostermundigenstr. 91-93 - CH-3000 Bern	10
J.J. GAGNEPAIN - Session de remise de prix / Award session CNRS - Dept. SPI - 3 rue Michel Ange - F-75794 Paris Cedex 16	
E. PHILIPPOT - Matériaux II / Materials II URA CNRS D 0407 - Université de Montpellier - F-34095 Montpellier Cedex	11
A. DORENWENDT - Standards atomiques II / Atomic Standards II Physikalisch-Technische Bundesanstalt - D-3300 Braunschweig	12
R.J. BESSON - Résonateurs III / Resonators III LCEP - ENSMM - 26 chemin de l'Epitaphe - F-25030 Besançon Cedex	13
P. KARTASCHOFF - Synchronisation dans les systèmes / Synchronization in systems Swiss Telecom PTT - Ostermundigenstr. 91-93 - CH-3000 Bern	14
J.M. HODE - Ondes de surface II / SAW II Thomson-Microsonics - B.P. 232 - F-06904 Sophia-Antipolis Cedex	15
P. THOMANN - Stabilisation de la fréquence laser / Laser Frequency Stabilization Observatoire Cantonal - Rue de l'Observatoire 58 - CH-2000 Neuchâtel	16

INDEX DES AUTEURS - AUTHOR INDEX

AUTEUR / AUTHOR	PAGES
ABRAMSON I.V.	87
ADLER B.	480
AKSYONOV V.M.	99
ALLAN D.W.	71
AL'SHIN B.	277
ANEMOGIANNIS K.	183
APRISHKIN G.	277
ASCARRUNZ F.G.	170
AUCOUTURIER E.	111
AUDOIN C.	147
AZOUBIB J.	283
BACHHEIMER J.-P.	423
BAERISWYL P.	46
BAHADUR H.	413
BAHOURA M.	101
BALLANDRAS S.	332
BARILLET R.	147
BARTLE K.A.	81
BAUMONT F.	81
BAUREGGER J.	183
BAVA E.	345
BAZZI O.	529
BECKER R.A.	251
BEDRICH S.	34, 485
BEIER J.	195
BENOIST E.	502
BERTHAUT A.	9
BETTERMANN D.	387
BEUTLER G.	46
BEZDELKIN V.V.	99
BIGLER E.	59, 332
BLIN T.	88
BOGOUSH M.E.	434
BONDU F.	353
BONELLO R.	220, 509
BONJOUR C.	59
BORISOCHKIN V.	277
BOURQUIN R.	127
BOURRAGUE Y.	502
BOUSSERT B.	384, 394
BOWER D.	245
BOY J.J.	9, 255, 438
BRENDEL R.	88
BRILLET A.	353
BRUNET M.	40
BUZEK O.	400
CACICOVSCHI D.M.	397
CAPELLE B.	289, 422
CARBONELLI M.	509
CEMUSOVA B.	400
CEREZ P.	111, 384
CERMAK J.	400
CHAUBET M.	297
CHEN D.P.	190
CHOMMELOUX L.	200
CHURCHILL N.	245
CLAIRON A.	101
CLAUDON P.	81
CLEVA F.	353
COCHET-MUCHY D.	289, 422
COCHEZ M.	59
COSTANZO G.	265
CUNIN B.	361
CUTLER S.	71
DAHER N.	417
DANIAU W.	93, 297
DANZY F.	404
DAVIS J.A.	81, 537
DE JONG G.	81
DE MARCHI A.	177, 265, 365
DESAINTFUSCIEN M.	349
DE SETA D.	509
DETAINT J.	289, 422
DEYZAC F.	9
DE YOUNG J.A.	81
DIMARCQ N.	111, 384, 394
DI MONACO O.	297
DINCA M.P.	369
DOBERSHTEIN S.A.	343
DOBROGOWSKI A.	217, 513
DOWSETT J.	245
DRULLINGER R.E.	166
DUDLE G.	115
DULMET B.	121, 301
DUPONT J.	349
DURAND S.	423
DUTREY J.-F.	40
ECKERT C.	361, 377
ENGELBERG J.	357
ERTMER W.	387
ESKELINEN P.	357
FERMIGIER B.	349
FILIATRE C.	454
FOISSY A.	454
FRANK A.H.	404
FREIK A.D.	20
FREY M.	462
FRILANDER J.	357
FRITSCHEL P.	353
GABARAEV O.G.	327
GAIGEROV B.	277
GALLIOU S.	301
GANKIN V.Z.	29
GAVIGNET E.	332
GEVORKYAN A.G.	29, 380
GHEORGHIU O.C.	397
GHEZALI S.	101
GIFFARD R.	71
GIORDANO V.	297, 377
GIRARDET E.	93
GIURGIU L.C.	397
GNIEWINSKA B.	24
GODONE A.	220
GOHIER R.	59
GOLDING W.M.	404
GOTOH T.	314
GOUZHVA Y.G.	29
GREENHALL C.	153
GROUZINENKO V.B.	434
GRUDLER P.	81
GUILLEMOT P.	88
GUSEV O.B.	327
GUZZO P.L.	438
HACKMAN C.	81
HAHN J.	485
HAMROUR M.R.	301
HARTING A.	34
HARTMANN C.	190
HELMCKE J.	408
HETZEL P.	81
HODE J.M.	200
HOUSSIN M.	349
HRUSKA C.K.	53
HUMPHREYS D.A.	516
IMAMURA K.	314
IMAOKA A.	317
JEANDEL G.	261
JENNI G.	132
JESSA M.	217
JONES E.C.	404
JUOMPAN B.	40
KALLIOMAKI K.	498, 521, 533
KHALIFA N.	529
KHURSHEED A.	177
KHUTORSCHIKOV V.	380
KIHARA M.	317
KIRCHNER D.	75, 81
KLAPPER H.	251
KLEPCZYNSKI W.	81
KLISCHE W.	311
KNIGHT D.J.E.	516
KNOOP M.	390
KOSHELYAEUSKY N.	80, 277
KOSTNER H.	480
KOUTCHEROV V.	277
KOYAMA M.	16
KRAMER G.	408
KRASNIKOVA S.N.	338
KREMPL P.W.	66
KRISPEL F.	66
KULAKOV S.V.	327
KUSTERS J.	71
LAMBERT C.	190
LANGE W. R.	541
LANTZ E.	373
LANDMANN A.	450
LARDET-VIEUDRIN F.	297
LATRACH L.	353
LAURENT P.	101

AUTEUR / AUTHOR	PAGES
LEA S.	101
LEE W.D.	166
LEUPRECHT G.	66
LIAO C.S.	525
LIN C.C.	525
LIN H.D.	525
LIPPHARDT B.	408
LISOWIEC A.	458
LONGET C.	141
MACHUI J.	183
MALEKI L.	1
MALYUKHOV V.A.	343
MAN C.N.	353
MANNERMAA J.	521
MANSFELD G.D.	20
MANSTEN T.	498, 533
MANZALINI A.	220, 509
MARGUIER J.	93
MARIANNEAU G.	59, 88, 93, 377
MARTIN G.	297
MASIUKIEWICZ A.	24
MASSARI F.	345
MATEESCU Irina	443
MATEESCU Ioan	443
MATOLA M.	357
McHUGH M.	373
McKINLEY A.	81
MEDVEDEV Yu.	277
MEMBREY F.	454
MIEHE J.A.	361, 377
MIHALCEA B.M.	397
MILETI G.	271
MOHANAN P.	431
MOMOSAKI E.	473
MORLOT G.	261
MOUREY M.	20, 377
MOURIER V.	502
MULLER J.H.	387
MYASNIKOV A.L.	29
NAU H.	311
NEFEDOV V.G.	327
NEUBIG B.	195, 469
NICULESCU A.	397
NOSEK J.	447
NOVERO C.	220
OLIVIER M.	297
OOMURA Y.	16
OURA N.	473
PALMIER D.	59
PASQUALI M.A.	255
PEARCE P.R.	81
PELLETIER V.	502
PENG J.L.	387
PERRONE F.	353
PERUCCHINI D.	509
PETIT G.	323
PETIT P.	111
PHARAOH K.I.	516
PHILIPPOT E.	59
PLANAT M.	377
PLESSKY V.P.	338
PLUMELLE F.	349
POIGNON C.	261
POWELL W.	81
POWERS E.D.	404
PUSHKIN S.	277
RAOELIJAONA F.	121
RATIER N.	88
RAVENSBERGEN M.	541
RESSLER H.	75, 81
RICHARD Y.	362
RIEGER V.	387
RIEHLE F.	408
RIHA G.	183
ROBICHON G.	141
ROBNIK R.	75
ROCHER J.	390
ROMISCH S.	365
RUBIOLA E.	265
RUSCHEWITZ F.	387
SAGNA N.	115
SANTERELLI G.	101
SALT D.	245
SCHILDKNECHT T.	46
SCHNATZ H.	408
SCHWAB M.A.	190
SEBASTIAN M.T.	251, 431
SELIMOTO H.	16
SELO G.	384
SENGSTOCK K.	387
SEYDEL E.	450
SHEU C.M.	525
SHI J.L.	525
SHIRLEY J.H.	166
SICCARDI M.	365
SMIRNOVA G.	380
SOLAL M.	200
SOLUCH W.	477
SORING A.	81
SREEMOOLANADHAN H.	431
STEHR U.	480
STEPANOV V.	277
SZULC W.	24
SZYMANIEC K.	101
TAKAGI M.	473
TAROT J.M.	362
TATARENKOV V.	277
TAVELLA P.	220
TELLIER C.	423
TEL'PUKHOVSKI N.	277
THEOBALD G.	394
THOMANN P.	115, 271
THOMAS C.	283
TORGUET R.	529
TOURDE R.	40
TRUCHOT P.	127
TULADHAR K.K.	132
UHRICH P.	40
URBANSKY R.	210
UTZINGER J.	46
VALENTIN J.-P.	454
VALENTIN M.	454
VECCHI G.	177
VEDEL F.	390
VEDEL M.	390
VEENSTRA L.	81
VENOT D.	147
VENTURA P.	200
VERNOTTE F.	373, 377
VIARD B.	255
WALLNOFER W.	66
WALLS F.L.	153, 166, 170, 227, 365
WALTER T.	153
WATANABE Y.	16
WEISS M.A.	24
WEISS K.	153
WOJCICKI M.	458
WOLF P.	323
WOLF M.	205
WRIGHT D.F.	492
YAMAKITA M.	473
YAMNI K.	255
YOLKIN G.	277
ZALAMANSKY G.	373
ZAREMBOVITCH A.	59
ZARKA A.	289, 422
ZELENKA J.	306, 443
ZHUKOV V.A.	327
ZORIN A.	80

Imprimé en France

par l'Imprimerie du
Conseil Général du DOUBS
Hôtel du Département
7 avenue de la Gare d'Eau
25031 BESANCON Cedex

Printed in France

by the Printing Office of
Conseil Général du DOUBS
Hôtel du Département
7 avenue de la Gare d'Eau
25031 BESANCON Cedex

Copyright by
9e Forum Européen Temps-Fréquence

Tous droits de reproduction et de traduction réservés
pour tous pays y compris les pays de la CEI