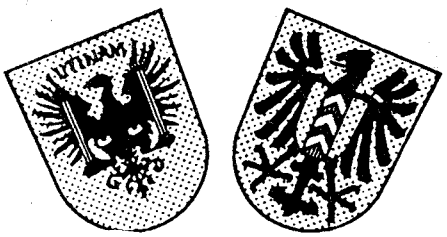


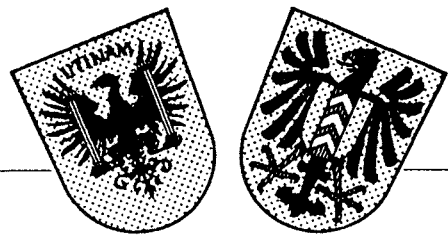
3eme Forum Européen Temps-Fréquence

3rd European Time and Frequency Forum

ACTES PROCEEDINGS

21-22-23 Mars 1989

March 21-22-23 1989



3e FORUM EUROPEEN

TEMPS-FREQUENCE

3rd EUROPEAN TIME AND

FREQUENCY FORUM

Organisé par les Communautés
Scientifiques de Besançon (FRANCE)
et de Neuchâtel (SUISSE)

Organised by the Scientific
Communities of Besançon (FRANCE)
and Neuchâtel (SWITZERLAND)

21 - 22 - 23 Mars 1989

21 - 22 - 23 March 1989

Le KURSAAL

Place Granvelle

25000 - BESANCON (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.

3e FORUM EUROPEEN TEMPS-FREQUENCE 3rd EUROPEAN TIME AND FREQUENCY FORUM

Président - Chairman ..... J.-J. GAGNEPAIN
Secrétaire Général - General Secretary ..... A. REMOND

COMITE FONDATEUR - FOUNDER COMMITTEE

R. BESSON
L.C.E.P. / E.N.S.M.M. Besançon, France

A. REMOND
OBSERVATOIRE Besançon, France

M. ECABERT
F.S.R.M., NEUCHÂTEL, SUISSE

B. SCHLUETER
OSA Neuchâtel, Suisse

J.-J. GAGNEPAIN
L.P.M.O. / C.N.R.S., Besançon, France

P. KARTASCHOFF
P.T.T. Berne, Suisse

Avec le soutien :

Ministère de la Recherche et de l'Enseignement Supérieur, Région de Franche-Comté, Conseil Général du Doubs, Ville de Besançon, C.N.R.S., D.R.E.T., C.N.E.T., S.F.M.C., E.N.S.M.M., Université de Franche-Comté, Observatoire de Besançon, et Laboratoires : L.C.E.P. et L.P.M.O.

With support of :

Fondation Suisse pour la Recherche en Microtechnique, OSCILLOQUARTZ SA (Neuchâtel), P.T.T. Suisse (Berne), Communauté de Travail du Jura (Franco-Suisse).

COMITE SCIENTIFIQUE - SCIENTIFIC COMMITTEE

C. AUDOIN - L.H.A. - France

J. BEAUSSIER - ONERA - France

R. BESSON - E.N.S.M.M. - France

M. BRUNET - C.N.E.S. - France

G. BUSCA - OSA - Suisse

H. de BOER - P.T.B. - Allemagne

H. DUCHAUSSOY - D.R.E.T. - France

J.J. GAGNEPAIN - L.P.M.O./C.N.R.S. - France

E. GRAF - OSA - Suisse

M. GRANDVEAUD - L.P.T.F. - France

B. GUINOT - B.I.P.M. - France

P. HARTL - Université Stuttgart - Allemagne

J. HENAFF - C.N.E.T. - France

P. KARTASCHOFF - P.T.T. - Suisse

D. KIRCHNER - Université Graz - Autriche

D.J.E. KNIGHT - N.P.L. - Angleterre

S. LESCHIUTTA - I.E.N. Torino - Italie

C. MAERFELD - THOMSON-Sintra DTAS - France

G. NARD - SERCEL - France

B. OEHMAN - STA - Suède

L. PROST - OFMET - Suisse

J. RUTMAN - C.E.P.E. - France

B. SERENE - E.S.A. - France

S. STARKER - DFVLR - Allemagne

P. THOMANN - OSA - Suisse

F. von WILLISEN - ASULAB - Suisse

W. ZINGG - MICRO CRYSTAL - Suisse

OVERSEAS CORRESPONDING MEMBERS

Dr. **A. BALLATO** - ERADCOM - U.S.A.

Dr. **H. HELLWIG** - NIST - U.S.A.

Dr. **K. HRUSKA** - York University - Canada

Dr. **K. NAKAGIRI** - RRL - Japon

Dr. **J. VANIER** - NRC - Canada

Dr. **F.L. WALLS** - NIST - U.S.A.

Dr. **G. WINKLER** - USNO - U.S.A.

Dr. **N. YANNONI** - US Air Force - RADC-ESE - U.S.A.

EXPOSANTS - EXHIBITORS

- **AUSTRON** - AUSTIN - TEXAS - U.S.A.

- **C.E.P.E.** - 44, avenue de la Glacière - 95100 ARGENTEUIL

- **D.R.E.T.** - 26, boulevard Victor - 75996 PARIS ARMEES

- **Q.K.P.** - 30, boulevard Galliéni - 92392 VILLENEUVE LA GARENNE

- **RACAL-DANA** - 18, avenue Dutartre - 78150 LE CHESNAY

- **S.I.C.N.** - 4, rue du Radar - 74008 ANNECY

LEGENDE DU QUARTZ ET DU CORINDON
(suite)

Il y a deux ans, un corindon, sans galon, vint à Besançon où tout est vibration
Prendre leçon d'un quartz issu d'une famille à double rotation.
Hélas ! Le maître, au cours de ses évolutions, fut frappé, par l'éclair, d'électrocution.
Ce jour-là, le corindon jura d'apprendre aux quartz à danser sans électricité.

Ce mois, revenant, pour sa mission, à Besançon où tout est vibration,
Il voit un quartz, à l'ombre d'un laboratoire, sur la route de l'Observatoire.
Paré d'électrodes dorées, surannées, un fil à chaque pied, ce cousin efféminé fait pitié.
D'un tir de laser, le corindon brise les fers du prisonnier puis, par pulsions,
l'éclaire de son rayon de lumière.
L'illuminé s'étire, frémit, palpite au rythme de la modulation, imposée par le corindon.
Il scintille de toutes ses facettes, danse, au gré des couleurs, la violette polka,
la verte bourrée, la brune sarabande, la blonde farandole, la rouge carmagnole.

C'est ainsi, qu'en 89, à Besançon où tout est vibration, le quartz, libéré de ses préjugés par la lumière du corindon, fit sa première révolution, sans le secours des électrons.

Prof. E. DIEULESAINT
Besançon, 23 mars 1989

PRESIDENTS DE SEANCES - SESSION CHAIRMEN

CONFERENCES D'OUVERTURE - OPENING CONFERENCES
J.J. GAGNEPAIN - C.N.R.S. / L.P.M.O., Besançon, France

RESONATEURS I - RESONATORS I
R. BESSON - L.C.E.P. / E.N.S.M.M., Besançon, France

MATERIAUX I - MATERIALS I
G.R. JOHNSON - Sawyer Research Products Inc., Eastlake, Ohio, USA

HORLOGES ATOMIQUES I - ATOMIC CLOCKS I
D.J.E. KINGHT - N.P.L., Angleterre

TEMPS-FREQUENCE I - TIME AND FREQUENCY I
S. STARKER - D.F.V.L.R., RFA

RESONATEURS II - RESONATORS II
H. DUCHAUSSOY - D.R.E.T., France

ONDES DE SURFACE I - S.A.W. I
C. MAERFELD - Thomson DTAS Sintra, France

OSCILLATEURS - OSCILLATORS
E. DIEULESAINT - Universtié P. et M. Curie, Paris, France

TEMPS-FREQUENCE II - TIME AND FREQUENCY II
S. LESCHIUTTA - I.E.N., Torino, Italie

RESONATEURS III - RESONATORS III
J.P. AUBRY - C.E.P.E., Argenteuil, France

TABLE RONDE - PANEL DISCUSSION
G. BUSCA - O.S.A., Neuchâtel, Suisse

BRUIT DE PHASE ET SYNTHESE DE FREQUENCE - PHASE NOISE AND FREQUENCY SYNTHESIS
H. de BOER - P.T.B., Braunschweig, France

HORLOGES ATOMIQUES II - ATOMIC CLOCKS II
C. AUDOIN - L.H.A., Orsay, France

RESONATEURS IV - RESONATORS IV
W. ZINGG - E.T.A., Grenchen, Suisse

TEMPS-FREQUENCE III - TIME AND FREQUENCY III
P. KARTASCHOFF - P.T.T., Berne, Suisse

MATERIAUX II - MATERIALS II
D. HAUDEN - C.N.R.S. / L.P.M.O., Besançon, France

ONDES DE SURFACE II ET FILTRES - S.A.W. II AND FILTERS
B.A. AULD - C.N.R.S. / L.P.M.O., Besançon, France

TABLE DES MATIERES - TABLE OF CONTENTS

CONFERENCES D'OUVERTURE

Time and frequency in Geosciences 1
B. Guinot, B.I.P.M., Sèvres, France

Environmental sensitivities of precision frequency sources 5
H. Hellwig, N.I.S.T., Gaithersburg, USA

RESONATEURS I

The susceptibility of BVA-SC resonators to photon ionization effects 11
J.J. Suter, A.G. Bates, J.M. Cloeren, J.R. Norton,
The Johns Hopkins Univ., Laurel, MA, USA
B. Schlueter, U. Peier, Oscilloquartz, Neuchâtel, Switzerland

Plotting of the vibrational distribution of a TS quartz resonator using a pulsed laser 22
R.J. Williamson, STC Technology, Harlow, Essex, U.K.

Dynamic high field relaxation of impurities in quartz crystal resonators 28
R. Brendel, L.P.M.O.-C.N.R.S., Besançon, France

The AT and SC resonators in a DC field of arbitrary direction 35
C.K. Hruska, Piezoelectricity Research Lab., York University,
Downsview, Ont. - Canada

MATERIAUX I

Simulation of the contour shape for etched quartz crystals 41
T. Leblois, P. Maître, C. Tellier, LCEP - ENSMM, Besançon, France

Technical aspects of $GaPO_4$ 50
G.Engel, P. Krempl and J. Stadler, AVL LIST GmbH, Graz, Austria

Improvement of berlinite's cristalline quality 57
E. Philippot, A. Goiffon, C. Avinens, USTL Montpellier, France
X. Buisson, R. Arnaud, O. Bignon, S.I.C.N., Annecy, France
A. Zarka, M.T. Sébastian, Université Paris VI, France

Quartz sweeping in α and β phases 61
J.P. Bachheimer, G. Dolino, Lab. Spectrométrie Physique,
Univ. J. Fourier, St Martin d'Hères, France

HORLOGES ATOMIQUES I

Millisecond pulsars and time scales 64
J.F. Lestrade, Bureau des Longitudes, France

Determining the effects of microwave power and C-field setting on the frequency of a cesium atomic frequency standard 69
S.K. Karuza, W.A. Johnson, F.J. Voit, Electronics Research Lab., The Aerospace Corp., USA

Studies of coupling to Ramsey cavity in cesium beam tube 73
L. Johnson, P. Thomann, Oscilloquartz SA, Neuchâtel, Switzerland

Long term frequency stabilization of extended-cavity 1.5 µm semiconductor lasers 81
M. de Labachelerie, K.M. Diomandé, P. Kemsu, P. Cerez, C.N.R.S. L.H.A. - Orsay - France

Application of squeezed diode laser light in rubidium atomic frequency standard 85
G.M. Saxena, National Physical Lab., New Delhi, India

TEMPS-FREQUENCE I

A dual frequency codeless GPS receiver measuring ionospheric effects and its application to time comparisons 89
M. Imae, Communications Res. Lab., Tokyo, Japon
W. Lewandowski, C. Thomas, B.I.P.M., Sèvres, France

Experience with two collocated C/A-code GPS-receviers of different type 94
D. Kirshner, Technical University, Graz, Austria
H. Ressler, Inst. für Weltraumforshung der OAW, Austria
S. Fassl, Forschungsgesellschaft Joanneum GmbH, Austria

Comparison of three GPS time receivers : a user's point of view 104
M. Granveaud, R. Tourde, L.P.T.F. - Paris - France
W.Lewandowski, B.I.P.M. - Sèvres - France

Time references from glonass satellites 121
I.D. Kitching, P. Daly, Univ. of Leeds, U.K.

RESONATEURS II

Crystal parameter measurement using the Hewlett Packard HP4195A 133
P.E. Morley, R.J.T. MARSHALL, STC Components, U.K.
R.J. Williamson, STC Technology, Harlow, Essex, U.K.

The quartz resonator automatic aging measurement facility 139
D. Beetley, General Electric CO, Largo, FL, USA

Automated measurement system for final frequency plating and characterisation of high precision SC-cut crystals 143
P. Stoermer, Ball Corp., Irvine, Ca., USA

Dual-mode temperature controlled quartz resonator 147
S. Galliou, M. Mourey, L.C.E.P. - E.N.S.M.M., Besançon, France

ONDES DE SURFACE I

The STW resonator. A new control element for high stability oscillators 152
B.A. Auld, L.P.M.O. - C.N.R.S., Besançon, France

Experimental study of planar stresses effects on surface acoustic wave devices .. 159
E. Bigler, S. Ballandras, D. Hauden, L.P.M.O. - C.N.R.S., Besançon, France

All quartz packaged SAW resonator 165
S. Calisti, P. Defranould, L. Penavaire, Thomson-Sintra ASM/DTAS, Valbonne - France
D. Hauden, L.P.M.O. - C.N.R.S., Besançon, France

A new technique for the temperature compensation of SAW oscillators 169
M.P. Cracknell, A.M. Harrison, D.J. Sharpe, STC Components, Harlow, Essex, U.K.

Investigation of SAW oscillator structures 173
P. Leblois, J. Groslambert, L.P.M.O. - C.N.R.S., Besançon, France

OSCILLATEURS

The design of VHF quartz crystal oscillators for severe vibration environments .. 180
M.K. Hobden, Marconi Electronic Devices Ltd, Lincoln, U.K.

Oscillateurs chauds 181
E. Girardet, L. Bidart, Q.K. Piézoélectronique, Paris, France

Mechanical and acoustic effects in low phase noise piezoelectric oscillators 187
E. Girardet, P. Renoult, L. Bidart, Q.K. Piézoélectronique, Paris, France

Temperature dependence of a UHF oscillator using a dielectric resonator 194
R.E.El Cheikh, M. Valentin, L.P.M.O. - C.N.R.S., Besançon, France

TEMPS-FREQUENCE II

Accurate delay calibration of satellite groundstations for two way time transfer 198
G. de Jong, VanSwinden Lab., Delft, Netherlands

Synchronization of time scales by television method using ECS satellites 204
O. Buzek, J. Cermak, URE-CSAV - Inst. of Radio Engng & Elect., Prague, Czechoslovakia
F. Cordara, P.G. Galliano, V. Petiti, Ist. Elettrotecnico Nazionale Galileo Ferraris, Turin, Italy
P. Tavella, Associazione Nazionale Galileo Ferraris Grant., Turin, Italy

Frequency stability and spectral purity requirements for ESA's data relay satellite 215
L.G. Bernier, Oscilloquartz, Neuchâtel, Switzerland
G. Busca, Observatoire cantonal, Neuchâtel, Switzerland

Lasso Experiments 220
J. Gaignebet, Groupe de Recherches de Géodésie Spatiale, Grasse, France
S. Leschiutta, Ecole Polytechnique, Turin, Italy

RESONATEURS III

Résonateurs à ondes élastiques de volume utilisant des matériaux amorphes ... 221
J. Beaussier, M. Breuzet, M. Gay, J. Uebersfeld, ONERA, Chatillon, France

Détermination et effets de modifications locales des fréquences de coupure des lames piézoélectriques planes 227
J. Détaint, J. Schwartzel, C. Joly, CNET, Dépt MCT/CMM, Bagneux, France

A 524 kHz thermally compensated quartz resonator for wrist-watch 235
C. Bourgeois, C.S.E.M. - Rech. et Dévelop., Neuchâtel, Switzerland

Influence des contraintes mécaniques internes sur l'hystérésis thermique et le vieillissement des résonateurs à quartz 238
J. Beaussier, ONERA, Chatillon, France

BRUIT DE PHASE ET SYNTHESE DE FREQUENCE

Frequency reference for both short and long term stability measurements 245
J. Groslambert, L.P.M.O. - C.N.R.S., Besançon, France

Ultra-low noise phase comparator for future atomic clocks 249
R. Barillet, L.H.A. - C.N.R.S., Université Paris-Sud, Orsay, France

The rejection problem in precise phase and frequency noise measurement 255
M. Olivier, A. El Ghazi, L.P.M.O. - C.N.R.S., Besançon, France

Coherent frequency synthesis up to 3.7 THz 259
A. Godone, E. Bava, Ist. Elettrotech. Nazionale Galileo Ferraris, Turin, Italy
M.P. Sassi, Istituto di Metrologia G. Colonnetti, Turin, Italy
C. Caldera, Politechnico di Torino, Italy

HORLOGES ATOMIQUES II

The manipulation of atomic beams by laser light pressure 264
W. Ertmer (invité), J.H. Müller, J. Nellessen, K. Sengstock, Institut für Angewandte Physik, Universität Bonn, FRG

Short-term stability of a small optically pumped cesium resonator 269
P. Thomann, F. Hadorn, Oscilloquartz S.A., Neuchâtel, Switzerland

Fluorescence anisotropy effects in an optically pumped cesium beam resonator . 274
G. Théobald, N. Dimarcq, V. Giordano, C.N.R.S. - L.H.A., Orsay, France

Progress in an optically pumped cesium beam resonator operating with a longitudinal magnetic field 277
A. Hamel, P. Petit, G. Théobald, P. Cerez, C. Audoin, C.N.R.S. - L.H.A., Orsay, France

Calculation of optical pumping with monochromatic lasers in a Cs beam 281
E. de Clercq, P. Mangin, L.P.T.F., Observatoire de Paris, France

RESONATEURS IV

Résonateur miniature à entrée et sortie optiques 293
E. Dieulesaint, D. Royer, P. de Hond, Université P. et M. Curie, Paris, France
B. Morbieu, Sté Crouzet - Div. Aérospatiale, Valence, France

Special miniaturised SC cut quartz crystal resonator for severe environments applications 296
R. Delaite, L.C.E.P. - E.N.S.M.M., Besançon, France

Topographie stroboscopique aux rayons X de résonateurs de quartz 301
Y. Zheng, A. Zarka, B. Capelle, Université P. et M. Curie, Paris, France
J. Détaint, J. Schwartzel, CNET - Dépt MCT/CMM, Bagneux, France

Influence of the definition of material constants in the study of crystal plates thermal behavior 305
B. Dulmet, R. Bourquin, L.C.E.P. - E.N.S.M.M., Besançon, France

TEMPS-FREQUENCE III

Loran-C signal synthesizer 313
E. Rubiola, Politecnico di Torino - Dip. di Elettronica, Turin, Italy

Digital frequency synchronisation with LF transmitters 319
P. Schumacher, Observatoire Cantonal, Neuchâtel, Switzerland

Service quality of a long distance telecommunication network in distributing timing signals (frequency) 326
T. Kulmala, PTL-Telecommunications Lab., Helsinki, Finland

Digital frequency synthesizer for atomic frequency standards 333
J. Viennet, C.N.R.S. - L.H.A., Orsay, France

MATERIAUX II

Origins of dislocations in single crystal cultured quartz 337
G.R. Johnson, R.A. Irvine, J.W. Foise, Sawyer Research Products Inc.
Eastlake, Ohio, USA

Tensorial characterization of the geometrical features for etched quartz plates . 346
C. Tellier, T. Leblois, L.C.E.P. - E.N.S.M.M., Besançon, France

Trace element analysis in quartz 356
B. Viard, J. Daumet, M. Khayar, H. Mérigoux, Lab. Cristallograhpie et Chimie Minérale, Université de Franche-Comté, Besançon, France

Ion beam milling : characterization and prevention of surface damage 362
N. Vialle, J.P. Aubry, C.E.P.E., Argenteuil, France

ONDES DE SURFACE II et FILTRES

SAW VCSO using hybrid technology 369
L. Penavaire, H. Gautier, Thomson-Sintra ASM/DTAS, Valbonne, France

Filtre RF faibles pertes à ondes de surface 372
J.P. Michel, C.E.P.E., Argenteuil, France

SAW filters for digital cellular radio applications 381
J.M. Hode, P. Dufilie, J. Desbois, Thomson-Sintra ASM/DTAS, Valbonne, France

A strategy for low cost IF filter for digital radiotelephone 387
J.P. Aubry, J.P. Michel, J.L. Le Corre, D. Blonde, P. de la Fournière,
C.E.P.E., Argenteuil, France

Observation of piezoelectric envelope solitons generated in the bulk acoustic wave radiation from interdigital transducers deposited on quartz plates 394
M. Planat, L.P.M.O. - C.N.R.S., Besançon, France

Time and Frequency in Geosciences

by B. Guinot

Bureau International des Poids et Mesures
Pavillon de Breteuil
F-92312 Sèvres Cedex, France

Time and frequency techniques play an essential role in many of the programs and projects for a global monitoring and understanding of our planet. The paper gives a general survey of these techniques and of their scientific and practical applications.

1. Introduction

The interest raised by geosciences has considerably increased during the last two decades and especially during the last few years. With the media opening a window on the entire world, attention is constantly attracted to the behaviour of our planet by natural catastrophes: seisms, tropical storms, droughts, floods ... We are aware of the dangers generated by a dramating increase of the world population: exhaustion of natural ressources, abnormal content of carbon dioxyde and methan in the atmosphere, ozone holes, pollutions, change of climate due to the destruction of the equatorial forests, to mention a few ones from a long list of alarming changes of our environment.

At the same time, space techniques offer the possibility to see the Earth as a whole. The observations from space platforms show us in real time the circulation of the atmosphere, of the oceans, the state of the surface of the oceans, the land use ... space geodesy, not only gives us the shape of the Earth, but also its minute deformations: our planet is alive, even at the scale of human lifetime.

An important factor in improving our knowledge of the Earth is the awareness of the interdependence of the various phenomenas. This knowledge must be global: extended to the entire planet and to all Earth disciplines. Many large scale projects for the study of the Earth have been launched or are being prepared. Our understanding of the evolution of the Earth, its atmosphere and oceans, will tell us more and more precisely what to do and what we should not do for a good management of our unique patrimony. Shall we be able to sacrifice immediate interests for the survival of the planet ? That is another problem.

Many of the techniques used in geophysical research, especially for the global phenomena, rely on an accurate dimensional measurement. For the distances to be measured, length is given by the flight time of electromagnetic signals. Long before the new definition of the metre based on a conventional value of the velocity of light (1983), space geodesy and very long baseline interferometry (VLBI) measured fictitious distances, by adopting a value of the velocity of light (not always the same value ... which was a source of difficulties). I would like to show here how time and frequency have an essential role, although ancillary, in the progresses of some branches of geosciences.

In these matters, there is no bi-univocal correspondence between the measurement technique and the geophysical study: the same technique serves several objectives and, inversely, the study of some particular phenomenon requires several technical approaches. Therefore I will first briefly describe typical methods of accurate dimensional measurements on the Earth and in space which require the time and frequency techniques. Then I will show how these measurements contribute to the knowledge of the Earth.

2. Dimensional measurements at the terrestrial scale

2.1. Doppler measurements

Some of the very first artificial satellites carried crystal clocks and emitted stable frequencies. The observed Doppler effect measured with ground clocks provides the relative positions of the satellites and a network of ground stations.

The Doppler method has been used with many geodetic satellites. It is also the method applied since 1960 for the system of positioning TRANSIT (or NNSS) of the USA. This system is still operational, but should be replaced by another t/f system, the Global Positioning System (GPS). In the most refined application of TRANSIT, a global geodesy at one metre accuracy level has been achieved. For some time, TRANSIT provided also the most precise values of the coordinates of the terrestrial pole of rotation, which moves with respect to the Earth itself.

The Doppler method is especially valuable for tracking satellites in low orbits such as those employed for missions of Earth observation. In France, the DORIS system is being experimented. Compared to TRANSIT, it is an inverted Doppler system, the measurements being made on board the satellite which interrogates automatic beacons equipped with crystal clocks.

Another application of Doppler measurements is the determination of relative velocity between satellites. For instance, in the GRM project, two satellites will be orbiting with a separation from 150 to 550 km and it is expected that their relative velocity will be measured with an uncertainty of 1 μm/s every 4 minutes. The method is a two-way Doppler which, to a large extent, is free from the oscillators instabilities; nevertheless the best available crystal clocks are required. The separation between the satellites depends on the irregularities of the gravitational field of the Earth: GRM should be able to map the geoid with an horizontal resolution of 100 m and heights to ± 0.1 m.

2.2. Distance measurements

Many of geophysical observations are based on the measurement of propagation time: for instance, the study of the Earth interior by the propagation of seismic waves. But we will restrict our discussion to the most precise applications where the time of flight of electromagnetic signal gives distances.

The two-way methods use the return signal from the target. In geosciences, microwave emissions are employed in radar altimetry from satellites which provides vertical distances, satellite to sea, to about

± 5 cm. The microwaves are also employed in the PRARE system of FRG, in development, for measuring distances satellite-Earth, with the timing equipment on board the satellite and relays on the ground; the expected accuracy is ± 10 cm. PRARE will also apply the Doppler method.

In spite of the inconvenience of not being an all weather technique, laser ranging on terrestrial and space targets is developed because it can bring a higher accuracy.

After an experimental period, some artificial satellites have been designed for laser ranging, with corner-cube reflectors. They have a spherical shape and a high density in order to minimize the non-gravitational forces, atmosphere drag, radiation pressure, which cannot be well modelled. Such satellites are already flying: STARLETTE (France, 1975), LAGEOS (USA, 1976), ETALON (URSS, 1989); other ones are in project : LAGEOS2 (Italy/USA), AJIZAI (Japan), STELLA (France).

Ground laser stations are constantly improved: their accuracy is now at the centimetre level, which requires that the time of flight be measured with uncertainties less than 100 ps. Two-color lasers will reduce further the uncertainties by measurements of the refraction and will open the millimetre era. These measurements are dependent on the frequency accuracy, but not critically: 1 mm/10 000 km = 10^{-10}. But there is a requirement of good short-term stability for the duration of the flight (~ 0.1 s).

Laser ranging is also applied to reflectors installed on the Moon by american astronauts and USSR automatic devices since 1969. The distance accuracy is of the order of 1 cm, say 2×10^{-11} in relative value.

Satellite laser ranging (SLR) and Lunar laser ranging (LLR) have some common goals: geodesy at the cm level, referred to the center of mass of the Earth, orientation of the Earth in space to less than ± 0.001". In addition SLR provides data on the geoid and LLR is a tool for studying the evolution of the Earth-Moon system.

In order to increase the number of sites positioned by SLR, mobile laser stations are developed; there is also the possibility to invert the technique, the laser borne by satellite firing on numerous ground reflectors.

The measurement of distances by unidirectional methods are much more demanding from the clocks, because they are based on synchronizations. The most imposing applications are the GPS of USA and the similar system GLONASS of USSR. For instance GPS will include 21 satellites with rubidium and cesium clocks which are kept synchronous, by software corrections, to about ± 10 ns. The simultaneous tracking of 4 satellites, or more, gives immediatly the three dimensional position and the time offset of the observer's clock. In normal use, the position is obtained with uncertainties of a few metres (as long as the system is not voluntarily degraded for unauthorized users). However the geodesists employ other modes of observation (by interferometry for instance) which allow the establishment of precise ephemerides, more precise than the broadcast ephemerides, and then an improved positioning. It is thus possible to obtain, in an operational mode, relative positions with uncertainties of 10^{-7} to 10^{-8} (10^{-8} corresponds to 1 cm over 1000 km).

2.3. Very long base line interferometry (VLBI)

VLBI is an astronomical technique which has been initially developed for mapping radiosources with high resolution (better than 0.001") for wave-lengths in the cm-dm domain. Later it has been shown that VLBI can also provide the orientation of terrestrial baselines with respect to the direction of extragalactic radiosources with uncertainties smaller than 0.001".

In positional VLBI, the basic measurement is the time difference of the arrival of quasar radiation in two stations A and B, which is measured by correlation techniques. As the angular resolution increases with distance, the length of the baseline AB is extended to several thousands km, making impossible the physical connection between the stations. The received signals are timed by local clocks, independent from each other. The requirements for the clocks are twofold:

(a) they must keep the phase of the received signals during the observation of a single source (15 minutes, for instance);
(b) they should be synchronous to better than 100 p s to take full advantage of the technique.

The latter requirement cannot be fulfilled. Fortunately, when observing several sources in different areas of the sky, it is possible to solve for the clock difference, assuming that this difference can be modelled with a small number of free parameters, the coefficient of a cubic, for instance. Only the hydrogen masers have a sufficient stability, up to 24 hours, for the positional VLBI.

VLBI provides the orientation of the Earth in space with uncertainties of the order of 0.001" (3 cm at the surface of the Earth) and the relative positions of the observing stations with an accuracy of about 1 cm. Mobile VLBI stations have been built for monitoring the coordinates changes of some critical locations of the Earth.

3. Applications to geosciences

To show the applications of the time and frequency techniques to the study of the Earth, I have selected some broad themes:

- the shape of the solid Earth,
- the shape of the geoid,
- the shape of the oceans,
- the rotation of the Earth.

However on must be aware that the progresses in these fields are strongly correlated. In particular the geoid plays a key role, because its knowledge is required for modelling the orbits of the satellites, which are, in turn, used for the other themes. The geoid is also a reference surface for the study of the oceans and object of study in itself.

3.1. Shape of the solid Earth

The terrestrial geodesy may be quite accurate locally, however the propagation of errors over long distances generates distorsions and the inconsistencies between major datums could reach several hundreds of metres. With the Doppler technique an enormous progress was achieved since it was possible in the 1970's to reach the one-metre accuracy all over the Earth. In addition, the Doppler techniques, as all the dynamical techniques which ultimately rest on the modelization of the orbits of artificial satellites and of the Moon, locate the center of mass of the Earth with respect to the observing stations.

However, it is the accuracy at the cm level which really opened a new era in geodesy and geophysics, because it gives access to the study of deformations during a reasonable fraction of human life.

The evaluation of spreading rates at ocean ridges has led to models of plate motions involving velocities of a few cm/year. However this is an average over millions of years. The geodesy at the cm level can provide a "snapshot" of these motions as well as intraplate motions. It might be interesting to compare the present velocities with the mean velocities over geological times. There is, for instance some indication that the motion of the Indian plate with

respect to the Eurasian plate is slowing down. But clearly, the detection of anomalies of the plate velocities will require much time and many observation sites to make sure that the results are not polluted by local motions.

The knowledge of the present crustal motions is required to understand what is happening at the limit of tectonic plates. It is a necessary complement to the mapping of deformations in seismic areas by local measurements over a few hundreds km, with an accuracy as good as possible, the uncertainties reduced to a few millimetres being a desirable goal. These studies require a dense grid of measuring sites. The data can be provided by purely terrestrial methods of distance measurement and by all techniques of space geodesy previously mentioned. However methods with sophisticated instruments on satellites and simple beacons on the ground are especially convenient. The immediate goal of all these measurements is the prediction of seisms.

3.2. Shape of the geoid

The geoid is the equipotential surface of the Earth which best fits the mean sea level. Its global shape was one of the first scientific results brought by artificial satellites which revealed, as early as 1959, the famous "pear shaped" figure of the Earth.

Nowadays, large scale irregularities of the geoid (resolution of about 1000 km) are derived from satellite tracking, almost exclusively by t/f techniques of section 2. When compared to a revolution ellipsoid which represents it at the best, the geoid shows bumps and holes reaching 100 metres. A much better spatial resolution is obtained over the oceans by satellite altimetry, the sea surface departing only by a few decimetres from the geoid. The GRM project would extend this resolution over the continents. Of course terrestrial gravimetry also contribute data: let us remark that absolute gravimeters timing the fall of a test mass are based on a t/f technique.

An accurate model of the geoid, in conjunction with the topography, the age of the crust, magnetism, is a surprisingly rich source of knowledge of the Earth tectonics and of understanding the Earth rheology and interior. This can be explained by the fact that the heterogeneities of the Earth's mantle have a signature on the geoid. The models of convection in the mantle, which involve the temperature and viscosity of the lithosphere and astenosphere are strongly constrained by the necessity of leading to theoretical effects on the geoid in conformity with the reality. The study of the geoid contributes to the understanding of the driving forces of the plate tectonics, of the transformations of the oceanic crust from its generation in mid-ocean ridges to its melting some 200 millions years later in subduction zones. It brings some light in volcanology and seismology in critical zones. Other spectacular results are, for instance, the discovery of a large area of anomalous lithosphere in South Central Pacific, the discovery of a sea mount which escaped soundings from the deformation of the geoid, the discovery of small scale convection and "hot spots". The laser-telemetry on LAGEOS seems to indicate that the flattening of the Earth is slowly decreasing; in conjunction with the fact that the flattening is greater than required by hydrostatic equilibrium, this result may be interpretated as the delayed answer of the Earth to the general deceleration of its rotation.

May be in the future the equipotential surfaces of the Earth's field will be mapped using the gravitational red-shift of frequencies: to an elevation of 10 cm corresponds a frequency offset of 10^{-17}.

We may add that the techniques of studying the Earth interior from the perturbations of the motion of artificial satellites has been successfully applied to Venus and Mars.

3.3. The shape of the oceans

After preliminary experiments starting in 1972, a dedicated oceanographic satellite SEASAT was launched in 1978, carrying a microwave altimeter. Unfortunately due to malfunction, only 3 months of SEASAT data are available. Nevertheless outstanding results have been obtained which led to ambitious missions: ERS1 (Space European Agency, 1990), TOPEX/POSEIDON (NASA/CNES, 1991).

A remarkable feature of these missions is that they require the ultimate advances in a wide range of space techniques, most of them being t/f techniques, in order to get the useful accuracy of vertical distance measurements:

- realization of the microwave radar and methods of analysis of the return signal;
- calibration of the radar by laser ranging, so that the total uncertainty of the vertical distance satellite/ocean be less than 4 cm;
- good modelization of the geoid, for the precise orbitography of the satellite and for use as a reference;
- satellite positioning with respect to ground station, using Doppler (TRANSIT, DORIS) and range measurements (GPS, PRARE) the goal being the vertical component of the satellite orbit to ± 10 to 15 cm.

Over one month average, the expected accuracy on mean sea level is expected to be of the order or below 5 cm. The sea currents and eddies are marked by alteration of the sea level of a few decimetres: it becomes possible to map them permanently. Other important results will be tide models, height of the waves, global field of the winds.

Let us cite some of the announcement objectives of TOPEX/POSEIDON to show the interest of such measurements:

• to conduct studies of the permanent and variable circulation of the ocean, and its interaction with the atmosphere, on global or basin scales, in order to contribute to our understanding of the role of the ocean in climate;

• to develop an improved understanding of the oceanic circulation on regional scales for periods of a few months;

• to obtain an improved understanding of the heat transported by the ocean;

• to conduct studies of tides models using altimeter observations (... geophysical interest of tidal energy dissipation, tidal loading of the lithosphere, and other such studies);

• to study the statistics of oceanic variability required by global models of oceanic circulation ...;

• to improve knowledge of the gravity field and associated geoid ...;

• to conduct studies of geophysics using either altimeter observations of the marine geoid or precise positioning of ground-based stations, contributing to a better understanding of lithospheric and mantle processes;

• to study surface wave physics, the growth or decay of waves in response to wind, and the interaction of waves with currents using altimeter observations of wave heights, surface currents, and wind.

One can note, in particular, the evident relationship of such studies with climatology, which is certainly of vital importance.

3.4. Rotation of the Earth

The motion of the rotation pole of the Earth with respect to the Earth itself was discovered in 1890. Although suspected long beforehand, the variation of the Earth rotation velocity was unquestionably proven in 1935-1950; however, its precise knowledge required the reference to atomic time, which is available since 1955.

The present measurement techniques are mainly satellite and lunar laser ranging and VLBI. The Earth rotation parameters (ERP), i.-e. the coordinates of the pole and the angle of rotation given by the Universal Time UT1, may now seem to be a by-product of accurate space geodesy, since their simultaneous determination with station coordinates and satellite position is required. However the ERP are quite important in themselves, because they are affected by the motions of the fluid parts of the Earth, atmosphere, oceans, fluid core : their variation is therefore a very complex phenomenon which is far from being well understood.

In this range of periods from a few days to a few years, the excitation of the irregularities of the ERP is dominated by atmospheric effects. Such effects contribute largely to the motion of the pole because there is a resonance between their annual component and the free nutation of the Earth, having a period of 1.2 years.

The ERP measurements may not be the best tool to understand the global behaviour of the atmosphere. But, inversely, it is conceivable to remove from the ERP the atmospheric (and may be also the oceanic) component in order to study other sources of excitation such as seisms, core/mantle interaction.

The latter excitation is poorly known and the ERP are one of the few wittnesses of what is taking place in the fluid core. As the core effects produce most probably long-term changes, over decades, centuries, a permanent monitoring of the ERP is necessary. Ideally that would require automatic observatories, in stable locations, where, as seen previously, the t/f techniques are essential.

Let us also mention, among other components of the ERP, the slowing down of the Earth rotation which is mainly due to dissipation in sea tides and to transfer of the Earth rotational energy to the Moon orbital energy. The latter will be precisely estimated from the increase of the Earth-Moon distance (≈ 1 cm/year) measured by LLR. This is a nice example of the indirect way by which a t/f technique can contribute to the knowledge of our planet.

4. Conclusion

Some of the programs of systematic monitoring of the Earth are very old (meteorology, measurement of magnetism, of the Earth rotation, seismology, ...). But many important projects are at their very beginning: monitoring of the deformations, of the sea surface, of ice cover, determination of the geoid, ...

In both cases the time and frequency techniques have often an essential role, either in improving the measurements or in opening new domains of study. The t/f techniques for this type of application have not yet reached their optimum — may be they will never do! Progresses can be considered in two directions:

- more stable and/or accurate time standards will continue to offer new possibilities;

- instruments of lesser cost, volume and mass and having a better reliability would allow to increase the number of observing sites on the ground and the capability of space stations.

5. Acknowledgements

No references are given: they would have been too numerous. Nevertheless, I would like to mention that I extensively used documents from CNES (France), ESA, GRGS (France), IFAG (FRG), JPL (USA), NASA (USA) [I hope that the reader knows these acronyms ...] and other organizations.

I also thank colleagues for loan of slides for the presentation of this communication at the Forum: G. Balmino, F. Barlier and the documentation service of the Observatoire de la Côte d'Azur, D. Kirchner, W. Schlüter.

ENVIRONMENTAL SENSITIVITIES OF PRECISION FREQUENCY SOURCES

Helmut Hellwig

National Institute of Standards and Technology
Gaithersburg, Maryland 20899

Abstract

This paper, presents issues that relate to the environmental sensitivity of the frequency of precision frequency sources. As a definition of precision sources, an aging rate of 10^{-9} per day or less is used. Included in this definition are high performance quartz crystal oscillators, rubidium gas cell devices, cesium beam standards and hydrogen masers. A summary is given of the typical environmental sensitivities of these four classes of precision frequency sources. This is followed by a discussion of interrelationships between two or more environmental stimuli and of the limits of measurability of the frequency changes imposed by the basis noise performance of the frequency source under test. The paper concludes with a review of standardization efforts in the area of time and frequency. This includes the IEEE Standard(1) "Standard Definition of Physical Quantities for Fundamental Frequency and Time Metrology," approved at the end of 1988. This new Standard was developed under the IEEE Project Authorization 1139(2). Also reported are the efforts towards a Standard characterizing environmental sensitivities which has been authorized by the IEEE as Project 1193. This paper is motivated by this ongoing standardization effort (3).

Introduction

It is self-evident that environmental sensitivities of precision frequency sources in most, if not all, applications limit their performance to levels inferior to those given by fundamental physical design principles. Therefore, the characterization and knowledge of the size of the effects and their correct measurement are of utmost importance to any user of precision frequency sources. In this paper, precision frequency sources are defined as those oscillators which feature frequency drifts or aging rates of less than 10^{-9} per day. This definition admits the very high performance quartz crystal oscillators, as well as cesium beam and rubidium gas cell atomic standards and hydrogen masers. Environmental sensitivities are of overriding importance in field use, such as on board of aircraft or ships or in other mobile application. However, even in protected laboratories, environmental sensitivities can become of considerable concern to the users such as the laboratories generating the time scales which form the Universal Coordinated Time (UTC).

This paper addresses the topic of environmental sensitivities in three parts. The first part is a summary of the environmental sensitivities of precision frequency sources. In tabular form the principal environmental stimuli are listed:

acceleration & gravitational potential
barometric effects
humidity
magnetic field
radiation
retrace (intermittent operation)
and temperature

Also considered are the typical stability performance characteristics in terms of frequency instability at one second averaging time as well as the best achieved stability (flicker floor) and the typical frequency aging per day.

The second part reviews selected areas where convolution of two or more transducing mechanisms has caused difficulties in characterizing and measuring the effects, such as interplay between humidity and temperature effects. A discussion is included which illustrates fundamental measurement limitations stemming from the basic, ultimate stability or noise performance of the devices. These limitations impose boundaries on the ability of the user to specify and, through testing, verify environmental sensitivities.

In the third and last section of this paper, a summary is given of the standardization efforts of the Standards Coordinating Committee 27 on time and frequency of the IEEE. These efforts resulted, at the end of 1988, (a) in the approval of the first IEEE Standard on time and frequency entitled "Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology", (b) in the approval of a new project aimed at characterizing environmental sensitivities. This paper is part of the effort to achieve consensus on this latter, new IEEE project.

Typical Environmental Sensitivities of Precision Frequency Sources

The following four tables summarize the data which, typically, are encountered with the four types of precision frequency sources: quartz crystal oscillators, rubidium gas cells, cesium beam standards, and hydrogen masers. The values reported in the tables are taken from data and procurement specifications available from the various manufacturers of the devices(4). Since there are several vendors for each of the four types of frequency sources, the values are rounded and represent certain compromises. A comprehensive discussion of these devices and the underlying physical and engineering principles can be found in Reference (5). Each table lists a number of entries with the following definitions:

Frequency stability - given is the square route of the two sample or Allen variance for sampling times of one second and one hundred seconds, as well as the best observed value (flicker floor) which in most cases is reached at averaging times of the order of a day. A measurement bandwidth of 1kHz was assumed.

Drift - given is the typical fractional (normalized) frequency change per day resulting from aging mechanisms.

Retrace - listed is the fractional frequency difference between a frequency source turned on from a cold start after sufficient warm-up, compared to its frequency during the last continuous operation previous to the cold start.

Temperature - listed is the fractional frequency change for a one degree Kelvin change within a temperature region around room temperature.

Acceleration - given is the frequency change encountered when changing a steady state acceleration of one g by one g in a specific direction.

Barometric change - given is the frequency change as a result of going from atmospheric pressure to vacuum.

Humidity - given is the frequency change when changing the environment from a zero humidity level to a level of 100% (but not condensing).

Magnetic field - given is the frequency change encountered as a result of applying a homogeneous magnetic field of one gauss (1 gauss = 10^{-4} tesla) to a device operating in a specific orientation with respect to the earth's magnetic field.

Radiation - given is the frequency change encountered after applying a dose of 10^{-4} rad of x-ray radiation.

In all the data given, we are assuming quasi-static conditions; i.e., we are making the assumption that the frequency source has stabilized and is in equilibrium with its environment before applying the environmental change or stimulus. Furthermore, after the application of the stimulus, the frequency standard again has reached a new equilibrium with its environment. It is obvious that these idealized conditions can never be truly met because they would not only require infinite time periods before and after application of the stimulus, but also an infinitely slow application of the stimulus itself. Therefore, to meet the challenge in developing standard test methods and uniform descriptions of the stimuli and the resulting effects, we have to agree on acceptable procedures approximating the quasi-static condition.

Sometimes, characteristics are important to the user which relate to a very dynamic environment; i.e., it may be important how the frequency source actually reacts or even tracks a particular temperature or humidity or radiation or acceleration profile. In such cases, test methods probably have to completely conform to the required profiles. In fact, experience has shown that extrapolation from one dynamic profile of an environmental stimulus to an altered profile typically are not meaningful and the results of such extrapolations are unreliable.

In addition to the typical values given in the four tables, each table also lists, in narrative form, the basic sensitivity mechanism.

Measurement Limitations

When specifying, characterizing, and evaluating precision frequency sources, we encounter three types of limitations: (a) fundamental limitations related to the basic frequency instability (noise) performance of a device under test, (b) limitations relating to the dynamics of the applied stimuli; e.g., measurements made clearly not under equilibrium conditions, and (c) convolution between two or more stimuli and/or physical transducing mechanisms. In addition, many parameters exhibit non-linear response. For example, near activity dips, quartz oscillators show gross changes in the magnitude and sense of the temperature coefficient. Magnetic field sensitivities are inherently non-linear and may even depend on temperature.

a) Fundamental Measurement Limitations

The following may appear trivial but, since too often test procedures and specifications are deficient in this regard, it is important to mention: It is self-evident that environmental sensitivity can neither be specified nor measured more precisely than is permitted by the fundamental frequency stability performance of the frequency source. Therefore, when writing test procedures or specifying frequency sources, great care must be taken to avoid conditions and statements which are incompatible with the fundamental performance of the source as, for example, represented by a plot of the square root of the two sample variance as a function of the sampling time. For example, if the flicker floor of a frequency source is 10^{-13}, it is not proper to specify a measurement requirement for certain environmental sensitivities at the 10^{-14} level.

Furthermore, one has to realize that the flicker floor is the best possible frequency stability of the device and, with the exception of quartz, is typically is reached only after long sampling time periods (hours to days). Test procedures must take this into consideration and allow sufficient elapsed time to perform these measurements before changes in environmental stimuli are mandated. In other words, the achievement of the quasi-static condition discussed above becomes an extremely time-consuming process if measurements are performed near the ultimate stability limitations of the device. If measurements and changes are to be performed faster, then the corresponding frequency stability at the allowable sampling time must be considered. For example, if the above-mentioned device with a stability floor of 10^{-13} has a stability of 10^{-10} at sample times of one second, then changes due to incrementally increasing stimuli levels in one second intervals cannot be measured with a greater precision than 10^{-10}. We note, that one can overcome this limitation somewhat by repeating and using correlation techniques.

In general, the use of statistical measures in non-equilibrium situations should be avoided. The following turn-on specification is an example of a specification which, when subjected to measurement verification, will predictably lead to ambiguous

(non-repeatable results): "The frequency stability in terms of the square root of the two sample variance shall be 10^{-10}, fifty seconds after turn-on." It is a fact that one simply cannot measure a statistical quantity with statistical confidence at a particular time or epoch. The test engineer is faced with an impossible dilemma: Should he perform a statistically not valid measurement (a snapshot measurement) with the consequence of non-repeatability in the second or third or nth test? should he assure statistical validity by making repeat measurements, which would put in question the particular 50 second timing requirement? Should the sampling start at 50 seconds, in which case the actual measurement time is later than 50 seconds? Should he spread the sampling around the 50 seconds mark, in which case earlier data which are not required to meet the specification are included in the sample? And, in any case, something is being averaged while a dynamic process is still causing changes; i.e., the oscillator has not settled down into an equilibrium condition.

b) Testing and Specifying Non-Equilibrium Conditions

In general, the condition of a non-equilibrium measurement occurs when the time constants involved in the transfer of the environmental stimulus into the sensitive components of the frequency source are longer than the timing of the measurement following the stimulus. For example, the temperature of the frequency source is raised by 20 degrees. The thermal time constant of a crystal oscillator could easily be two hours due to the insulation and thermal mass of the quartz crystal oven.

If a measurement is to be taken thirty minutes after termination of the temperature stimulus, one will get a result which is highly dependent on the particular manner and time profile of the temperature change. It will be different whether it was an abrupt 20 degree change, or a ramp, or any other imaginable profile. In addition, it is of obvious concern whether the temperature coupling of the oscillator to the temperature chamber was excellent or more remote (e.g., direct physical base plate contact to a high capacity heater or just air circulation). Furthermore, if close coupling and fast ramping was possible, then the initial condition of the oscillator will also be of importance and affect the actual results of the stipulated measurement, 30 minutes after cessation of the stimulus.

Related concerns can easily be imagined for all other environmental stimuli ranging from retrace (length of off-time, condition prior to turn-off, condition during the off-time, etc.) to acceleration, humidity, radiation, etc. As indicated above, in cases where a non-equilibrium measurement is essential, many more conditions than just the specification of the stimulus and its timing are important. They include:

the original conditions of the frequency source under test;

the particular profile of the application of the stimulus;

the way the stimulus is applied to the oscillator (the degree of coupling of the stimulus to the oscillator);

the question to what degree statistical measures such as the Allen variance, are useful in a non-equilibrium, dynamic situation and in situations which involve non-linear responses to stimuli.

c) Interaction of Two or More Environmental Stimuli

In an ideal measurement situation, it is desirable to have the environmental stimulus applied in such a way that it uniquely and exclusively affects the frequency source under the test. Often, this is not possible and, frequently, it is not even obvious that other effects may be present. The following is a list of the most commonly encountered situations where the application of a particular stimulus is leading to the reaction of the frequency standard to an associated stimulus. The associated stimulus, in the majority of cases, is temperature.

1. Acceleration (tip-over test) — An instrument is tested for acceleration by turning it 180° in the earth's magnetic field. If done properly, this is a quasi-static test and should yield accurate measurements as a result of the applied change of two g. However, the positioning upside down of the instrument invariably changes its internal heat distribution. A significant portion of heat transfer, in most frequency sources, is through air convection. The convective heat transport will be altered significantly if the force of gravity is reversed. Thus, tip-over tests, more often than not, are really tests of temperature sensitivity due to the convective heat transport component.

2. Radiation — If a radiation dose is applied in a relatively short time period (short compared to the thermal time constant of the unit under test), then the heating associated with the radiation loading will cause thermal effects in the unit, especially when measurements are made within the time period of the thermal time constant of the unit. Thus, frequency excursions observed in response to a radiation shock very frequently are reactions to thermal shock rather than radiation shock.

3. Humidity — In addition to affecting conductivity and dielectric constants, changes in humidity will change the internal heat conductance of the unit, especially as regards convective coiling and heat transfer (6). In addition, the time constants for effective change in the humidity levels within the unit are not necessarily equal to its thermal time constants, causing potentially very erratic results when tests are performed in non-equilibrium conditions as regards humidity

or temperature or both. The unit may have critical areas which are inadequately vented and, as a result, involve local time constants which may be substantially different than the global time constants of the unit. After all, humidity is only a meaningful concept when the area under consideration is in complete thermal equilibrium.

4. Temperature -- In temperature testing, a temperature range is being administered ranging from a certain specified low temperature (typically below room temperature). Various forms of ramping are applied and the change can take the unit all the way to its upper temperature limit from an initial "soaking" at the lower temperature limit; or the reverse may be prescribed. The temperature changes may be administered through direct physical (base plate) contact with the heating or cooling source, or through general placement in a refrigerator/oven with a combination of radiative and convective cooling and heating. Obviously, significant differences can occur between the ramping of the heating/cooling source (the oven/refrigerator) and the unit itself, depending on the effectiveness of heat transfer. Thus, at any given moment within the temperature profile, the temperature of the unit, of its internal components, and its outside surfaces may be different among each other, and different from that of the heating or cooling source itself.

   In addition to the presence of substantial temperature gradients, yet another phenomenon often is not being considered seriously. When the unit is taken from low temperature to high temperatures or the reverse, significant changes in humidity level will occur in the oven chamber and in the unit internally (especially if the oven is not humidity controlled). Worse, actual condensation and even the formation of ice can occur within the unit, especially when a cold soaked unit is taken up to higher temperatures. Condensation of moisture can have serious effects on the functionality of the device, including but not limited to, electrical short, infiltration of moisture into sensitive electronics, and even permanent, lasting changes or damage of electronic or physical components.

   Condensation effects can be eliminated with testing under vacuum. However, when ramping an operating unit through 1/10 to 1/1000 of atmospheric pressure, electrical discharges may occur within the unit. This may affect its performance as well as inflict permanent damage to sub-assemblies and electronic components.

Standardization Efforts

Measurements are only of value if they are repeatable and can be compared with the results of others. Systems designer, manufacturer, and user desire a common baseline to compare specified performance with actual results, and to assure compatibility with other systems or networks particularly with those having a common interface. One prerequisite for this is the agreement on the basic quantities which describe the performance of frequency sources in the frequency and time domain. They allow the quantification of frequency instability and time error as well as the ability to predict performance in the sense of timekeeping. The following quantities have been in use in a consistent way for about two decades: frequency, phase, and time departure; frequency, phase and time instability; two sample (Allen) variance and the corresponding deviation; time interval error; confidence limit. These quantities have been defined and confirmed by the recently approved IEEE Standard (Project designation: 1139) which lists the mathematical and physical definitions of these quantities, as well as notes on application and interrelationships between the frequency and time domains.

The above listed quantities, as defined by the new IEEE Standard, also must be used to describe environmental sensitivities. The IEEE has approved a project (Project designation 1193) to draft an IEEE Standard aimed at defining standard procedures to measure environmental sensitivities of precision frequency sources. Its primary purpose will be the assurance of consistency and repeatability of environmental sensitivity measurements, and the portability of results on particular frequency sources between the various segments of the time and frequency community.

It was the intent of this paper to discuss some of the challenges which lie ahead in the task of quantifying environmental sensitivities and to describe some observations about correlations and interrelationships. This paper also is intended to solicit input from the many practitioners and from laboratories engaged in the testing, verification, and acceptance of precision time and frequency sources.

References

(1) Information about Standards issued by the Institute of Electrical and Electronics Engineers (IEEE) can be obtained from the IEEE Standards Office, 345 East 47th Street, New york, NY, 10017-2394

(2) Allan, D.; Hellwig, H.; Kartaschoff, P.; Vanier, J.; Vig, J.; Winkler, G.; Yannoni, N. "Standard Terminology for Fundamental Frequency and Time Metrology." Proc. 42nd Annual Symposium on Frequency Control, published under 88CH 2588-2 by the IEEE, 445 Hoes Lane, Piscataway, NJ 08854

(3) The author is Chairman of the Technical Committee 3 on Time and Frequency of the Instrumentation and Measurement Society of the IEEE, as well as Chairman of the Standard Coordinating Committee 27 on Time and Frequency, which was formed with the cooperation of the Ultrasonic Ferroelectricity and Frequency Control (UFFC) and Microwave Theory and Techniques (MTT) Societies of the IEEE.

(4) Austron; Ball Corp. (Efratom); EG&G; Frequency Electronics; Frequency and Time Systems; Hewlett-Packard; Oscilloquarz; Piezo Crystal; Sigma Tau;

(5) Hellwig H. "Microwave Frequency and Time Standards"; Chapter 10 in "Precision Frequency Control" E.A. Gerber and A. Ballato, Eds., Academic Press. 1985.

(6) Walls, F.L. "The Influence of Pressure and Humidity on the Medium- and Long-Term Stability of Quartz Oscillators" Proc. 42nd Annual Symposium on Frequency Control, published under 88CH 2588-2 by the IEEE, 445 Hoes Lane, Piscataway, NJ 08854.

Figure 1 QUARTZ CRYSTAL OSCILLATOR

	Stability		Drift	Retrace	Temp.	Accel.	Barom.	Humidity	Magnetic	Radiation
	1 s	floor	per day		(per K)	(per g)	(0 to atmos. pressure)	(0 to 100%)	(per gauss)	(per krad)
typical value	10^{-11}	10^{-12}	10^{-10}	10^{-9}	10^{-10}	10^{-9}	10^{-10}	10^{-9}	10^{-11}	upset plus temporary & permanent worsening of the drift rate
best value	10^{-13}	10^{-13}	10^{-12}	10^{-10}	10^{-12}	10^{-10}	none	none	none	
basic mechanism	- -	- -	changes in the quartz crystal	changes in the quartz crystal	crystal lattice property	crystal mounting & lattice property	crystal enclosure and mounting	electronics and convective cooling	none	changes in crystal lattice (upset mostly a thermal effect)

Figure 2 RUBIDIUM GAS CELL OSCILLATOR

	Stability 1 s	floor	Drift per day	Retrace	Temp. (per K)	Accel. (per g)	Barom. (0 to atmos. pressure)	Humidity (0 to 100%)	Magnetic (per gauss)	Radiation (per krad)
typical value	10^{-11}	10^{-13}	10^{-12}	10^{-11}	10^{-12}	upset, 10^{-12}	10^{-11}	10^{-12}	10^{-12}	upset, none
best value	10^{-12}	10^{-14}	10^{-13}	10^{-12}	10^{-13}	10^{-14}	10^{-13}	none	10^{-13}	none
basic mechanism	- -	- -	gas cell physics	gas cell physics	gas cell physics, microwave power	gas cell physics (crystal for upset)	gas cell dimensions, cavity dimension	electronics and convective cooling	atomic transition	none*

Figure 3 CESIUM BEAM TUBE STANDARD

	Stability 1 s	floor	Drift per day	Retrace	Temp. (per K)	Accel. (per g)	Barom. (0 to atmos. pressure)	Humidity (0 to 100%)	Magnetic (per gauss)	Radiation (per krad)
typical value	10^{-11}	10^{-13}	10^{-15}	10^{-12}	10^{-13}	upset, 10^{-13}	10^{-12}	10^{-12}	10^{-12}	upset,none
best value	10^{-12}	10^{-14}	10^{-16}	10^{-13}	10^{-14}	upset, 10^{-13}	10^{-15}	none	10^{-14}	none
basic mechanism	- -	- -	electronic parameters, beam detection	electronic parameters	electronic parameters	beam position (crystal for upset)	design dependent	electronics and convective cooling	atomic transition	none

Figure 4 HYDROGEN MASER STANDARD

	Stability 1 s	floor	Drift per day	Retrace	Temp. (per K)	Accel. (per g)	Barom. (0 to atmos. pressure)	Humidity (0 to 100%)	Magnetic (per gauss)	Radiation (per krad)
typical value	10^{-12}	10^{-14}	10^{-14}	10^{-12}	10^{-13}	upset, 10^{-13}	10^{-12}	10^{-12}	10^{-12}	upset,none
best value	$5X10^{-13}$	$5X10^{-16}$	10^{-16}	10^{-13}	10^{-14}	upset, 10^{-15}	10^{-15}	none	10^{-14}	none
basic mechanism	- -	- -	electronic parameters, wallshift	electronic parameters	cavity tuning and wallshift	beam position (crystal for upset)	design dependent	electronics and convective cooling	atomic transition	none*

The Susceptibility of BVA–SC Resonators to Photon Ionization Effects

J. J. Suter, A. G. Bates, J. M. Cloeren, J. R. Norton

The Johns Hopkins University
Applied Physics Laboratory
Laurel, MD USA

B. Schlueter and U. Peier
Oscilloquartz, SA
Neuchatel, Switzerland

I. Introduction

Quartz crystal oscillators are employed in satellites for the generation of precision time and frequency signals. These crystal resonators must be insensitive to ionizing radiation in order to meet their performance requirements. The susceptibility of quartz crystal resonators to ionizing radiation (proton, electron and gamma) manifests itself as frequency shifts. Previous experiments have indicated that these shifts are dependent on total radiation dose, dose rate, particle flux or fluence and kinetic energy levels [1, 2 and 3]. Some recent investigations did indicate that the frequency susceptibility of quartz crystals to lower levels of radiation (<10 rad [Si]) is not correlated to bulk effects, i.e., impurity levels in the quartz crystal blank. The present work further substantiates this hypothesis by presenting the results of radiation tests on BVA SC quartz crystal resonators. These resonators had above average aluminum defect center concentrations (≈18 ppm) but exhibited some of the lowest radiation sensitivities on a per accumulated dose basis. It is apparent from these tests that electrodeless resonators, like the BVA, have good radiation responses for simulated low-earth orbit radiation levels (0.6 rad for 42 minutes). The objective of our current study is to investigate the radiation sensitivity of SC cut BVA resonators.

II. Radiation Study

To quantify the effects of radiation on quartz crystal resonators, materials studies have been conducted on the basic radiation susceptibility mechanisms in quartz crystals [1]. These studies have focused on the quartz crystal resonator's susceptibility to cosmic rays, electron, proton, neutron, and X-ray radiation. Based upon its unique interaction mechanism with matter, each of these forms of radiation interacts differently with the quartz crystal. In general, the radiation a satellite encounters in space is due to trapped

electrons and protons in the Van Allen belts and is enhanced by activity on the sun. It has been shown that trapped electrons in the Van Allen belts do not contribute significantly to the total radiation doses received by instrumentation located on the TOPEX spacecraft behind 5 gm/cm^2 of equivalent aluminum shielding [4]. Therefore, our study on the susceptibility of quartz crystal resonators to low-earth radiation focused on proton radiation induced effects. Specifically, we have investigated the effects of ionizing radiation levels, encountered by the TOPEX spacecraft, on the stability of quartz resonators. Typically, the doses accumulated by a quartz crystal blank behind 5 gm/cm^2 of equivalent aluminum shielding is about 0.6 rad [Si] and dose rates not exceeding 0.014 rad [Si] per minute. This study reports on the effects of these dose rates on BVA-SC resonators.

Protons encountered in space by quartz crystal oscillators lose their energy mainly by Coulomb interactions with the electrons in a solid. Photons emitted by a Cobalt 60 source and interacting with the quartz crystal resonator have been shown to cause the same magnitude of frequency shift in quartz crystal resonators as protons [Ref. 2]. This effect can be explained by the fact that the probability of Compton collisions of 1.25 MeV photons from a Cobalt 60 source depends only on the density of the absorber and not on its atomic number. Therefore, the frequency susceptibility of quartz crystal resonators to low-earth orbit radiation can be emulated by gamma rays from Cobalt 60 sources.

The low-earth orbit radiation spectrum typically consists of cosmic rays, trapped protons and electrons, and solar flare protons [Ref. 4]. The density of the cosmic rays is low and therefore not a hazard to quartz crystal resonators. It was also shown that the contributions of electrons to the total radiation dose received by an orbiting satellite is several orders of magnitude smaller than the proton radiation behind a specified amount of shielding. Therefore, protons are the main radiation hazard for spacecraft electronics and, in particular, quartz crystal resonators. The contributions of solar flare protons generated during solar activity can be predicted on a yearly basis, allowing the technical activities of a spacecraft mission to be scheduled around these events. However, a satellite is intermittently exposed to trapped proton radiation in the Van Allen belts. The low dose rates allow for the accumulation of only tenths of a rad in the crystal blank over one orbit and, therefore, it is important to study the effects of low doses ($<$1 rad) on the radiation susceptibility of quartz crystal resonators. It should also be pointed out that the accumulation of proton radiation during each orbit takes place over certain intervals when the satellite traverses specific dense radiation zones within the Van Allen belts, while a certain amount of time per orbit is spent in radiation free environments. For the TOPEX satellite, which is scheduled to be launched in 1991, the maximum dose behind 5 gm/cm^2 of equivalent aluminum shielding is approximately 8000 rad [Si] over its three year mission (see Fig. 1). The energy spectrum of the TOPEX orbit is shown in Figure 2.

The quartz crystal in the spacecraft will accumulate the radiation over specific intervals depending on the time spent in radiation zones. Figure 3 illustrates the magnitude of these radiation zones as a function of orbit time, the South Atlantic Anomaly (SAA), represents the most intense radiation zone. The peak value of the accumulated dose behind 5 gm/cm^2 of aluminum shielding in the South Atlantic Anomaly is about 10^{-3} rad/sec; the TOPEX

spacecraft spends approximately 42 minutes of its orbit in this radiation anomaly. The South Atlantic Anomaly represents an average dose rate of approximately 0.6 rad for 42 minutes or 0.014 rad/min. Since the relative accumulated doses for the other zones are at least two orders of magnitude lower, it is safe to assume that most of the radiation accumulated in a quartz crystal resonator is contributed by the South Atlantic Anomaly. In other words, the spacecraft spends the remainder of the orbit ($\approx$70 min) in a relatively radiation free environment.

Thus, it is important to study the effects of this annealing on the frequency stability in alpha quartz. It is readily apparent that one cannot simulate the effects of low-earth orbit radiation on quartz crystals by high dose and high dose rate radiation, because such tests do not take into account the radiation free intervals of actual low-earth orbits, nor do they expose the crystal to the correct low dose rates which have been shown to activate different radiation susceptibility mechanisms in alpha quartz. Therefore, the present study was conducted to examine the radiation susceptibility of SC BVA resonators to low dose rate ionizing radiation.

III. Resonator Studies

For the resonator susceptibility test, quartz crystal resonators were exposed to 1.25 MeV gamma radiation from a Cobalt 60 source (refer to Figure 4). The dose rates at the quartz crystal blank were kept at 0.014 rad [Si] per minute during all tests. The small attenuation of the gamma radiation through existing oscillator shields was taken into account in each experiment so that a consistent dose rate at the quartz crystal was established. The quartz crystal resonators were of the electrodeless BVA SC type. They were manufactured from cultured premium Q swept quartz. The 5.0 MHz sinusoidal output of the resonators was applied to a mixer which was made part of a frequency stability data acquisition system which has been described in several previous papers [Refs. (2) and (3)].

Four BVA-SC resonators were exposed to the Cobalt 60 radiation. The radiation exposures consisted of a 42 minute exposure to 0.6 rad [Si] at a dose rate of 0.014 rad/min followed by a 70 minute annealing interval. These tests were repeated several times to study long term trends. Typical results of these tests are shown in Figs. 5 and 6 and summarized in Table I. It can be seen that on a per rad basis, the radiation sensitivity varies quite substantially. Specifically, one resonator (S/N 1) did not show any radiation response. This phenomenon is not easily explained since all resonators were manufactured from the same manufacturer's lot of swept quartz. Furthermore, no correlation could be accounted for between radiation susceptibility and the resonator's individual electrical parameters (refer to Table I).

Figure 5 shows a typical response of the BVA-SC resonators to an exposure of 0.6 rad [Si] over 42 minutes. The resonator has a negative frequency shift of -1.6×10^{-11} (df/f per rad). A small recovery occurs after the 42 minutes of exposure and continues during the 70 minute annealing interval. Successive exposures of 0.6 rad [Si] result in similar frequency shifts. After five successive exposures and annealing periods, the resonator seems to recover but exhibits a permanent frequency offset. This phenomenon has

been observed in many different classes of crystal resonators [Refs. (2) and (3)]. All BVA-SC resonators were tested as manufactured (not pre-conditioned) and exhibited negative frequency shifts when exposured to gamma radiation (refer to Table I).

Preconditioning of quartz crystal resonators with 20 krad [Si] of gamma radiation has been shown to reduce their irradiation susceptibility [Refs. (1) and (2)]. This pre-irradiation of an assembled quartz crystal resonator supplies enough energy to overcome the binding energy of some cations located interstitially within the resonator. The cations can, therefore, freely move in the crystal resulting in a quartz crystal having a lower entropy. This more perfect crystal should be less susceptible to ionizing radiation than a crystal which has not been pre-irradiated.

An experiment was conducted to see if this hypothesis is also true for low radiation exposures of quartz crystal resonators. BVA-SC resonator serial number 23 was, therefore, preconditioned with 20 krad of gamma radiation. The radiation tests were repeated using the same 0.6 rad [Si] exposure followed by 70 minutes of annealing. The response of this resonator is shown in Figure 6. The radiation sensitivity of this preconditioned resonator is 1.3 X 10^{-11} (refer to Table I). The preconditioning caused a sign reversal in the frequency shift. However, the magnitude of the frequency shift is quite similar to that of the resonator before preconditioning. This experiment showed that the current level of preconditioning (20 krad) does not dramatically reduce the resonator's radiation sensitivity. One may conclude that 20 krad of preconditioning may not be the optimum pre-conditioning level. Further experimentation will be required to fully understand the relation between preconditioning dose and its effect (if any) on the motion of cations in the quartz resonator. The sign reversal of the frequency shift after preconditioning still indicates that bulk effects in the quartz resonator might still be affected. Additional experiments will need to be conducted to study this phenomenon. Unfortunately, theoretical models of the piezoelectric effect in quartz resonators do not explain this sign reversal. Only further experimentation can presently offer some insight on the true underlying causes of these complex phenomena.

IV. References

1. J. C. King and D. R. Koehler, "Radiation Effects in Resonators" in Precision Frequency Control, E. A. Gerber and A. Ballato eds, Academic Press, NY 147-158 (1985).

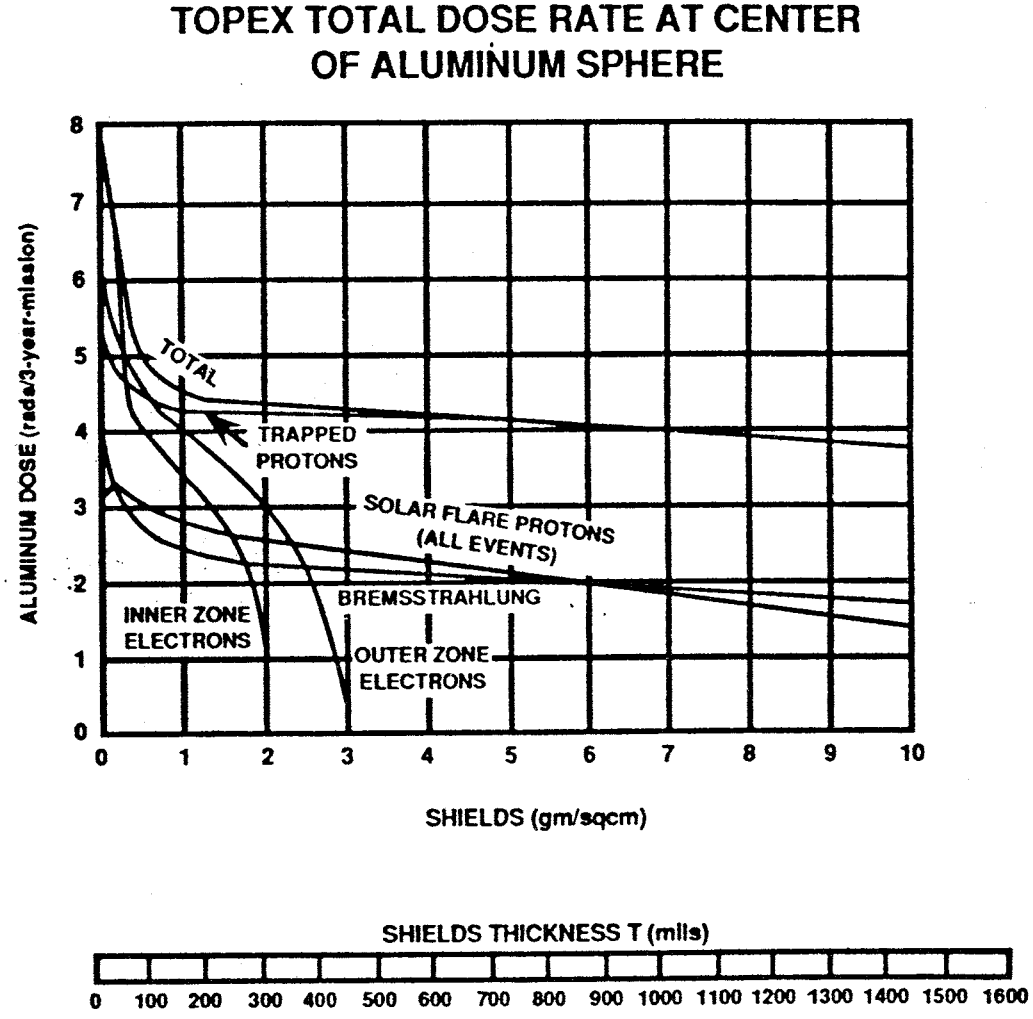
2. J. R. Norton, J. M. Cloeren, and J. J. Suter, "Results from Gamma Ray and Proton Beam Radiation Testing of Quartz Resonators," IEEE Transactions on Nuclear Science, NS-31(5), 1230-1233 (1984).

3. J. J. Suter, A. G. Bates, and J. M. Cloeren, "The Radiation Susceptibility of BVA-AT Resonators to Simulated Medium Earth Orbit Radiation," Proc. 2nd European Frequency and Time Forum, Neuchatel, Switzerland, 679-686 (1987).

4. P. A. March et al, "TOPEX Radiation Environment," JPL Document D-2116 (1985).

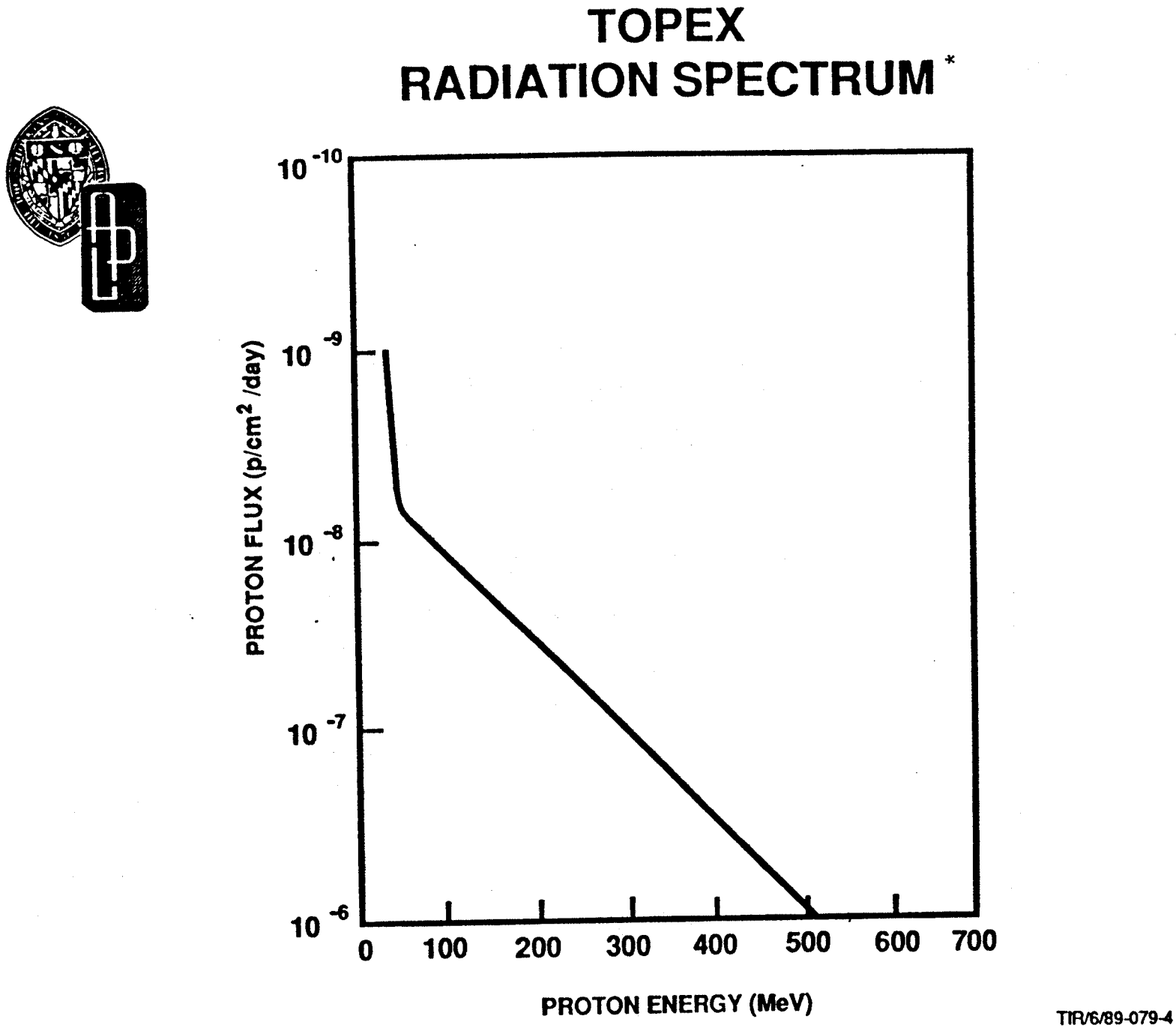
Figure 1.

TOPEX RADIATION DOSE*



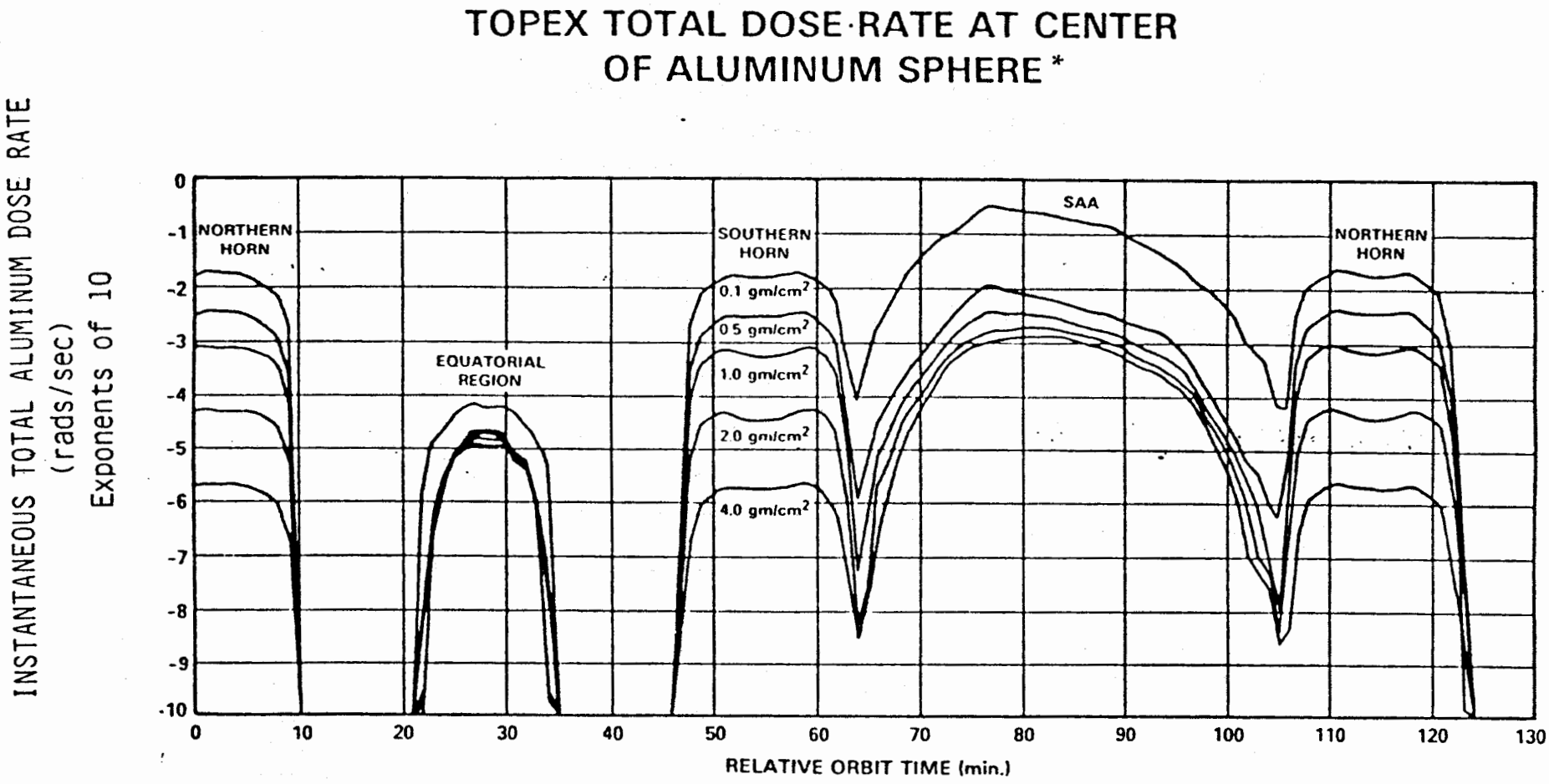
* After Reference 4.

Figure 2.



* After Reference 4.

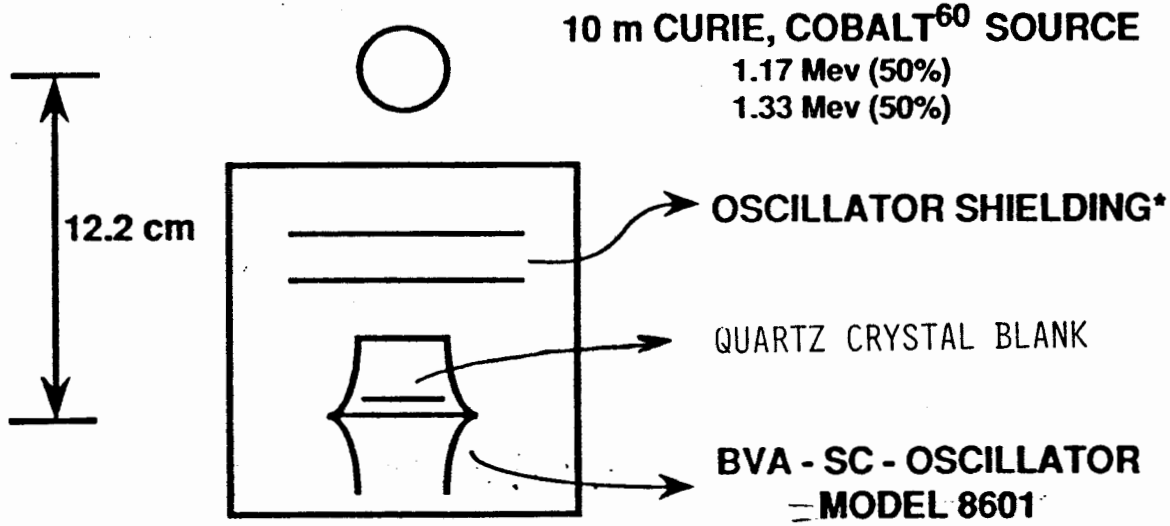
Figure 3.



* After Reference 4.

Figure 4.

EXPERIMENTAL CONFIGURATION



* 2.77 gm/cm^2 EQUIVALENT ALUMINUM (INTENSITY OF GAMMA RAYS AT BVA QUARTZ CRYSTAL BLANK: 94%)

o DOSE RATE AT CRYSTAL BLANK: 0.014 rad (Si) / MINUTE

TIR/6/89-079-1

Figure 5.

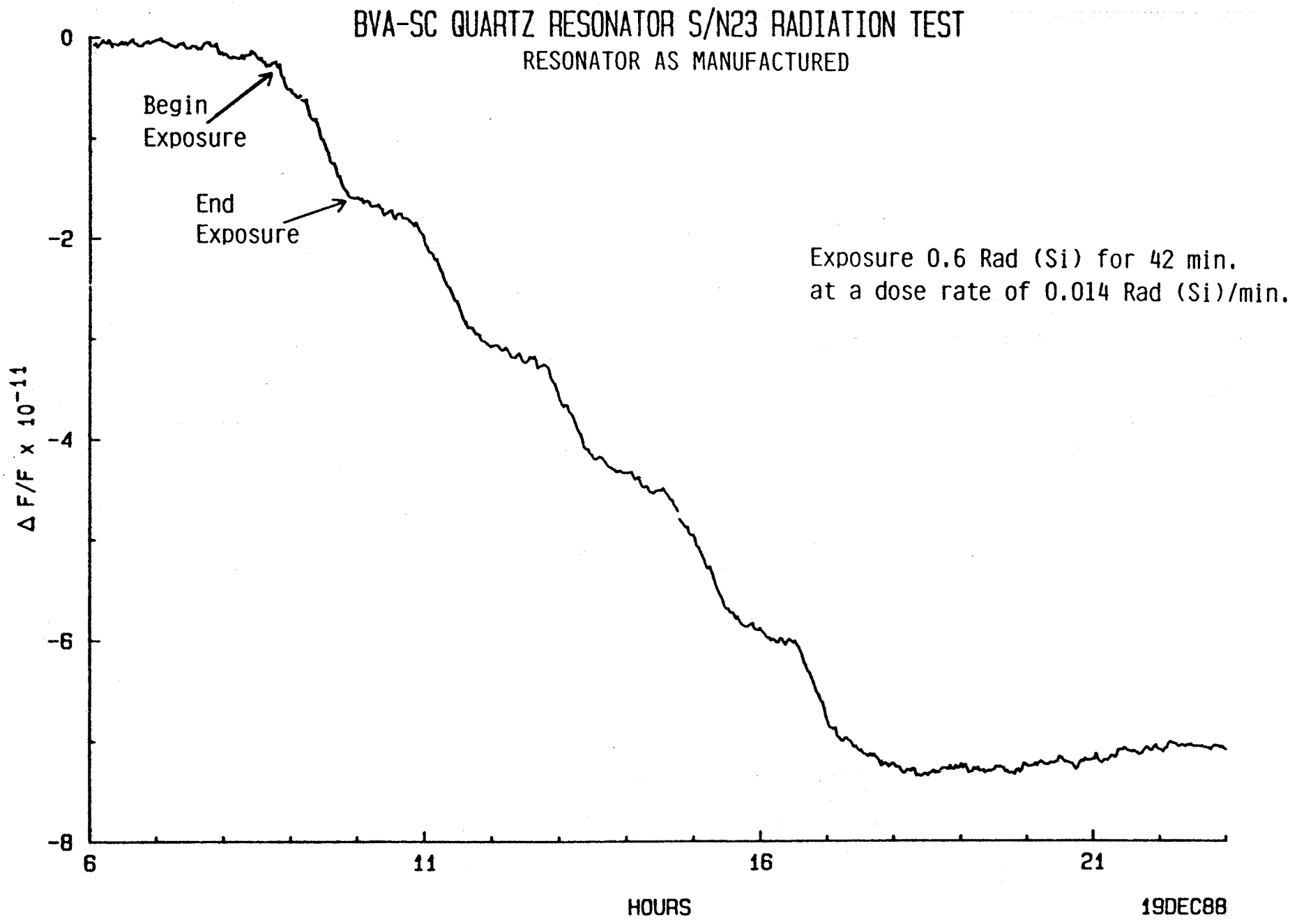

BVA-SC QUARTZ RESONATOR S/N23 RADIATION TEST
RESONATOR AS MANUFACTURED
Begin Exposure
End Exposure
Exposure 0.6 Rad (Si) for 42 min.
at a dose rate of 0.014 Rad (Si)/min.
0
-2
-4
-6
-8
Δ F/F x 10⁻¹¹
6
11
16
21
HOURS
19DEC88

Figure 6.

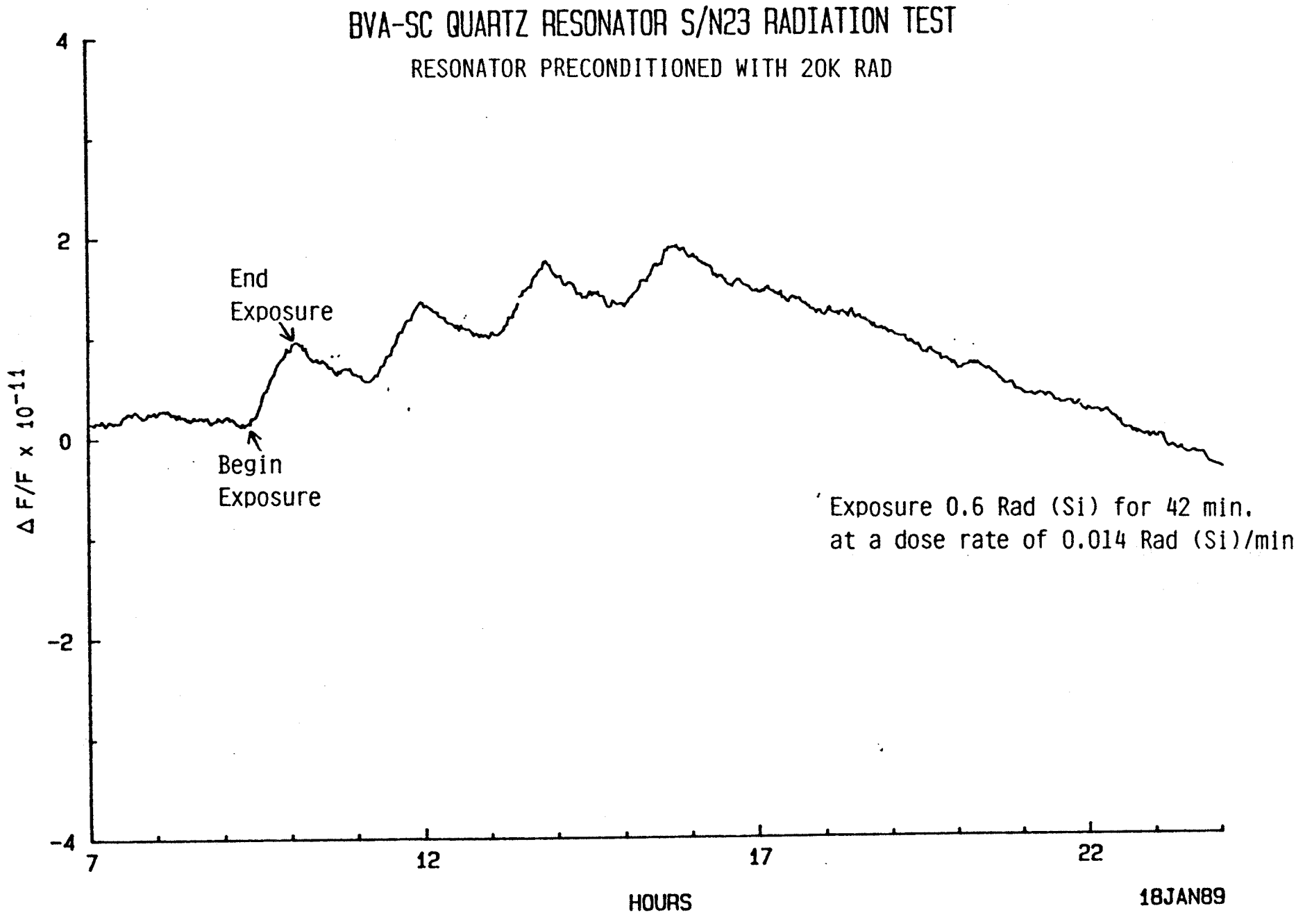

BVA-SC QUARTZ RESONATOR S/N23 RADIATION TEST
RESONATOR PRECONDITIONED WITH 20K RAD
End
Exposure
Begin
Exposure
Exposure 0.6 Rad (Si) for 42 min.
at a dose rate of 0.014 Rad (Si)/min.
ΔF/F x 10⁻¹¹
4
2
0
-2
-4
7
12
17
22
HOURS
18JAN89

Table I.

SC BVA RESONATOR ELECTRICAL CHARACTERISTICS*

Resonator Serial Number	Q	C_o [Pf]	f_o [Hz]	R [ohm]	Radiation Sensitivity per Rad+
1	2604448	4.12	4999982.8	47.6	0 (no response)
10	2582817	4.06	4999993.9	48.6	-3.67×10^{-11}
23	2469491	3.93	4999993.0	52.6	-1.6×10^{-11} $+1.3 \times 10^{-11}$ **
24	2615656	3.99	4999983.2	47.7	-1.6×10^{-11}

* Data supplied by Oscilloquartz, S.A.
(Resonators manufactured of Premium Q, swept z-grown alpha quartz)

+ For a 0.6 Rad [Si] exposure for 42 min. at a dose rate of 0.0142 Rad [Si] per min.

** Preconditioned wtih 20K Rad.

Plotting of the Vibrational Distribution of a TS Quartz Resonator using a Pulsed Laser

R.J.Williamson

STC Technology Ltd, Harlow, Essex. England

Summary

A method of plotting the vibrational distribution of a quartz resonator using a light beam as a probe to locally perturb the standing wave of the resonating area has been developed. The perturbation manifests itself as a transitory change in the frequency of the resonance, which is proportional to the magnitude of the standing wave. Thus by rastering the pulsed light beam across the area of the resonator and monitoring the change in its resonant frequency, the vibrational distribution can be mapped.

In practice, the crystal resonance is maintained by a phase-lock loop with a long time constant, and the change monitored is the phase across the crystal. This change is small and is typically buried in the 1/f noise of the oscillator. By using a lock-in amplifier at the repetition rate of the laser, this small change can be measured. The measurement uses commonly available equipment and can plot the vibrational distribution in less than a minute.

1. Introduction

A number of methods of plotting the vibrational distribution of quartz resonators have been developed, the most well known being Lang topography using X-rays. Others using interferometry are only really applicable to modes with a predominantly normal component, such as flexural modes and at low frequencies. The present method, however, is applicable to thickness shear modes at high frequencies. For such devices the only other method was to use a toothpick which is gently tracked across the surface of the resonator whilst the electrical response of the crystal is monitored. A mixture of dampening and loading by the toothpick alters the crystal in both the frequency and amplitude domain. Great skill is needed for this method since the likelihood of breaking the crystal was high. In the present method the laser beam replaces the toothpick and the effect is just to modulate the crystal frequency. It is proposed that the mechanism is by the force frequency effect whereby the light induced heat pulse creates a localised compression in the quartz due to thermal expansion.

2. Equipment

The laser used for the current work was a Questek 2040 excimer laser. In operation, excited dimers are formed between a noble gas and a halogen such as argon fluoride which emits at 193 nm. The output energy of the laser can be varied between 20 and 200 mJ per 15 ns pulse with a repetition rate up to 100 Hz. The output beam is about 10 * 20 mm² so that the fluence is between 10 and 100 mJ/cm² The large beam size enables a very simple means of beam placement as shown in figure 1. A computer-controlled X-Y table moves a brass foil containing a number of apertures thus enabling different areas to be irradiated but with the same fluence. In front of the foil is a beam expander which has a twofold purpose. It reduces the fluence available

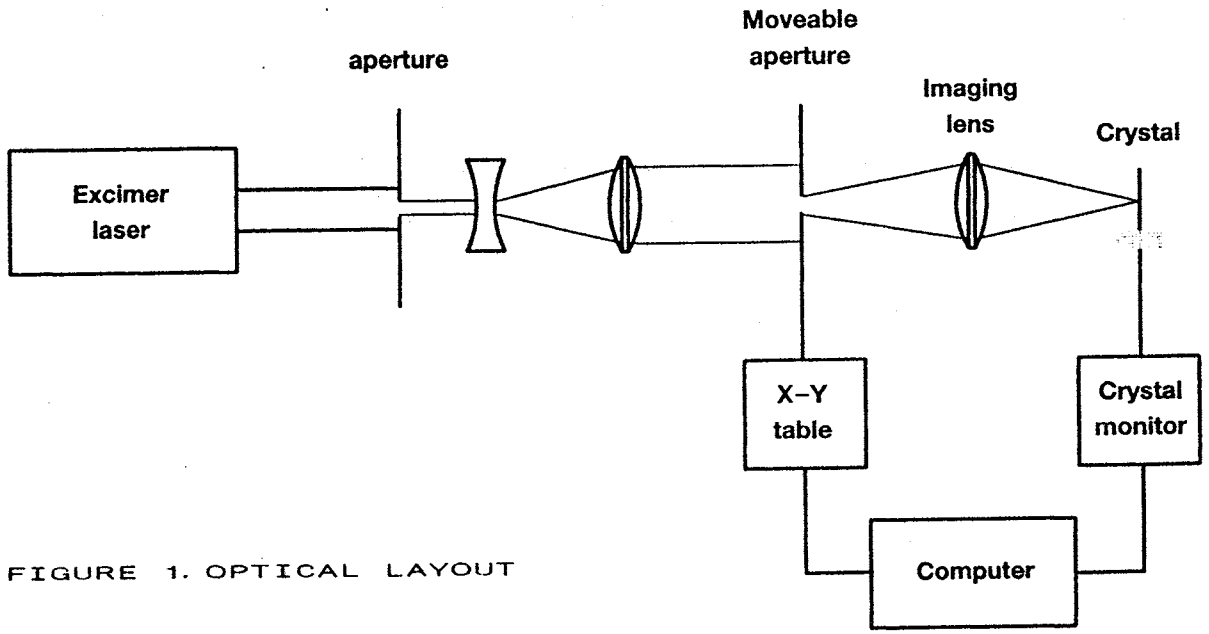


FIGURE 1. OPTICAL LAYOUT

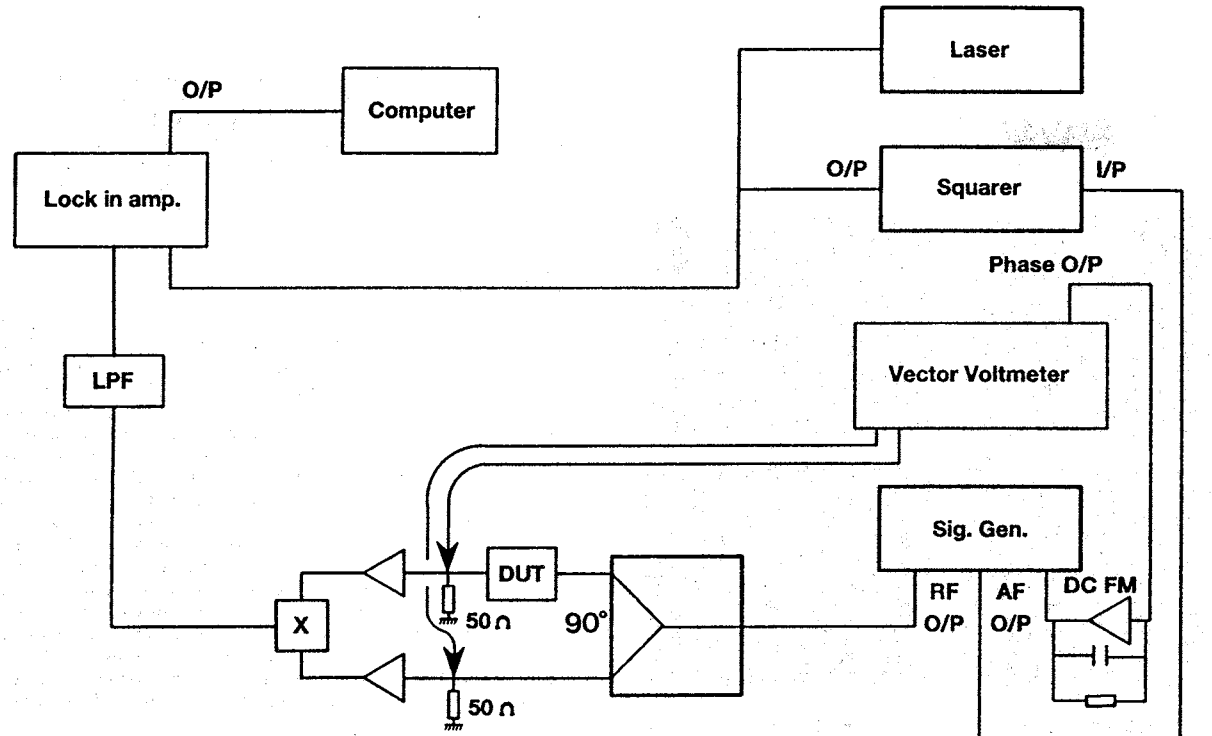


FIGURE 2. MEASUREMENT LAYOUT

from the laser and evens out the intensity distribution across the beam. After the foil a simple lens system is positioned at twice the focal length from both the foil and the crystal. This forms an inverted image with magnitude -1 of the foil and reduces the diffraction effects from the apertures.

The crystal measurement equipment is shown in figure 2. The signal generator used was a Rohde and Schwarz SMG since it has extensive modulation capabilities. The vector voltmeter was the ubiquitous HP 8405A and was used to find the resonance. Unique to what is in essence a modified IEC 444 crystal measurement circuit is the quadrature hybrid from Anzac (JH 131). This power splits with a ninety degree phase shift within a few degrees over the range 20 MHz to 200 MHz. This considerably simplifies the circuit since line stretchers are not required when the frequency is changed. The crystal is measured in a 100 ohm circuit made from SMA coaxial attenuators. This is not ideal from the point of view of sensitivity since it substantially increases the loaded Q of typical resonators, however it made for a much simpler jig. Two identical amplifiers from Minicircuits (ZFL500LN) were used to bring the signal levels up to about 7 dBm to maximise the sensitivity of the HP 10514 double balanced mixer used for phase detection. The output of the mixer is low pass filtered to remove the sum component and then measured using a Stanford Research SR510 lock-in amplifier. This is also fed from the same signal source that is used to trigger the laser (the AF output of the SMG). The lock-in amplifier is in essence a narrow band detector so that the wanted signal can be measured buried in noise.
The phase output from the vector voltmeter can be used instead of the mixer, in which case both the quadrature hybrid and the amplifiers are unnecessary. However, the rise time of the phase modulation will be slugged. Furthermore, that output is noisier with supply interference as well as the 20 kHz IF remnants. Nonetheless, that output has been successfully used for plotting.

3.Method

For plotting the vibrational distribution of a quartz resonator, the laser beam is rastered across the electrode of the crystal whilst it is sitting on resonance. The laser spot sizes used were 0.15 mm and 0.3 mm. depending on the size of the resonator electrode and the degree of detail required. The laser fluence used was about 3 mJ/cm^2 since this gave adequate response without damage to the electrode. Repetition rates used were typically 40 Hz since this was a compromise between avoiding both 1/f noise and supply interference whilst maintaining reasonable gas lifetimes for the laser.

To avoid any drift in the crystal frequency whilst plotting, the output from the mixer can be fed via a very low pass filter into the DC FM input of the signal generator. However, since a synthesised source was used and the ambient was well controlled, this proved unnecessary. The rastering and voltage measurement is under the control of a computer using the IEEE interface. It proved necessary to take an average of several measurements at each point since the pulse to pulse variation of the laser energy was at least 20%. The data is then used to generate contours which are drawn on the screen of the computer.

4 Results

This method has been used to plot the vibrational distribution of both overtones and inharmonics. Figure 3 gives a sequence of plots of a 50 MHz fundamental and its overtones up to the 11th at 550 MHz, illustrating that the method is very sensitive. The circle in the background corresponds to the position of the 1mm electrodes and the dots are the measurement positions. In this case a 0.3mm laser spot was used with no overlap of the spots.

Despite the large spot used, the main characteristics of this resonator can be seen. There is a wedge in the quartz toward the top left hand corner, the effect of which gets more pronounced with higher overtones.

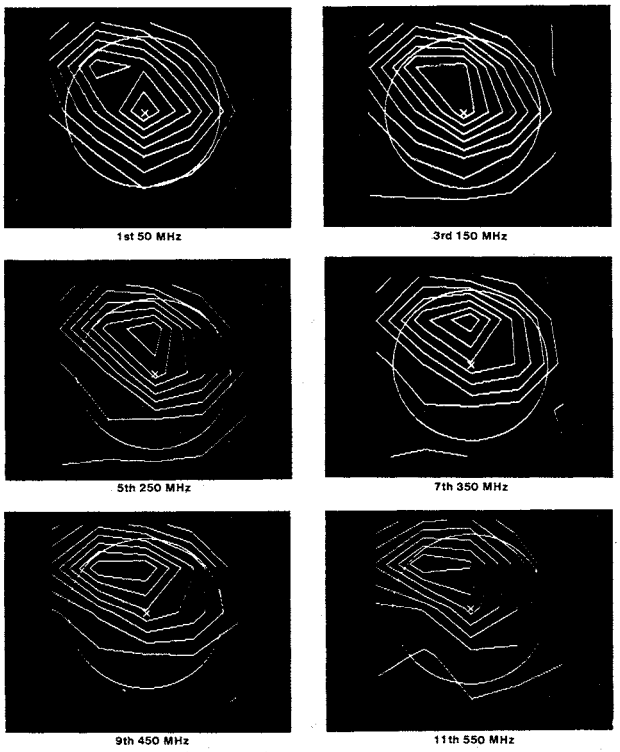


FIGURE 3. 50 MHZ FUNDAMENTAL AND OVERTONES

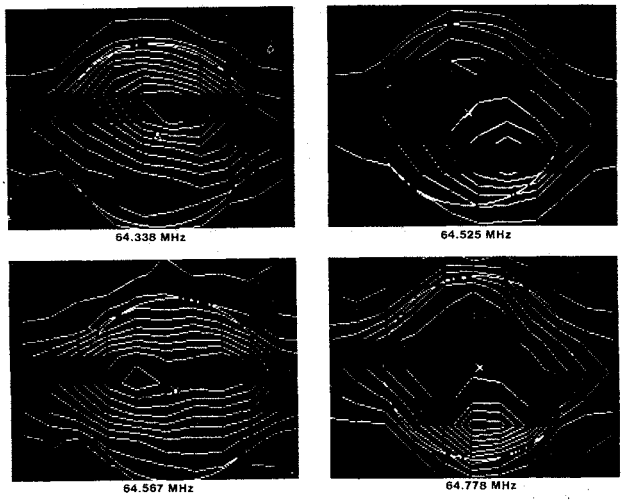


FIGURE 4. 64 MHZ FUNDAMENTAL AND INHARMONICS

Figure 4 gives a sequence of a 64.3 MHz fundamental with some of its inharmonics. In this case the laser spot diameter was 0.15mm so that more points could be used and greater detail is revealed. There is a limit on the effective spot size and hence on the resolution achievable because of lateral thermal diffusion which will smooth out detail. This will affect the plots of the inharmonics most where the nodal lines are suppressed. However, the general shape is still evident. The crystal used was of the standard key-hole design with the tabs positioned on both sides of the electrode (the circle). It can be seen that the vibrational distribution extends down the tabs for the higher inharmonics.

5 Mechanism

In order to elucidate the mechanism responsible for the observed phase shifts used for the plotting process, the time constants and the magnitudes of the phase shifts have been investigated. It is assumed that a fraction of the laser energy from each 15 ns pulse is absorbed by the electrode and converted into heat which then diffuses into the quartz. It is also assumed that for beam sizes large compared with the thickness of the quartz, lateral diffusion may be ignored initially. After a characteristic time, the absorbed heat will have raised the local temperature of the cylindrical section under the laser spot, thus building up a stress field which causes the observed phase shift. The details of this model have been tested experimentally as described in the following

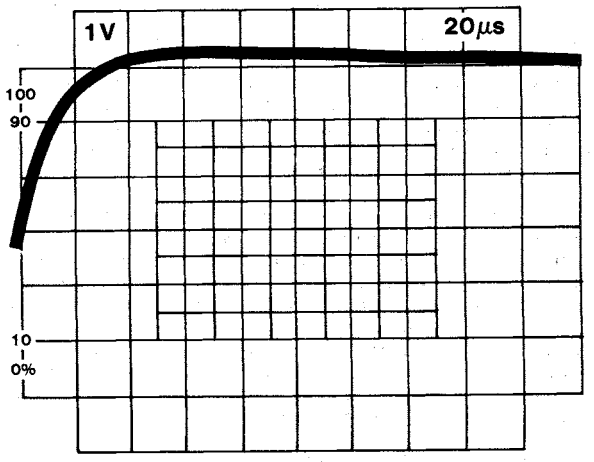


FIGURE 5. RISING EDGE OF PULSE

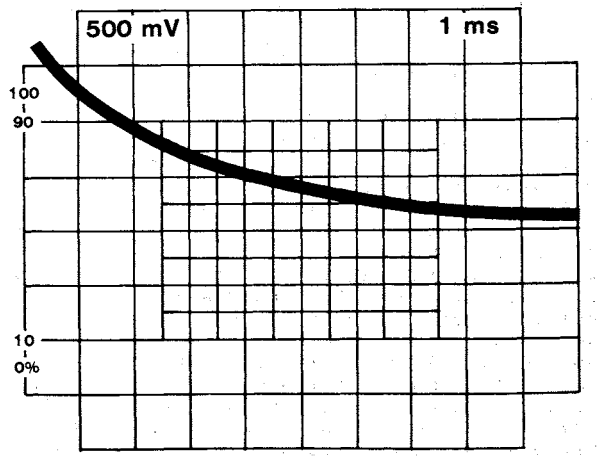


FIGURE 6. FALLING EDGE OF PULSE

5.1 Time Domain

The output of the mixer from a 50 MHz fundamental crystal irradiated by a single laser pulse is given in figures 5 and 6. This shows that the rise time of the phase shift is exponential and is about three orders of magnitude faster than the fall time. Figure 7 plots this rise time (measured assuming an exponential form) for a number of crystals between 50 MHz and 114 MHz and shows that it is proportional to $1/f^2$. This relationship confirms that the phase pulse may be the result of a laser induced heat pulse diffusing through the thickness of the quartz.

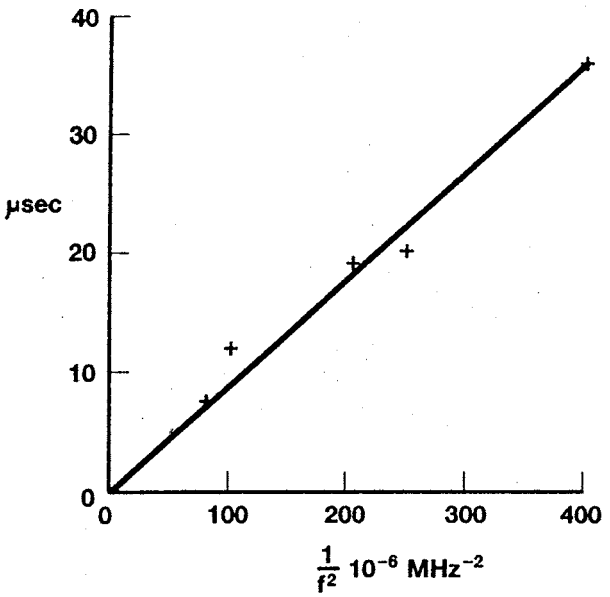


FIGURE 7. PLOT OF RISE TIME WITH $1/f^2$

A mathematical model of heat absorption on one side of a thin substrate is given by Danielson and Sidles [1]. It includes the optical absorption depth of the substrate. In the present case, the absorption in the quartz surrounding the electrode is small. For the metallic electrode, however, the absorption depth is typically less than 100Å. This is much less than the thickness of electroding of a crystal so that the heat from the light pulse will be generated in the electrode and no light will penetrate into the quartz. This simplifies the heat diffusion equation for a short laser pulse on one side (x=0) of a substrate of thickness l to

$$T(x,t)/T_\infty = 1 + 2\sum_{n=1}^{\infty}\cos(n\pi x/l)\cdot\exp(-n^2\tau) \qquad (1)$$

where $T_\infty = Q/\rho c l$ is the uniform temperature after a long time, Q is the absorbed fluence, ρ is the density of the substrate and c is its specific heat. $\tau=\pi^2 Dt/l^2$ is a dimensionless time with D the thermal diffusivity. This is presented graphically in figure 8.

For a quartz crystal resonator vibrating in thickness shear, its frequency is given by

$$f = C/l \qquad (2)$$

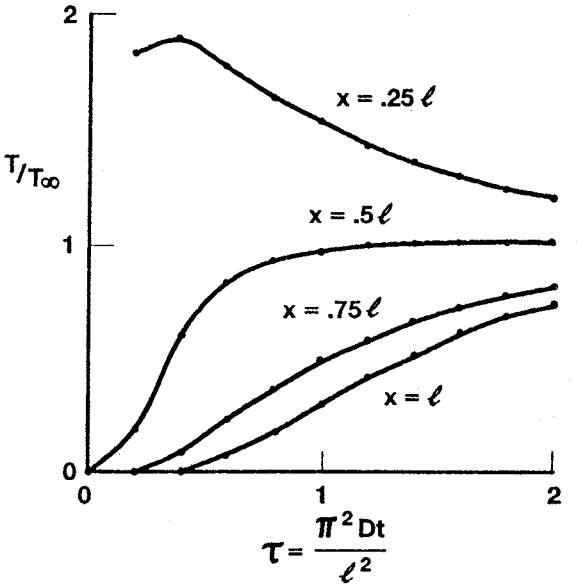


FIGURE 8. PLOT OF HEAT TRANSFER INTO SUBSTRATE

where C is the appropriate frequency constant for the cut of quartz. Thus the dimensionless time becomes

$$\tau = \pi^2 Dtf^2/C^2 \qquad (3)$$

or

$$t = (\tau C^2/\pi^2 D)\cdot 1/f^2 \qquad (4)$$

Thus the actual time is proportional to this dimensionless time and $1/f^2$. For AT quartz D=4.3 mm²/sec in the thickness direction and C=1660 kHz.mm. Taking t to be the rise time of the phase pulse for a crystal at frequency f we get τ=1.3. From figure 6 this corresponds to the time for the back face of the crystal to have reached half its final temperature thus confirming that transport of heat is the mechanism responsible for the observed time constant of the phase shift.

5.2 Amplitude Domain

It is concluded that the time response of the phase pulse is governed by the thermal diffusion time of the laser induced heat pulse. It is now proposed that the mechanism responsible for the frequency modulation is the force frequency effect, with the force supplied by the thermal expansion of the quartz. This is given by [2,3]

$$df/f = Kt_q \qquad (5)$$

where K is a constant for a given crystallographic orientation and t_q is the lateral stress across the whole resonator. For AT and BT-cuts, K is $-2.75*10^{-11}$ and $2.65*10^{-11}$ m²/N, respectively. Thus these two cuts are expected to show opposite phase shifts. This change of sign has been demonstrated.

We can test this hypothesis further by comparing the measured frequency shift for a crystal with theory. For these tests the laser spot was made large enough to cover the whole device area. The maximum frequency deviation occurs once the heat pulse has equilibrated through the thickness of the crystal but before lateral diffusion becomes

significant. At this point the temperature increase of the quartz is T_∞. The resulting thermally induced lateral stress will be

$$t_q = \alpha C T_\infty \qquad (6)$$

where α is the average lateral expansion coefficient and C is the average lateral elastic constant. This gives

$$df/f = K\alpha CQ/\rho c l \qquad (7)$$

For a 6MHz crystal with aluminium electrodes irradiated by 3.4mJ and assuming the absorption at 193 nm to be about 0.3, this should give a frequency shift of 6.6 ppm. To measure this, it is necessary to calculate the mixer conversion factor and the Q of the crystal. This can conveniently be achieved by applying a known frequency modulation and measuring the output from the mixer. With this calibration, the measured frequency shift was 10.4 ppm which is remarkably close to the theoretical in view of the assumptions made.

5.3 Vibrational distribution

It is important to remember that equation 7 holds only for a uniform stress, that is where the laser beam covers the whole resonator area. To use this mechanism to plot the vibrational distribution a small spot is used and it is necessary to show that the phase and hence frequency deviation resulting from a laser pulse at a particular point is proportional to the thickness shear amplitude at that point. (A similar problem is that of proving that the mass sensitivity of such a resonator is proportional to the amplitude of vibration. This is a proof that is often evaded in books on microbalances [4].) A solution to both problems lies in the variational method used in quantum mechanics [5] but first applied by Lord Rayleigh in 1873 [6].

As described above, the effect of the laser beam is to alter the elastic constant of the quartz within the area of the beam. For a thickness shear resonator, the potential energy within that beam is proportional to this elastic constant times the square of the vibrational amplitude. Integrating this over the area of the device, gives the total potential energy. Now, the natural resonant frequency is the frequency at which the time-averaged kinetic energy of the system, which is proportional to the square of the frequency, equals the time-averaged potential energy. Thus any change in the potential energy of the system produces a corresponding change in the resonant frequency. It follows that for a given local change in the elastic constant, where the amplitude is maximum the change in potential energy and hence frequency is greatest. The converse is also true so that the frequency shift must follow the amplitude.

A similar line of reasoning, but invoking kinetic energy, can be used to explain the mass sensitivity of a thickness shear resonator.

5.4 Effect of anisotropy

The force frequency effect for quartz is caused by elastic nonlinearities (third order elastic constants). The direction of the force is important because of the anisotropy of quartz. The effect described above is the integral in all directions. For the BT cut the coefficient varies slightly with angle,

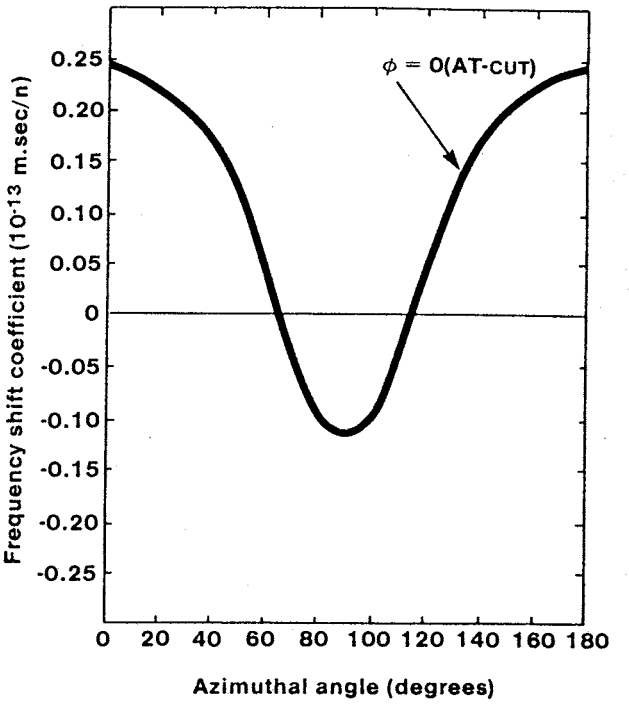


FIGURE 9. FORCE FREQUENCY EFFECT FOR AT QUARTZ

but for many of the other cuts [7] including the AT cut (figure 9) the coefficient goes through zero and hence changes sign in some directions.

This anisotropy revealed itself when plotting AT cut resonators. It was observed that in some areas of the resonator, always near the edge and often in two places nearly opposite one another, the phase pulse changed sign. This happened when the maximum of the slope of the amplitude was about ±90° to the X axis. Since the laser spot in the present investigation had a size only about an eighth the resonator diameter, there could be an appreciable difference in amplitude across the spot on the slope of the amplitude distribution. The effect of the spot is proportional to the integral of the amplitude times the angular value of the effect, so that the resultant can be dominated by one direction and hence change sign.

$$df/f = \int A^2(r,\theta)\cdot K(\theta)d\theta \qquad (8)$$

This anisotropy can be diminished but not eliminated by reducing the beam diameter. Apart from signal to noise considerations, the lateral diffusion of heat along the electrode will extend the area of quartz perturbed. The thermal diffusivity of aluminium is 81 mm²/sec and the ratio of the thermal diffusion lengths in the electrode and the quartz is given by

$$l_{Al}/l_Q = \sqrt{D_{Al}/D_Q} = 4.35 \qquad (9)$$

Therefore if the laser spot diameter is similar to the quartz thickness the lateral diffusion will dominate.

The result of the anisotropy in the force frequency effects for some cuts is a differing sensitivity along the two axes. For instance, a circular vibrational distribution will appear elliptical. In principle this can be taken into account when forming the contours if the orientation is known. Alternatively this property can be used to measure the orientation of the device.

6 Conclusions

The method described here is capable of plotting the vibrational distribution of fundamental thickness shear resonators, their overtones and their inharmonics. It uses a pulsed mode laser to probe the surface of the resonator in a raster pattern thus building up a contour map of the distribution. The parameter measured is the magnitude of the induced phase modulation as the crystal is momentarily knocked off frequency. It is proposed that this perturbation is caused by thermally induced compressive stress through the thickness of the quartz under the laser beam. Because of the elastic nonlinearities in quartz this gives rise to a frequency shift.

The basic principle of the method is applicable to any device whose operating parameter has a temperature coefficient. For instance, the distribution of current density in a circuit or IC could be plotted by monitoring the modulation in the current through it at the modulation frequency. The sensitivity of the method is obtained by the use of a lock-in amplifier or wave analyser to measure this modulation in noise.

7 Acknowledgement

The author wishes to thank Dr A.F.B.Wood of STC Quartz Crystal Division and R.A.Heinecke of STC Technology Ltd for valuable discussions. This work has been carried out with the support of the Procurement Executive, Ministry of Defence and also by STC Quartz Crystal Division.

8 References

[1] G.C.Danielson and P.H.Sidles, "Thermal diffusivity and other non-steady state methods", Thermal Conductivity, Vol.2, Ed. R.P.Tye Acedemic Press Inc. (London) Ltd

[2] E.P.EerNisse, "Quartz resonator frequency shifts arising from electrode stress", Proc. 29th AFCS, pp 1-4

[3] E.P.EerNisse, "Stress effects in quartz crystal microbalances", Methods and Phenomena, Vol.7, Ed. C.Lu and A.W.Czanderna, Elsvier Scientific Publishing

[4] H.K.Pulker and J.P.Decosterd, "Applications of quartz crystal microbalances for thin film deposition process control",Methods and Phenomena, Vol.7, Ed. C.Lu and A.W.Czanderna, Elsvier Scientific Publishing

[5] L.I.Schiff, "Quantum Mechanics", p171, McGraw-Hill Book Company

[6] Lord Rayleigh, "Theory of Sound", 2nd rev. ed., vol.1, Sec.88, reprinted by Dover, New York

[7] A.Ballato, E.P.EerNisse and T.Lukaszek, "The force frequency effect in doubly rotated quartz resonators", Proc 31st AFCS, pp 8-16

©1989 STC PLC
STC TECHNOLOGY LTD
London Road, Harlow, Essex, CM17 9NA England

DYNAMIC HIGH FIELD RELAXATION OF IMPURITIES IN QUARTZ CRYSTAL RESONATORS

R. Brendel

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

It is shown in the present paper that the amplitude and phase of the resonant frequency of a crystal containing ionic impurities, submitted to a slowly-varying alternating electric field, are related to the concentration and the diffusion constant of the defects. An experimental setup for recording dynamic frequency relaxation is presented, as well as experimental results. By identifying the experimental data obtained at different frequencies with theoretical curves, it is possible to obtain both the impurity concentration and the diffusion constant. Moreover, experiments carried out at various temperatures permit extraction of the migration energy of the impurity.

Introduction

High precision in positioning devices requires high stability time bases. These devices are mainly quartz crystal oscillators whose frequency may be affected by various perturbations, such as thermal or mechanical disturbances, ionizing radiations, etc.

Some of these perturbations alter frequency stability through the mechanism of impurities contained in the crystal. High field frequency relaxation is a method used to understand how the acoustical properties of the material can be influenced by impurities. The method requires observing the frequency variation of a resonator submitted to a high electric field, which can be either continuous or alternating, and then deducing the impurity features from the records.

Electroelastic effect and the static relaxation method

It has been known for a long time that the frequency of some quartz crystal resonators can be changed by applying a strong quasi static electric field to their electrodes. This effect was first observed and studied by C.K. Hruska [1,2] who showed that it is due to a nonlinear piezoelectric phenomenon called the polarizing or electroelastic effect.

Ever since, much effort has been spent in attempting to determine the nonlinear piezoelectric coefficients responsible for the electroelastic effect, either by using the resonator method (frequency shift measurement) [3-7] or the transit time method (pulse-echo technique) [8,9]. Recent works show that the initial discrepancies between the results obtained by both methods are being progressively reduced [10].

When performing electroelastic effect measurements by the resonator method, the initial frequency shift is often followed by a slow relaxating behaviour (see Fig. 1), first explained by J. Kusters [6] as being due to a screening effect produced by ionic impurities moving through the crystal under the applied quasi static electric field. Other experiments have since been reported [11,12] together with a theoretical interpretation of the relaxation phenomenon. It was shown that the relative frequency shift of an infinite plate with thickness 2h is related to the internal electric field E by the relation (1)

$$\frac{\Delta f}{f} = \frac{K}{2h\,\rho_o\,V^2} \int_{-h}^{+h} E \cos^2\left(\frac{m\,n\,x}{2h}\right) dx \qquad (1)$$

where ρ_o is the density of the material, V the velocity of the acoustical vibration for a given mode, m the overtone, x the abscissa along the normal to the plate and K a parameter whose value depends upon the crystallographic orientation of the plate [13]. For a perfect crystal, the internal electric field is a constant, so Eq. (1) is reduced to the homogeneous electroelastic effect

$$\frac{\Delta f}{f} = \frac{K\,E}{2\rho_o\,V^2} \qquad (2)$$

The static relaxation method consists of measuring the ratio between the steady state frequency shift $\Delta\omega_D$ and the initial homogeneous frequency shift $\Delta\omega_H$ (see Fig. 1). It can be expressed as a function of the applied voltage and the impurity concentration. Nevertheless the kinetic properties of the charge motion, such as the diffusion constant or the mobility can be hardly obtained by this method.

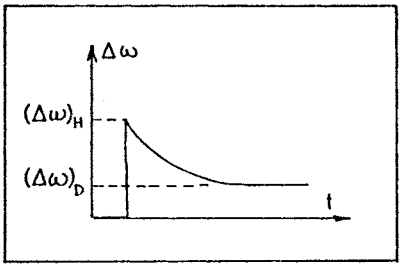


Fig. 1 : Under the influence of an applied DC voltage, the resonator frequency experiences an initial frequency shift $(\Delta\omega)_H$ due to the electroelastic effect. It is followed by a relaxation phase created by the impurity diffusion screening effect, leading to a steady state frequency shift $\Delta\omega)_D$.

The dynamic high field relaxation method

In order to obtain the kinetic parameters of the ionic diffusion, the experimental setup previously used has been modified. The static DC field is replaced by a low frequency sine generator driving a high voltage amplifier (see Fig. 2), the synthesizer is locked onto the oscillator and the error voltage proportional to the frequency variation of the quartz crystal resonator in recorded, together with the applied voltage. The high voltage frequency range lies between 10^{-4} Hz and a few Hz.

The records obtained at various frequencies of the high voltage show (see Fig. 3), as could be expected, that the amplitude of the frequency variation decreases when the high voltage frequency decreases. This can be explained by noting that at high frequencies, the impurities have hardly had time to move before the field is inverted. Consequently, their small motion cannot substantially modify the internal electric field and the frequency shift follows the relation (2), where E is a sine function of time as in a pure crystal. On the other hand, at low frequencies, the impurities move further under the influence of the electric field and give rise to a screening of the field, leading to a decrease in amplitude of the frequency variation.

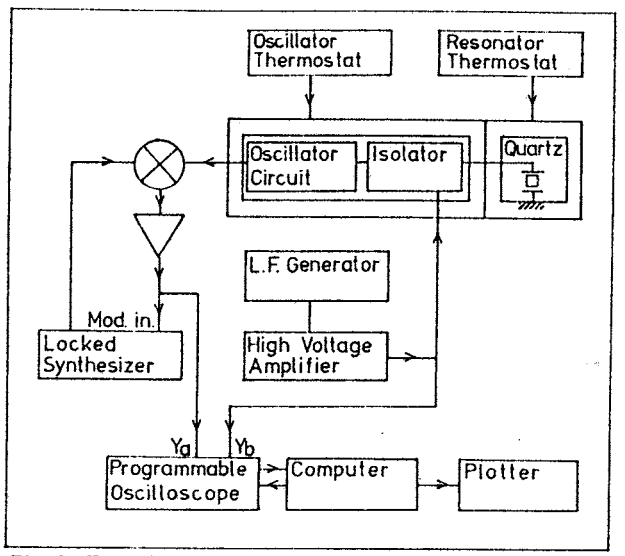
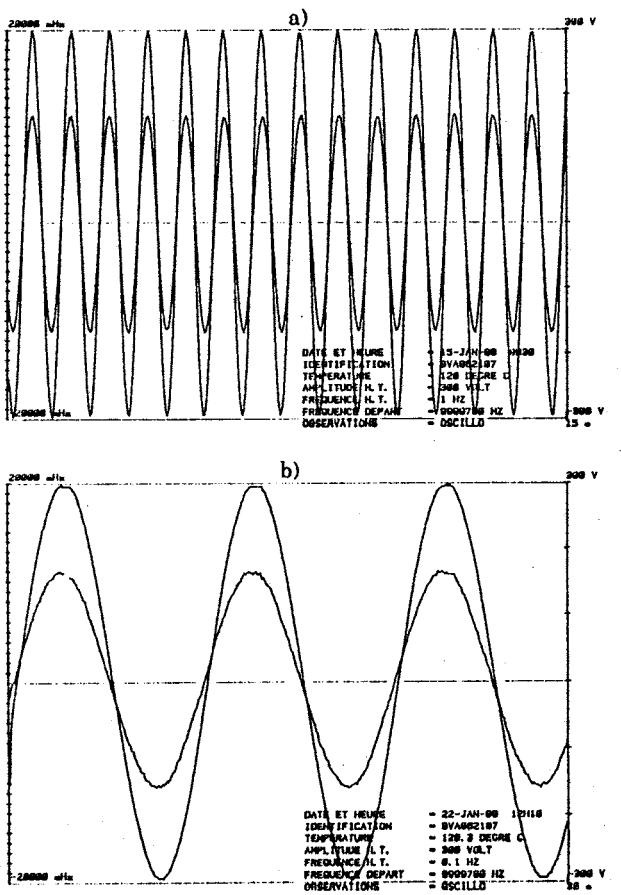


Fig. 2 : Experimental setup block diagram for the dynamic high field relaxation measurements



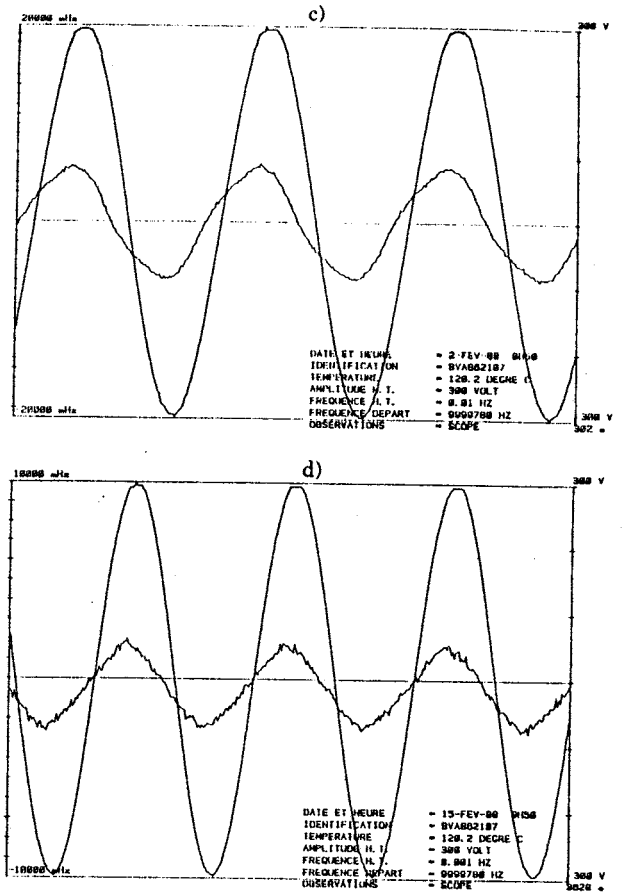


Fig. 3 : Examples of frequency shift records obtained with high voltages at various frequencies F_{HV}. The larger trace is the high voltage amplitude (right scale), the smaller trace is the induced frequency shift (left scale). As the high voltage frequency F_{HV} decreases, the frequency shift amplitude decreases, but remains approximately sinusoïdal
a) $F_{HV} = 1$ Hz b) $F_{HV} = 0.1$ Hz
c) $F_{HV} = 0.01$ Hz d) $F_{HV} = 0.001$ Hz
The crystal under test here is a 10 MHz 3rd overtone electrodeless resonator

Theoretical interpretation of the dynamic relaxation

The most common interpretation of the ionic conductivity in quartz assumes that during the growth of the crystal some Si^{4+} atoms are replaced by Al^{3+} associated with a monovalent cation such as Na^+, Li^+, ..., achieving charge neutrality in the crystal [14]. In addition, it is assumed that cations can neither leave nor enter the crystal from the outside (blocking electrodes). When submitted to an external electric field the cations move towards the negatively charged electrode leaving behind them uniformly distributed negative charges. Only the diffusion of one kind of impurity in a one-dimensional model is considered here.

In this case, the cationic concentration p(x,t) obeys a set of nonlinear parabolic partial differential equations :

charge conservation :

$$\frac{\partial p}{\partial t} = D \frac{\partial^2 p}{\partial x^2} - \mu \frac{\partial (pE)}{\partial x} \qquad (3)$$

Gauss' theorem :

$$\frac{\partial E}{\partial x} = \frac{p - n_o}{e} \varepsilon \tag{4}$$

$$E = -\frac{\partial \Phi}{\partial x} \tag{5}$$

where $E(x,t)$ is the internal electric field, $\Phi(x,t)$ the potential, e the electronic charge, ε the permittivity of the crystal along the direction of diffusion and n_o the initial impurity concentration. In the absence of external electric field the electrical neutrality of the crystal requires $p = n_o$ every where.

In the above equations, D and μ are respectively the diffusion constant and mobility of the moving specie related by Einstein's relation

$$D = \mu . U_T \quad , \quad U_T = \frac{k_B T}{e} \tag{6}$$

where k_B is Bolzmann's constant and T the absolute temperature.

Because of the blocking electrodes, we have the additional relation

$$\int_{-h}^{+h} n_o dx = \int_{-h}^{+h} p\, dx = 2 n_o h \qquad \text{at all times} \tag{7}$$

Steady state solution of the dynamic relaxation equations

When a low frequency high voltage sinusoidal potential of amplitude Φ_o is applied to the electrodes of the vibrating plate, the boundary conditions can be written

$$\begin{aligned} \Phi(-h,t) &= 0 \\ \Phi(+h,t) &= \Phi_o e^{j\Omega t} \end{aligned} \tag{8}$$

Although the diffusion equations are nonlinear, experiments show that if the high voltage period is short enough, the impurities cannot move far from their equilibrium position, so that the internal electric field is not too much distorded. In such a case the resonant frequency variation of the plate is roughly sinusoidal (see Fig. 3). That observation leads one to look for a harmonic potential solution which can be expressed to first approximation in the form

$$\Phi(x,t) = \phi(x)\, e^{j\Omega t} \tag{9}$$

where $\phi(x)$ is a time independent complex function whose boundary values (8) become

$$\begin{aligned} \phi(-h) &= 0 \\ \phi(+h) &= \Phi_o \end{aligned} \tag{10}$$

By taking the derivatives of Eq. (9) with respect to x and t as many times as required and putting Eqs. (4) and (5) in Eq. (3) one obtains the nonlinear equation

$$j \frac{\varepsilon \Omega}{e} \phi'' e^{j\Omega t} = D \left| \frac{\varepsilon \phi''''}{e} - \frac{\phi''}{U_T} \left(n_o - \frac{\varepsilon \phi''}{e} e^{j\Omega t} \right) + \frac{\varepsilon \phi' \phi'''}{e U_T} \right| e^{j\Omega t} \tag{11}$$

where ϕ', ϕ'', ϕ''' and ϕ'''' are successive derivatives of ϕ with respect to the space variable x.

If the higher harmonic terms such as $e^{2j\Omega t}$ are neglected in Eq. (11) this equation is reduced to a simple linear differential equation

$$\phi'''' - \alpha^2 \phi'' = 0 \tag{12}$$

where

$$\alpha^2 = \kappa + j\lambda \qquad \kappa = \frac{n_o e}{\varepsilon U_T} \qquad \lambda = \frac{\Omega}{D} \tag{13}$$

Furthermore, condition (7) is then equivalent to

$$\int_{-h}^{+h} \phi'' dx = 0$$

which gives

$$\phi'(-h) = \phi'(h) \tag{14}$$

This result shows that the instantaneous electric field has the same value at the two boundaries. By solving Eq. (12) with condition (14) it can be shown that the first harmonic potential solution has the following form

$$\Phi(x,t) = \frac{\Phi_o}{2} \left| \frac{\sinh \alpha x - (x/h) \sinh \alpha x}{\sinh \alpha h + j \frac{\lambda \alpha h}{\kappa} \cosh \alpha h} + \frac{x}{h} + 1 \right| e^{j\Omega t} \tag{15}$$

Equation (5) then gives the electric field

$$E(x,t) = -\frac{\Phi_o}{2h} \left| 1 + \frac{\alpha h \sinh \alpha x - \sinh \alpha h}{\sinh \alpha h + j \frac{\lambda \alpha h}{\kappa} \cosh \alpha h} \right| e^{j\Omega t} \tag{16}$$

and Eq. (4) expresses the local excess charge

$$p(x,t) - n_o = -\frac{n_o \Phi_o}{2 U_T} \frac{(1 + j\frac{\lambda}{\kappa}) \sinh \alpha x}{\sinh \alpha h + j \frac{\lambda \alpha h}{\kappa} \cosh \alpha h} e^{j\Omega t} \tag{17}$$

Steady state harmonic frequency shift

Equations (13) show that for a perfect crystal (i.e. a crystal without impurities) Eq. (16) reduces to an uniform electric field

$$E_h = -\frac{\Phi_o}{2h} e^{j\Omega t} \tag{18}$$

Substituting this value into Eq. (1) one finds the homogeneous dynamic frequency shift

$$f_h = \left(\frac{\Delta f}{f} \right)_h = \frac{K}{2 \rho_o V^2} \left(\frac{-\Phi_o}{2h} \right) e^{j\Omega t} \tag{19}$$

The diffusion dynamic frequency shift $f_d = (\Delta f/f)_d$ would be obtained in a similar way, but by substituting Eq. (16) into Eq. (1).

The relevant parameter of the dynamic high field relaxation is the ratio between f_d and f_h, that is

$$\frac{f_d}{f_h} = \frac{\frac{1}{1 + (\frac{\alpha h}{m\pi})^2} + j \frac{\lambda}{\kappa} \frac{\alpha h}{\tanh \alpha h}}{1 + j \frac{\lambda}{\kappa} \frac{\alpha h}{\tanh \alpha h}} \tag{20}$$

The homogeneous frequency shift f_h can be measured from an experiment at low temperature or with a large enough high field frequency to insure that the impurities have not time to move.

Relation (20) allows one to connect the amplitude and phase of the frequency shift of the plate to the kinetic diffusion parameters (diffusion constant and mobility).

Figures 4 and 5 show the amplitude and phase of the ratio f_d/f_h given by Eq. 20 as a function of the variable $\lambda = \Omega/D$ (log scale) for different values of the parameter $\kappa = (n_o e)/(\varepsilon U_T)$.

The curves obtained look different for κ larger or smaller than a limiting value between 10^{11} and 10^{12}, corresponding to an impurity concentration of about 10^{18} charges/m^3 at 120°C.

The calculation was done for a 0.56 mm thick quartz plate vibrating in the 10 MHz third overtone.

These curves show, as expected, that when the high voltage frequency increases the diffusion frequency tends towards the homogeneous frequency of a pure crystal.

This feature appears more clearly in a Nyquist plot (Fig. 6) where the imaginary part is plotted versus real part of the ratio f_d/f_h. The curves obtained then look like half circles which are described in a clockwise sense with increasing high voltage frequency.

All half circles pass through the point (1,0) and their radii increase as the impurity concentration increases, tending towards a limiting half circle of radius 1/2 going through the origin when κ tends toward infinity.

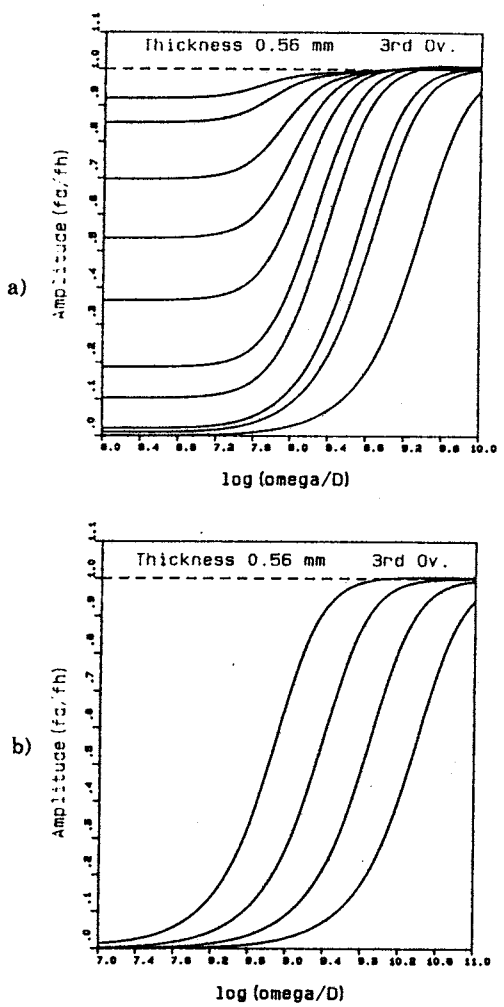


Fig. 4 : Theoretical amplitude of the complex ratio f_d/f_h (Eq. 20) as a function of the variable $\lambda = \Omega/D$ for different values of the parameter $\kappa = n_o e / (\varepsilon U_T)$ proportional to the impurity concentration n_o.
Here, the parameter κ has the values :
a) 1, 2, 5.10^8 ; 1, 2, 5.10^9 ; 1, 5.10^{10} ; 10^{11} ; 10^{12}
b) 10^{11} ; 10^{12} ; 10^{13} ; 10^{14}
The curves shift toward the right as κ increases

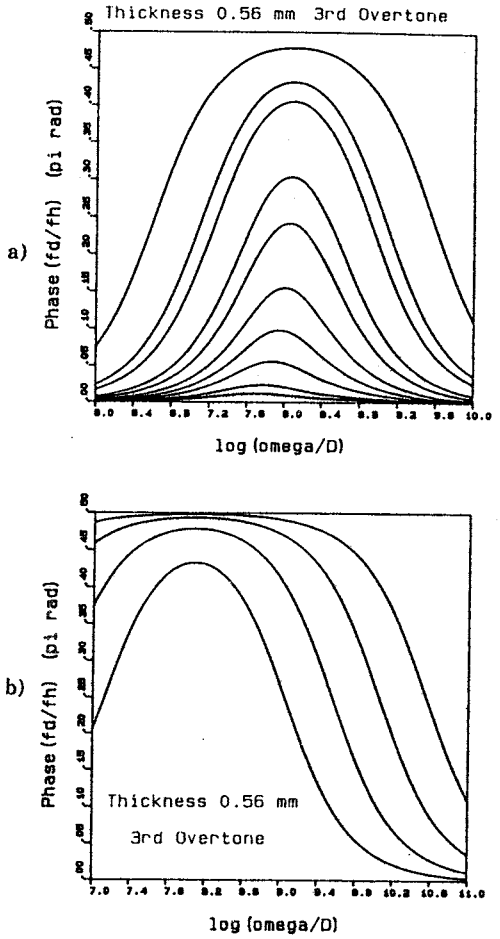


Fig. 5 : Theoretical phase of the complex ratio f_d/f_h, under the same conditions as on Fig. 4.
The curves shift toward the top as κ increases.

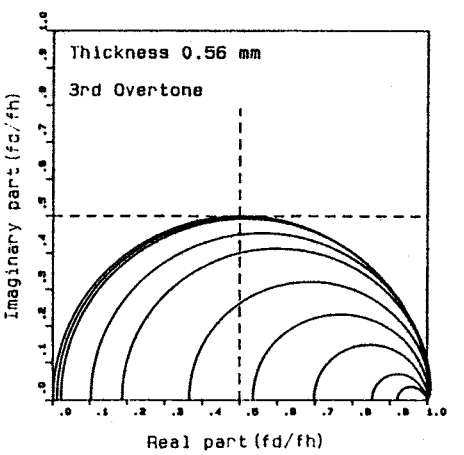


Fig. 6 : Theoretical curve of the complex ratio f_d/f_h (Eq. 20) in a Nyquist plot, for the same conditions as in Fig. 4a and 5a. As κ increases the half circle radii increases, tending toward a limiting curve of radius 1/2.

Experimental results

From records such as those presented Fig. 3 it is possible to plot the amplitude and phase of the ratio f_d/f_h as a function of the high voltage excitation frequency for different values of the temperature (100, 120 and 140°C on the example shown Figs. 7 and 8). The experiment was performed with a 10 MHz third overtone electrodeless SC cut synthetic quartz crystal resonator (BVA). The amplitude of the high voltage was 600 V peak-to-peak, but the results do not depend appreciably on the strength of the field. In fact they are not different from the records obtained with a high voltage amplitude of 400 V peak-to-peak (Figs. 10 and 11).

The amplitude curves obtained have the form of the theoretical plots. Although the phase measurement is somewhat inaccurate because of the appearance of higher harmonics (see Fig. 3) it is obvious the phase difference goes through a maximum, as predicted theoretically.

As the temperature increases, the curves shift towards the higher frequencies. This means that the ratio $\lambda = \Omega/D$ increases with temperature. Since the frequency range is the same in all cases this implies that the diffusion constant decreases with increasing temperature, as expected.

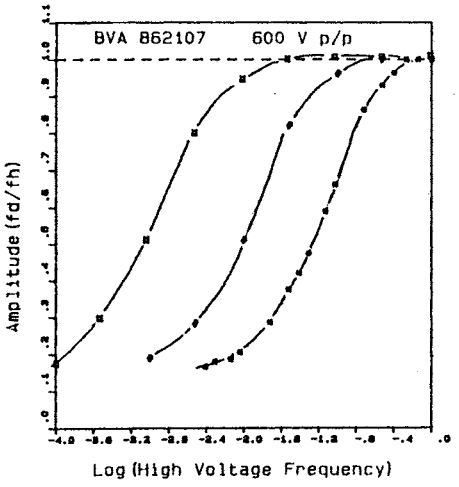


Fig. 7 : Experimental amplitude of the ratio f_d/f_h, obtained with a 10 MHz third overtone SC cut electrodeless quartz resonator at different temperatures. From left to right : 100°C, 120°C, 140°C. The high voltage amplitude here is 600 V peak-to-peak.

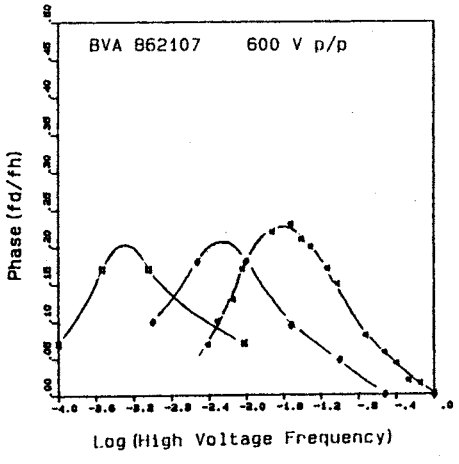


Fig. 8 : Experimental phase shift of the ratio f_d/f_h under the same conditions as in Fig. 7

Figure 9 shows the same experimental points as in Figs. 7 and 8 in a Nyquist plot. Whatever the temperature, the points are distributed on the same curve although the point on the curve corresponding to a given frequency changes with the temperature. This means that the coefficient κ, proportional to the impurity concentration, does not depend appreciably on the temperature in the range used.

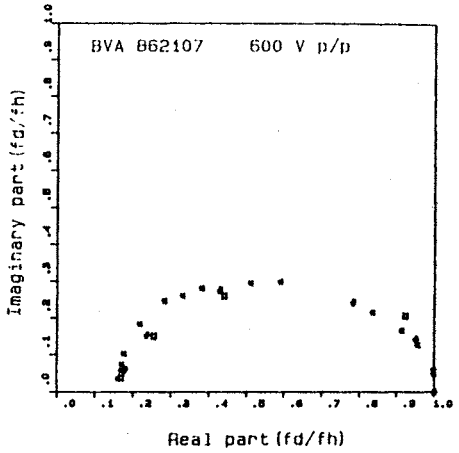
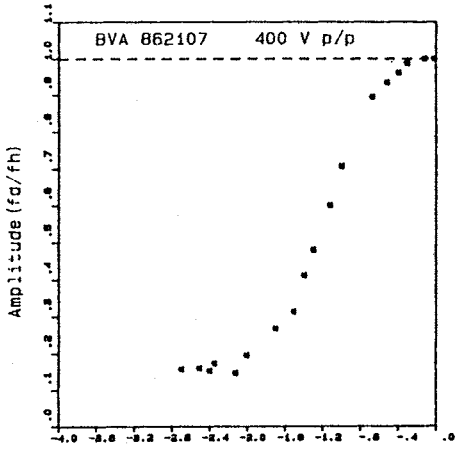


Fig. 9 : Experimental data of Figs. 7 and 8 presented in a Nyquist plot. Whatever the temperature, the points lie approximately on the same curve.



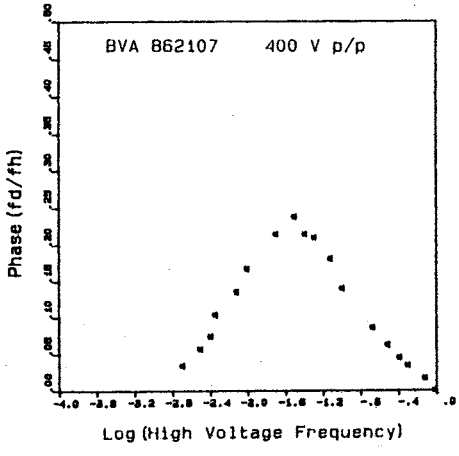


Fig. 10-11

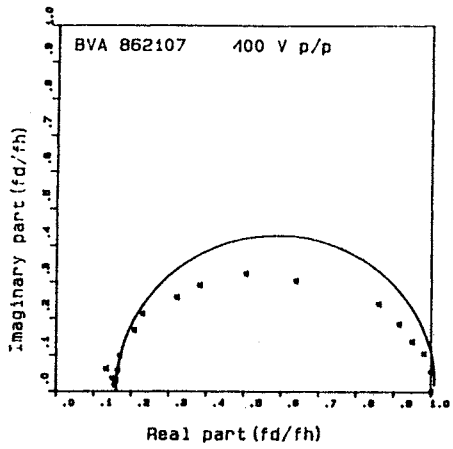


Fig. 10-11-12 : A sequence of the same kind as Figs. 7, 8, 9, but with 400 V peak-to-peak. Note that these curves do not differ appreciably from the previous ones.

Diffusion parameters

When the corresponding theoretical half circle is superposed on the experimental Nyquist plot it is seen that the two plots do not fit at all points (Fig. 12).

A possible explanation could lie in the fact that in the calculation only the first harmonic is considered. In actual fact higher harmonics exist and may distord the curve.

Nevertheless the plots in Fig. 12 allow one to obtain an approximate value of the coefficient κ.

In the present case κ is found to be about 6.10^9 m^{-2}, which corresponds to an impurity concentration of about 5.10^{16} charges/m^3. For this value of κ the theoretical phase curves (Fig. 5) show that the maximum phase difference is obtained when the ratio $\Omega/D \approx 1.1\ 10^8$ m^{-2}. Knowing the excitation frequency for which the phase difference is maximum it is possible to obtain the diffusion constant and the mobility. Table I summarizes the results obtained at various temperatures.

It is well known that the diffusion constant follows Arrhenius' law

$$D = D_o\, e^{-\frac{W_m}{k_B T}} \qquad (21)$$

In order to determine the migration energy W_m, the diffusion constant is usually plotted in an Arrhenius plane whose coordinates are respectively the inverse absolute temperature and the natural logarithm of the diffusion constant. In such a plane, relation (21) is a linear function, the slope of which is proportional to the migration energy.

Figure 13 shows the result obtained with the data listed on Table I, the points form a line whose slope corresponds to the migration energy W_m = 1.45 eV.

Such a value is typically in the range of the known values of the ionic diffusion energy. Here, the ions involved are probably sodium which is the most common impurity found in synthetic quartz crystals.

Conductivity measurements made in the past give a diffusion constant for sodium of about 1.04 eV [16].

Alternatively, radioactive tracer ^{22}Na was used to measure the diffusion constant, either parallel to the C-axis (0.87 eV) or normal to the C-axis (1.78 eV) [17].

It is to be noted that the plate used in the high field relaxation method is a doubly rotated cut so that the diffusion occurs neither parallel nor normal to the C-axis. Consequently the value of 1.45 eV obtained here is consistent with the range of the known data.

Table I

T (°C)	frequency of the maximum phase shift	D (m^2/s)	μ (m^2/V/s)
100	$5\ 10^4$	$2.8\ 10^{11}$	$8.71\ 10^{10}$
120	$5\ 10^3$	$2.8\ 10^{10}$	$8.24\ 10^9$
140	$4\ 10^2$	$2.24\ 10^9$	$6.3\ 10^8$

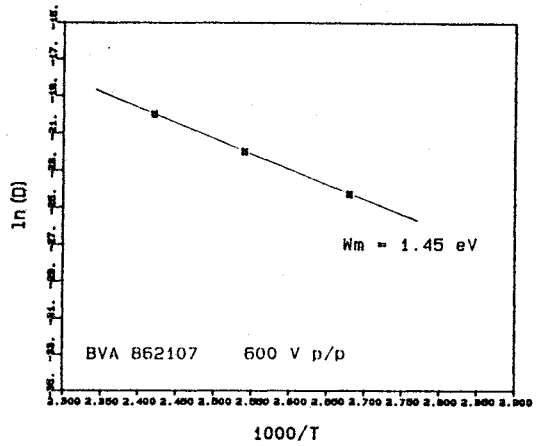


Fig. 13 : The impurity migration energy can be obtained from an Arrhenius' plot. In this case it is found to have a value of 1.45 eV.

Conclusion

The dynamic high field relaxation method is among the most sensitive methods for dealing with ionic impurity characterization, although its implementation requires some restrictive conditions. The crystal has to be piezoelectrically excited and has to exhibit an electroelastic effect. In addition, certain electrode effects, not yet explained, lead to rather complicated dynamic relaxation phenomena. To avoid these effects, electrodeless resonators must be used [13].

At present effort is being focused on understanding these electrode effects in order to extend the measurement method to more easily available electroded resonators. Furthermore, from simultaneous measurements of current and frequency shift it is possible to obtain new information about the impurity contained in the crystal and to improve the measurement accuracy of the kinetic parameters.

It should be noted that the dynamic high field relaxation method can eventually be useful for studying the effect of radiation on quartz crystal resonators. In fact some preliminary experiments show that radiation seems to induce on swept resonators, a weak high field relaxation with a large time constant [15], the origin of which remains unclear.

Acknowledgement

The author wishes to thank Prof. B.A. Auld for his interest and his help in revising the manuscript.

References

[1] C.K. Hruska, Czechoslovak Journal of Physics, B11, 150 (1961).

[2] C.K. Hruska, Czechoslovak Journal of Physics, B12, 338 (1962).

[3] C.K. Hruska, IEEE Trans. on Sonics and Ultrasonics, SU-18, 1 (1971).

[4] C.K. Hruska, IEEE Trans. on Sonics and Ultrasonics, SU-25, 198 (1978).

[5] C.K. Hruska, IEEE Trans. on Sonics and Ultrasonics, SU-25, 390 (1978).

[6] J.A. Kusters, Proc. of the 24th AFCS (1970), p. 46.

[7] R. Brendel, J. Appl. Phys. 54, 5339 (1983).

[8] G.A. Reider, E. Kittinger, J. Tichy, J. Appl. Phys. 53, 8716 (1982).

[9] E. Kittinger, J. Tichy, W. Friedel, J. Appl. Phys. 60, 1465 (1986).

[10] C.K. Hruska, R. Brendel, J. Appl. Phys. 65, 715 (1989).

[11] R. Brendel, J.J. Gagnepain, Proc. of the 36th AFCS (1982), p. 97.

[12] R. Brendel, J.J. Gagnepain, J.P. Aubry, Proc. of the 40th AFCS (1986), p. 121.

[13] R. Brendel, 1st European Freq. and Time Forum (1987), p. 153.

[14] J.A. Weil, Rad. Effects 26, 261 (1975).

[15] G.Marianneau, R. Brendel, Rapport DRET 86/044 (1987).

[16] J. Verhoogen, Am. Mineralogist 37, 637 (1952).

[17] G.H. Frischat, Berichte der Deutsch-Keram. Ges. 47 (1970), pp. 238, 313, 364, 635.

This work was supported by *Direction des Recherches Etudes et Techniques.*

THE AT AND SC RESONATORS IN A DC FIELD OF ARBITRARY DIRECTION

Carl K. Hruska

Piezoelectricity Research Laboratory, York University
Toronto, Ontario M3J 1P3, Canada

Abstract

The complete set of the effective electroelastic constants of quartz is calculated for the first time using the data on the frequency-dc field dependence of quartz resonators.

The newly obtained electroelastic constants are used to map the dc field-induced frequency shifts of the thickness modes of quartz plates as a function of the dc field direction. Applications made to the AT and the SC resonators show the locus of the dc field directions with a zero frequency-dc field effect and the size and variations of the effect for the dc fields in the lateral directions (relative to the resonator plates).

The determined electroelastic constants are in good agreement with their counterparts obtained by a different method and the available experimental data confirm the predictions of the frequency-dc field dependence made.

Introduction

The principal objective of this paper is the mapping of dc field-induced frequency shifts of the thickness modes of quartz plates as a function of the dc field direction.

Ideally the study would be undertaken in terms of the nonlinear theory using the true (thermodynamically defined) material constants of quartz but the available applications of the nonlinear theory do not cover the case under consideration.

On the other hand, past experience shows that predicting and mapping the frequency-dc field dependence has always been dealt with very successfully in an approximate form using the effective electroelastic constants [1,2,3]. A suitable approximate expression for the frequency-dc field dependence covering the case of an arbitrarily directed biasing field is obtained here by extending the formula used with success for the biasing fields in the resonator plate thickness and in its lateral directions [4,5].

Since the mapping to be made requires that all effective electroelastic constants be known, the obtained expression is first used to find the missing values [3] of the electroelastic constants intended to be used. Subsequently, it is applied to the AT and the SC resonators and used for mapping the frequency-dc field dependence in several selected cases which may be of interest or serve an illustrative purpose.

This work uses material constants and experimental quantities originating from several different sources. During the calculations done here it was impossible to avoid mixing the data determined at various temperatures: 20•C, 25•C, or room temperature. Consequently, the results of this work can at best be said to apply to room temperature.

All numerical quantities stated in this paper are given for right-hand quartz and the frame of reference according to the IRE 1949 Standard [6].

Linear coefficient of the frequency-dc field dependence

Consider a rectanguar quartz plate (Fig.1.) with its thickness t, width w and length l parallel to an orthogonal frame of reference X', Y' and Z'. The position of the plate relative to the basic reference frame X, Y and Z in the quartz crystal is arbitrary and described by the rotational symbol $(xzlw)\phi/\theta$ according to [6].

Assume that the body of the plate can be subjected to a uniform dc electric field E' of an arbitrary direction defined by the angles σ, $\sigma > 0$, and τ, where the angle τ is regarded as positive when measured in the direction from +Y' to +Z'.

When the plate is subjected to a dc field $\Delta E'_1$ it generally changes the frequency f of its thickness vibrations by Δf. The relative frequency change can be recorded as

$$\Delta f/f = (1/f)(df/dE')_{E'=0}\cdot\Delta E' + (1/2f)(d^2f/dE'^2)_{E'=0}\cdot(\Delta E')^2 .$$

The quadratic component is negligible and the frequency-dc field dependence of the thickness vibrations is sufficiently described in terms of its linear coefficient L

$$(1) \qquad L = (1/f)(df/dE')_{E'=0} ,$$

which can be expressed as

$$(2) \quad L = c_1 d_{11} + c_2 d_{14} + a_1 f_{111} + a_2 f_{113} + a_3 f_{114} + a_4 f_{122} + a_5 f_{124} + a_6 f_{134} + a_7 f_{144} + a_8 f_{315} .$$

Here d_{11} and d_{14} are the piezoelectric strain constants and the quantities

$$(3) \quad f_{111},\ f_{113},\ f_{114},\ f_{122},\ f_{124},\ f_{134},\ f_{144},\ f_{315}$$

are the eight effective electroelastic constants of quartz; both are related to the basic reference frame in quartz.

The coefficients c_1, c_2, and a_j, $j = 1,2,\ldots,8$, are relatively complicated functions of the four angles ϕ, θ, σ, and τ, of the mode of the thickness vibrations of the resonator and of the linear material constants of quartz: elastic stiffnesses, piezoelectric stress constants and dielectric permittivities.

The derivation of Eq.(2) follows the steps described in detail for the case of the dc field in thickness [4] except for the differentiation which has to be made for a dc field E' of an arbitrary direction. The derivation presents no difficulty; it is only tedious and the details are unwieldy. For the purpose of this work its result has been given only in a simple qualitative form.

Eq.(2) can be used in determining the effective

electroelastic constants of quartz (3) from observations of the linear coefficient L. If all eight electroelastic constants are available, Eq.(2) can be used for the mapping of the linear coefficient L as a function of the four angles ϕ, θ, σ, and τ.

Experimental data to calculate the unknown electroelastic constants

The most recent attempt to determine the effective electroelastic constants of quartz was based on observations of the linear coefficient L for the plates vibrating in thickness with the dc biasing field also in the thickness direction of the plate (σ = 0°). The work [3] resulted in the determination of six combinations of the eight constants f_{ijk} in (3) as shown in Table I. The remaining electroeastic constants f_{122} and f_{134} can also be calculated using the thickness vibrations if only the observations of L are made for the dc biasing fields of directions other than that of the thickness of the plates ($\sigma \neq 0°$).

Such observations have been made using the thickness vibrations of a plate of orientation (*xzlw*)105.07/-9.24 and a lateral (σ = 90°) dc field of various directions $\tau \in [-75° ; 60°]$ as shown in Fig.2. In this experiment the linear coefficient L was measured for the fundamental frequency of the thickness modes C and A. Its respective values L^C and L^A are recorded in Table II. More details can be found in [7].

The experimental values of the linear coefficient L in Table II were calculated as a ratio $\Delta f/(f.\Delta E')$, where Δf was the linear component of the change in the resonance frequency f caused by the biasing dc field $\Delta E'$. For the type of the electrodes used the dc field in the space between the electrodes was not homogeneous; its average value obtained by a numerical solutions of the Poisson equation was accepted as $\Delta E'$.

The vibrations of the plate described above are excited in the central unelectroded part of the resonator but spread throughout its entire body due to the untrapping effect of the mass of the exciting electrodes. On the other hand, the lateral dc field can reach and affect only the vibrations in the central part of the resonator. As a result, the measured frequency shift Δf is substantially smaller in absolute value than it would be if the whole vibrating mass of the resonator were exposed to the biasing field [8]. Consequently, the experimental values of L^C and L^A in Table II do not match the theoretical value of L in Eq.(3) derived on the assumption that the whole body of the resonator is exposed to the biasing field.

To rectify the situation each value of L in Table II should be multiplied by a (unknown and generally different) factor K > 1. Numerical estimates made in the course of work [8] with the help of the effective electroelastic constants borrowed from [9] suggest that - in the first approximation - the factor K is dependent on the mode of vibrations and independent of the direction of the lateral field. Consequently, to make a corresponding correction of the observed values L^C and L^A in Table II, only two different factors K, K^C and K^A, need to be used, one for each mode of vibrations. The corrected values pertaining to the mode C and A are thus formally recorded as $K^C L^C$ and $K^A L^A$, respectively.

Calculation of the electroelastic constants f_{122} and f_{134}

The twenty observations of L_i^C (i = 1,2,...,10) and L_i^A (i = 11,12,...,20) from Table II have been formally corrected by the introduction of the (hitherto numerically unknown) factors K^C and K^A as suggested in the previous section. Eq(3) has been applied to the corrected observations $K^C L_i^C$ and $K^A L_i^A$.

A system of linear equations has been obtained consisting of two sets of equations; ten equations (i = 1,2,...,10) for the mode C

$$(4)\quad K^C L_i^C - C_i = a_{11} f_{111} + a_{12} f_{113} + a_{13} f_{114} + a_{14} f_{122} + a_{15} f_{124} + a_{16} f_{134} + a_{17} f_{144} + a_{18} f_{315},$$

and ten (i = 11,12,...,20) for the mode A

$$(5)\quad K^A L_i^A - C_i = a_{11} f_{111} + a_{12} f_{113} + a_{13} f_{114} + a_{14} f_{122} + a_{15} f_{124} + a_{16} f_{134} + a_{17} f_{144} + a_{18} f_{315},$$

where $C_i = c_{11} d_{11} + c_{12} d_{14}$, for i = 1,2,...,20. The values of the quantities C_i and a_j, j = 1,2,...,8, have been calculated for the fixed orientation angles ϕ = 105.07°, θ = -9.24°, σ = 90° as a function of the lateral field direction angle τ and the mode of vibration pertinent to individual observations. The values of the linear material constants necessary for the calculation have been taken from [10].

At this point the number of unknown quantities in Eqs.(4) and (5) was then reduced by employing the earlier results from Table I. Having isolated all terms containing unknown quantities on the right hand side the linear system could be rewritten

$$(6)\quad b_{i0} - C_i = b_{i4} f_{122} + b_{i6} f_{134} - L_i^C K^C$$

where i = 1,2,...,10, and

$$(7)\quad b_{i0} - C_i = b_{i4} f_{122} + b_{i6} f_{134} - L_i^A K^A$$

where i = 11,12,...,20, and, further

$$b_{i0} = 2.88\, a_{i1} - 0.29\, a_{i2} - 0.91\, a_{i3} - 0.80\, a_{i5} + 0.91\, a_{i7} + 0.71\, a_{i8},$$

$$b_{i4} = a_{i4} + 0.42\, a_{i2} + 0.26\, a_{i3} + 0.76\, a_{i5} - 0.50\, a_{i7} - 0.60\, a_{i8},$$

$$b_{i6} = a_{i6} - 0.25\, a_{i2} + 0.67\, a_{i3} + 0.67\, a_{i5} + 0.51\, a_{i8},$$

where i = 1,...,20 .

Our system of twenty linear equations (Eqs.(6) and (7)) now contains only four unknown quatities K^C, K^A, f_{122} and f_{134}. Their most likely values have been determined using the least square fit. The results of this process and their standard errors are listed in Table III.

Table IV presents the complete set of the effective electroelastic constants of quartz obtained by combination of the results in Tables I and III. It is the first set of electroelastic constants that has been determined using the frequency-dc field dependence of quartz resonators (called the resonator method).

Table IV also contains the most recent values of the effective electroelastic constants obtained from transit time observations [9]. The difference between the two sets of constants is about 30 percent on the average.

The correction factors K^C and K^A are a byproduct of our calculations. The suitability of their values can

be visualized by comparing the corrected values of L^C and L^A from Table II with their counterparts calculated using Eq.3.This is done in Table V.

Application to the AT and SC cuts

Using Eq.(3) and the effective electroelastic constants just calculated predictions of the magnitude of the linear coefficient L of the frequency-dc field dependence have been made for the resonators of the orientation (*xzlw*)90°/-35.25° (the AT cut) and (*xzlw*)111°/-33.50° (the SC cut).

Figures 3 and 4 show the results for the dc fields of σ = 90° and 0°≤ τ≤ 180° (the lateral field).

As the actual conditions of the experiment do not allow the creation of a uniform dc field acting in the whole volume of the resonator the calculated values of L are typically larger than their measured counterparts. Consequently, an agreement in the shape of the experimental and the theoretical curve, such as seen in Figs. 3 and 4, is to be regarded as the best confirmation of the theory that can be achieved.

Figures 5 - 8 show the loci of orientations of the dc field which will generate no frequency change (L = 0) as may be of interest in high stability applications. Also shown are the directions of the dc field corresponding to the extreme values of L.

Figures 9 and 10 present a mapping of the linear coefficient L for dc field directions in a two dimensional domain of the angles $\sigma \in$ (0°,90°) and $\tau \in$ (90°,270°).

Conclusion

This paper deals with the interactions between the the three thickness modes of vibrations of quartz plates and an electric field acting on the body of the plates (frequency-dc field dependence). Its principal result is an expression for the linear coefficient of the frequency-dc field dependence applicable to plates of any orientation and to any direction of the dc field.

The expression was used to calculate the locus of the electric field directions for the AT and the SC resonators where the resonator frequency is insensitive to the electric field. This may be of importance when designing the resonator electrodes while trying to minimize the effect of the exciting field on the resonator frequency stability.

Detailed mapping of the linear coefficient of the frequency-dc field dependence was also made for the C mode of the AT and the SC resonators to show how strongly the frequency-dc field effect depends on the lateral field direction.

The above calculations required the knowledge of all eight effective electroelastic constants of quartz which this work also attempted to calculate for the first time from the observations of the frequency-dc field dependence of quartz resonators.

The work used several simplifying assumptions both in the derivation of the theoretical expression for the linear coefficient of the frequency-dc field dependence as well as in the experimental data reduction process.

In spite of the simplifications the obtained results seem to be quite harmonious. The effective electroelastic constants agree surprisingly well with similar quantities obtained independently and by a different experimental method. The agreement is by far the best reached hitherto for the electroelastic constants from different sources. Also the predictions of the linear coefficient of the frequency-dc field dependence are strongly supported by the available experimental data.

A more rigorous treatment of the problem will have to wait for a model of the thickness vibrations subjected to a dc biasing field with the exciting field divorced from the thickness direction of the plate. The results obtained here represent a promise and an encouragement for further work in this direction.

Acknowledgments

The author wishes to acknowledge the assistance with the experimental part of the work received from Prof. R. Besson, Dr. J. F. Darces, Dr. G. Genestier, Prof. H. Merigoux and Mr. M. Kucera, M.Eng.Sc. Miss P. Hruska helped with the final stages of preparation of this paper. This work was supported by the York University-Faculty of Arts Fellowship and by the National Sciences and Engineering Research Council of Canada.

References

[1] K. Hruska and A. Khogali, "Polarizing effect with alpha-quartz rods and the electroelastic tensor," IEEE Trans. Sonics Ultrason., vol. SU-18, pp. 171-176, July 1971.

[2] E. Kittinger and J. Tichy, "Mapping of the electroelastic effect in quartz based on fundamental material constants," to appear in the Proc. 1st European Frequency and Time Forum, Communautes Scientifiques de Besancon (France) et Neuchatel (Suisse), pp. 222-225, March 1987.

[3] C. K. Hruska, H. Merigoux and M. Kucera, "Linear frequency- dc field effect with the thickness modes of alpha-quartz plates", Journal of the Canadian Ceramic Society, vol. 57, No.2, pp. 53-57, July 1988.

[4] K. Hruska, "Polarizing effect with piezoelectric plates and second-order effects," IEEE Trans. Sonics Ultrason., vol. SU-18, pp. 1-7, Jan. 1971.

[5] C. K. Hruska, "Lateral dc electric field in quartz resonators", Proc. of the 2nd European Frequency and Time Forum, Communautes Scientifiques de Besancon (France) et Neuchatel (Suisse), Neuchatel, pp.621-627, March 1988.

[6] Standards on Piezoelectric Crystals, 1949, Proc. IRE 37, pp. 1378-1395, Dec. 1949.

[7] C. K. Hruska and M. Kucera, "The thickness modes of a quartz plate resonator in a lateral dc field", IEEE Trans. Sonics Ultrason., vol. SU-32, pp. 600-3, July, 1985.

[8] C. K. Hruska, "Lateral dc electric field in quartz plate resonator", Proc. 2nd European Frequency and Time Forum, Communautes Scientifiques de Besancon (France) et Neuchatel (Suisse), pp.621-627, March 1988.

[9] E. Kittinger, private communication.

[10] R. Bechmann, "Elastic and piezoelectric constants of alpha-quartz", Phys. Rev., vol. 110, pp. 1060-1061, June 1, 1958.

List of captions

Table I. Values of the effective electroelastic constants and their combinations according to [3]. The errorsr are standard errors. Given in N/(V.m) for 25°C, right hand quartz and a frame of reference according to IRE Standard [6].

Table II. Observations of the linear coefficient of

the frequency-dc field dependence for the fundamental frequency of the thickness modes C and A of a quartz plate of orientation (*zxlw*)105.07°/-9.24°. Obtained for lateral dc fields (σ = 90°) of various directions τ . The values L^C and L^A are raw observations not corrected for the effect of energy 'untrapping'. Given in 10^{-12} m/V, for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Table III. Effective electroelastic constants f_{122} and f_{134} determined in this work. Given in N/(V.m), for room temperature, right hand quartz and a frame of reference according to the IRE Standard [6].

Correction factors K^C and K^A have been obtained simultaneously. The errors are standard errors.

Table IV. Complete set of the effective electroelastic constants determined using the thickness modes of plates and the dc electric field in the thickness and the lateral directions. Obtained by combination of results in Tabs.I and III. Corresponding values from an source [9] are stated for comparison. All values are given in N/(V.m), for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Table V. Comparison of the experimental and calculated values of the linear coefficient L with the inclusion of the correction factors K^C and K^A. The comparison characterizes the degree of fit attained during the least square process leading to determination of the effective electroelastic constants f_{122} and f_{134} in Tab.III. The values are given in 10^{-12}m/V, for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

f_{111}	-2.88 + 0.03
$f_{113} - 0.42\ f_{122} + 0.25\ f_{134}$	-0.29 + 0.05
$f_{114} - 0.26\ f_{122} - 0.67\ f_{134}$	-0.91 + 0.04
$f_{124} - 0.76\ f_{122} - 0.67\ f_{134}$	-0.80 + 0.04
$f_{144} + 0.50\ f_{122}$	0.91 + 0.02
$f_{315} + 0.60\ f_{122} - 0.51\ f_{134}$	0.71 + 0.04

Table I.

τ	L^C	L^A
-75°	-0.858	-3.350
-60°	-1.613	-3.390
-45°	-1.600	-5.700
-30°	-1.708	-7.292
-15°	-1.794	-6.892
0°	-1.460	-5.632
15°	-1.587	-4.379
30°	-0.953	-3.771
45°	-0.585	-1.906
60°	-0.189	-0.839

Table II.

f_{122}	= 0.93	±	0.12
f_{134}	= 2.13	±	0.20
K^C	= 4.59	±	0.30
K^A	= 1.65	±	0.20

Table III.

	this work	[14]
f_{111}	-2.88	-2.80
f_{113}	-0.84	-0.54
f_{114}	0.77	0.46
f_{122}	0.93	1.55
f_{124}	1.35	1.41
f_{134}	2.13	1.61
f_{144}	0.44	0.18
f_{315}	1.24	0.92

Table IV.

τ	$K^C L^C$		$K^A L^A$	
	exp	cal	exp	cal
-75°	-3.937	-4.575	-5.538	-6.717
-60°	-7.402	-5.943	-5.604	-8.635
-45°	-7.342	-6.906	-9.422	-9.965
-30°	-7.838	-7.399	-12.054	-10.616
-15°	-8.233	-7.387	-11.392	-10.543
0°	-6.700	-6.872	-9.310	-9.752
15°	-7.283	-5.888	-7.238	-8.296
30°	-4.373	-4.503	-6.233	-6.274
45°	-2.685	-2.812	-3.151	-3.825
60°	-0.867	-0.928	-1.387	-1.116

Table V.

Figure 1. Linear coefficient L of the frequency-dc field dependence is a function of four angles: ϕ, θ (defining the orientation of the resonator (*xzlw*)ϕ/θ and the direction of the resonator thickness X') and σ and τ (defining the direction of the dc field vector E'). The angle $\sigma \geq 0°$ measures the angle between +X' and the field vector E'. The angle τ is positive when measured in the direction from +X' to +Z'. The axes X, Y, and Z form the basic reference frame in quartz according to IRE Standard [6].

Figure 2. Geometry of the resonator electrodes for the measurement of the linear coefficient L in a lateral field E'. The electrodes cover the opposite parts of the crystal plate; its unelectroded part is marked by a dotted borderline. Other remarks similar to those made for Fig.1.

Figure 3. Dependence of the linear coefficient L on the dc field direction angle τ for σ = 0° (lateral field). Obtained for the fundamental frequency of the mode C of an AT resonator (orientation (*xzlw*)90°/-35.25°). The shape of the predicted curve (solid line) agrees well with the available observations (dotted line). Observations were made on three resonators of type I [8] of of overall dimensions 19 mm x 19mm x 0.64 mm at the frequency of 2.6 MHz, using electrode separation of 5 mm and a dc potential across the electrodes of 2 kV. Given for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Figure 4. Dependence of the linear coefficient L on the dc field direction angle τ for σ = 0° (lateral field). Obtained for the fundamental frequency of the mode C of an SC resonator (orientation (*xzlw*)111°/-33.50°). The shape of the predicted curve (solid line) agrees well with the available observations (dotted line). Observations were made on two resonators of type II [8] of overall dimensions 19 mm x 19mm x 1.01 mm at a frequency of 1.8 MHz, the circular mesa height 0.1 mm, mesa diameter and

electrode separation of 7 mm and a dc potential across the electrodes of 2 kV. Given for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Figure 5. Locus of dc field directions (σ,τ) where the linear coefficient L = 0 as obtained for the AT cut. The locus is a straight line identical for all three thickness modes. Points P_1 and P_2 indicate the dc field directions for which the linear coefficient L reaches absolute extreme values; at P_1 these are -12.1, 0.7 and -15.3 (all in 10^{-12} m/V) for the mode C, B and A, respectively; at P_2 the extreme values are the same in magnitude but opposite in sign. Given for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Figure 6. Locus of dc field directions (σ,τ) with zero linear coefficient L for the thickness mode C of the SC cut. Point P indicates the dc field direction, where L = $-14.5\ 10^{-12}$ m/V reaching the maximum absolute value of L for this mode. Given for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Figure 7. Locus of dc field directions (σ,τ) with zero linear coefficient L for the thickness mode B of the SC cut. Point P indicates the dc field direction, where L = $-4.7\ 10^{-12}$ m/Vreaching the maximum absolute value of L for this mode. Given for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Figure 8. Locus of dc field directions (σ,τ) with zero linear coefficient L as obtained for the thickness mode A of the SC cut. Point P indicates the dc field direction, where L = $-14.5\ 10^{-12}$ m/V reaching the maximum absolute value of L for this mode. Given for room temperature, right hand quartz and a frame of reference according to IRE Standard [6].

Figure 9. Mapping of the linear coefficient L for a two dimensional domain of the dc field directions defined by the angles $\sigma \in (0°,90°)$ and $\tau \in (90°,270°)$. Valid for mode C of the AT cut. The coefficient L is plotted in 10^{-12} m/V.

Figure 10. Mapping of the linear coefficient L for a two dimensional domain of the dc field directions defined by the angles $\sigma \in (0°,90°)$ and $\tau \in (90°,270°)$. Valid for mode C of the SC cut. The coefficient L is plotted in 10^{-12} m/V.

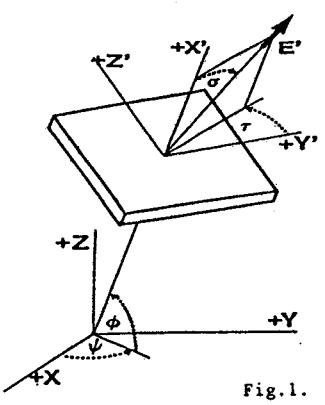


Fig.1.

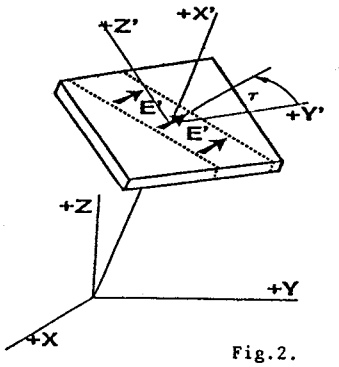


Fig.2.

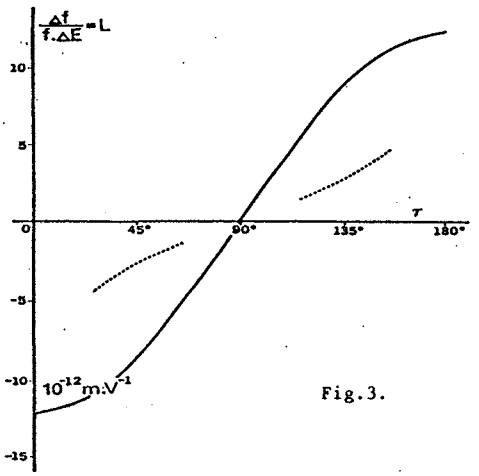


Fig.3.

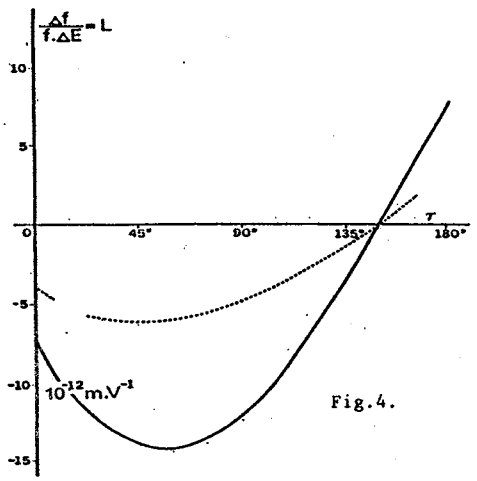


Fig.4.

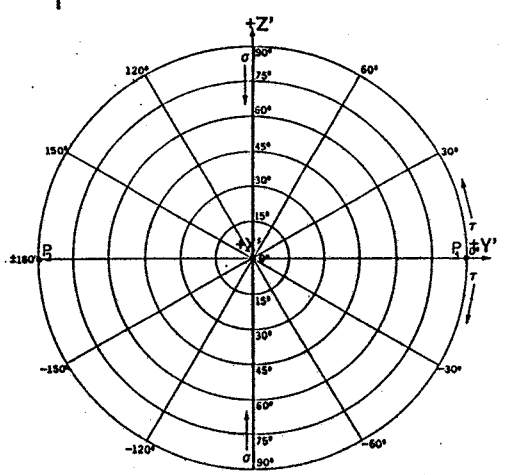


Fig.5.

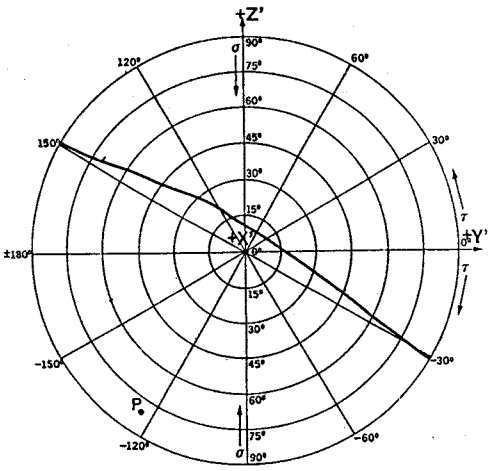


Fig.6.

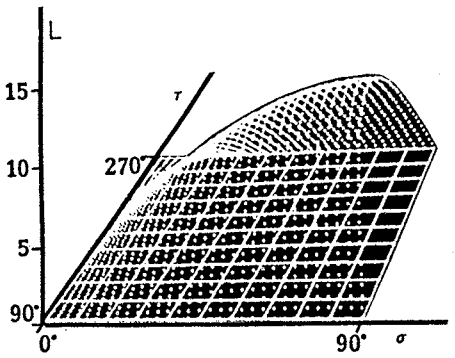


Fig.9.

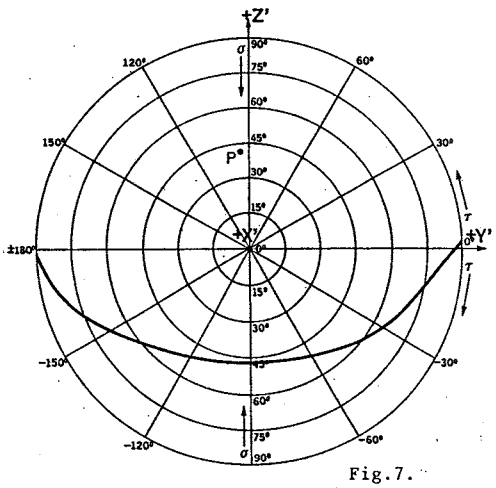


Fig.7.

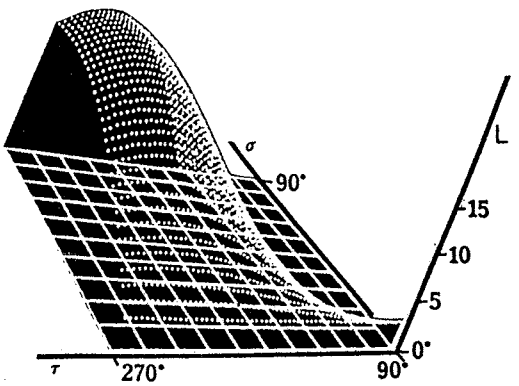


Fig.10.

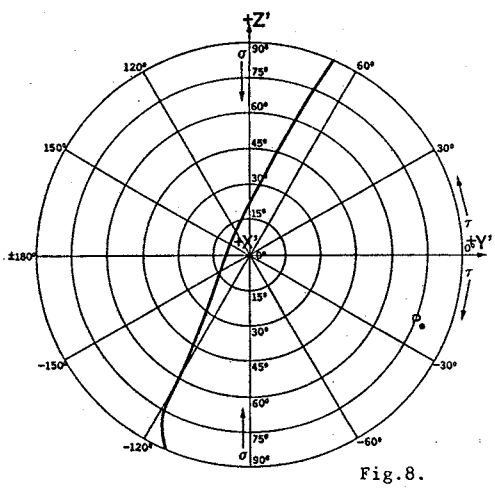


Fig.8.

SIMULATION OF THE CONTOUR SHAPE FOR ETCHED QUARTZ CRYSTALS

T. LEBLOIS, P. MAITRE and C. TELLIER

Laboratoire de Chronométrie, Electronique et Piézoélectricité (E.N.S.M.M.)
Route de Gray - La Bouloie - 25030 BESANCON CEDEX - FRANCE

ABSTRACT

Using a tensorial analysis of the dissolution, the successive dissolution shapes of a starting circular section are derived from the equation of the representative surface of the dissolution slowness for the quartz crystal. Firstly the polar diagrams of the dissolution slowness vector $\vec{L}$ in the (X, Y), (X, Z) and (Y, Z) planes are presented as deduced from experimental data. The theoretical dissolution shapes and the corresponding out-of-roundness profile graphs of etched X, Y and Z quartz plates are then derived. The comparison of the theoretical profile graphs with the experimental ones for plates etched in concentrated ammonium bifluoride solutions shows a satisfactory agreement. Experimental results on the dissolution profile graphs related to various singly-rotated quartz plates reveal also that the dissolution of quartz crystals is completely orientation dependent. Finally the application of the method to the dissolution shapes of singly-rotated quartz plates is described.

1 - INTRODUCTION

The interest in preparing quartz resonator plates by chemical etching has been revived in the past few years (1-9). But a lot of experimental data (2, 6-13) related to both dissolution figures formation and variations of etch rate with surface orientations have undoubtly proved that for quartz plates immersed in concentrated $NH_4F.HF$ or HF solutions the changes in the shape of dissolution figures with orientation can be explained only in terms of the variations of the etch rate. Hence, if it is important to understand how an etchant produces dissolution etch pits or etch hillocks which can markedly affect the performance of a resonator by inducing a marked decrease in the Q-factor (14, 15) it is also necessary, for a chemical attack governed by orientation, to consider the changes in the general shape of the resonator on repeated etchings (16).

Thus the work reported in this paper started as an attempt to explain the observed dissolution shapes of section of quartz crystals in the framework of the recent tensorial analysis of the dissolution proposed by Tellier et al (17, 18). This tensorial method which constitutes a wide extension of the two-dimensional kinematic theory of the dissolution developped several years ago by Frank (19) allows us to determine in three dimensions the propagation vector $\vec{P}$ (20) related to a surface element of any orientation (Ψ, θ) from the equation of the representative surface of the dissolution slowness vector, $\vec{L}$. Hence we can track during the dissolution the various moving surface elements which propagate within the bulk crystal along linear trajectories (19, 20) and then we can construct numerically the successive general dissolution shapes of a crystal.

2 - THE THREE FUNDAMENTAL POLAR DIAGRAMS OF THE DISSOLUTION SLOWNESS FOR THE QUARTZ CRYSTAL

The general form, $L(\Psi, \theta)$ of the equation for the dissolution slowness surface is intimately connected with the symmetry of the crystal into consideration (17). Thus taking into account the degree of symmetry of the quartz crystal to reduced the number of dissolution constants for the dissolution tensors involved in the equation of the slowness surface a generalized form

$$L_{XYZ}(\Psi, \theta) = L_o + \sum_{j=1}^{m} A_j \cos^{2j}\theta + \sum_{k=1}^{n} B_k \cos^{2k+1}\theta \sin 3\Psi + \sum_{\ell=1}^{n} C_\ell \cos^{2\ell+1}\theta \sin\theta \cos 3\Psi \qquad (1)$$

was finally found (17).

The projection of the slowness surface in the reference planes (OXY) (OXZ) and (OYZ) gives the corresponding fundamental polar diagram for the dissolution slowness as

$$L_{XY}(\Psi, \theta = 0°) = A_Z + B_Z \sin 3\Psi \qquad (2)$$

with

$$A_Z = L_o + \sum_{j=1}^{m} A_j \qquad (3)$$

and

$$B_Z = \sum_{k=1}^{n} B_k \qquad (4)$$

$$L_{XZ}(\Psi = 90°, \theta) = L_o + \sum_{j=1}^{m} A_j \cos^2\theta - \sum_{k=1}^{n} B_k \cos^{2k+1}\theta \qquad (5)$$

$$L_{YZ}(\Psi = 0°, \theta) = L_o + \sum_{j=1}^{m} A_j \cos^{2j}\theta + \sum_{\ell=1}^{n} C_\ell \cos^{2\ell+1}\theta \sin\theta \qquad (6)$$

respectively.

It should be pointed out that when we are concerned with a polar diagram we can only treat two-dimensional problems but a theory in two-dimensions suffices to derive the dissolution

shapes for an initially circular quartz plate. Moreover to use the fundamental polar diagrams we have to determine from experimental data the unknown constants, L_o, A_j, B_k and C involved in Equations (1) to (6).

For the L_{XZ} polar diagram this was done by evaluating the dissolution slowness from a magnified profile graph of a deeply etched Y-cut section **(16)** which represents the changes in roundness of the circular starting shape. We adopted this procedure rather than the procedure which consisted of measuring the etch rate of various Y-rotated quartz plates because the dissolution slowness differs for the two faces of an Y-rotated plate.

To avoid the design and preparation of a great number of singly rotated quartz plates (X-rotated quartz plates) with angle of cut, , varying from 0° to 180° by two degrees only the unknown constants appearing in Equation (6) were determined using the profile graphs related to the etched section of four differently oriented singly rotated quartz plates. Then, the establishment of the L_{XY} equation in which the number of unknown constants is finally reduced to 2 is easily accomplished by estimating the values of these constants from the previous results for the L_{XZ} and L_{YZ} equations.

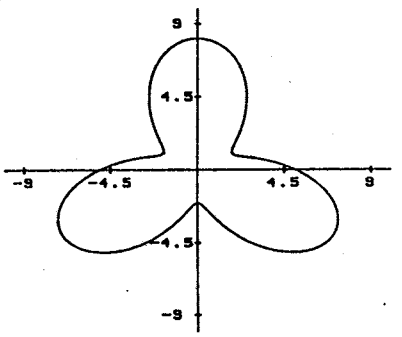


Figure 1 : Polar path of the dissolution slowness L_{XY}.

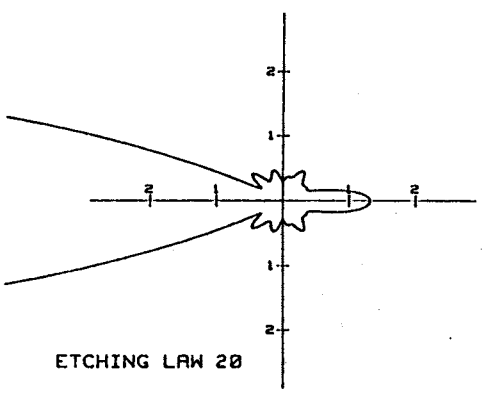


Figure 2 : Polar path of the dissolution slowness L_{XZ}.

The theoretical shapes of the L_{XY}, and L_{XZ} polar diagrams are displayed in Figures 1 to 2. We observe marked differences in the shape of these fundamental polar graphs. Thus the dissolution of quartz crystals in concentrated $NH_4F.HF$ solution is highly anisotropic as predicted by some authors **(21)** and one can reasonably expect a significant disparity in the dissolution shapes related to differently oriented quartz plates. It may be also of interest to note that

(i) the L_{XY} polar diagram of Figure 1 satisfies the three-fold symmetry about the Z axis.

(ii) Even if the L_{XZ} polar diagram exhibits a rather complicated shape it follows the two-fold symmetry related to the X axis.

3 - DISSOLUTION SHAPES OF THE X-, Y- and Z-CUT PLATES

3.1/ Experimental : nature of dissolution shapes

The etching phenomena described in this work were produced by a concentrated $NH_4F.HF$ solution etchant. Four specimen of the X-, Y- and Z- cut circular plates (15 mm in diameter, 2.5 mm in thickness) were cut from the same synthetic quartz crystal. After cutting the planar surfaces and the circular contours were lapped with a 5 m abrasive. In addition, the contours were polished to remove damaged surface layer. For each orientation (Ψ, θ) with (Ψ, θ) respectively equal to ($\Psi = 90°, \theta = 0°$), ($\Psi = 0°, \theta = 0°$) and ($\Psi = 0°, \theta = 90°$) one of the specimen possesses a flat mark to locate easily the direction of the X or Y axes. The plates were etched at a constant temperature of 333 $\pm$ 1° K for successive periods of time.

The changes in shape of the starting circular section were studied by using a Talyrond analyser which generates the least square circle so that the dissolution profiles were displayed at relatively large magnifications with the superimposed reference circle to give graphs which for convenience are called "profile graphs".

After each isothermal etching the profile graphs obtained with the various specimens of a given orientation behave similarily whereas distinctive dissolution shapes are produced on differently oriented specimens showing that certainly the dissolution of quartz in concentrated $NH_4F.HF$ solutions is solely orientation dependent as suggested earlier by several workers **(7-13, 21-23)**. Figures 3, 4 and 5 illustrate the typical changes in the dissolution shape of X-, Y- and Z-cut plates with starting circular shape at three different stages of etching. It should be pointed out that the profile graphs shown in these figures are obtained with a numerical filtering and thus do not include the roughness of the true profile graphs of circular sections. The contour of Z-cut plates does not contain surface elements which dissolve very rapidly as observed for example for the contour of X- and Y-cut plates. Then the typical dissolution shape of Z-cut sections is long to be formed and with this procedure we avoid some difficulties which can arise in determining the exact dissolution profile of the Z section in the first stage of etching.

(i) The observed dissolution shapes exhibit the expected symmetries ; this establishes without ambiguity the important role played by the surface orientation in the dissolution process.

(ii) The profile graphs present successive local minima, m, and maxima, M. The minima seem to correspond to specific angles Ψ_{mi} whose values remain unchanged with subsequent etching. Perhaps the angles related to the minima are correlated with the formation of facets of fixed orientation.

3.2/ Theoretical derivation of the dissolution shapes

The derivation of successive dissolution shapes for X, Y and Z circular sections from the corresponding polar diagram of the dissolution slowness can be accomplished by a numerical simulation (12, 13, 17, 18). It is sufficient to note that the dissolution vector, $\vec{L}$, at a point M of the starting circular section is associated with a planar surface element whose projection in the plane of the polar diagram is now identified with the tangent to the circular section at the point M (Figure 6). The numerical simulation calculates the propagation vector $\vec{P}$ (17, 18, 20) and after suitable tests to distinguish converging and diverging trajectories generates the successive dissolution shapes of the starting circular section of given orientation (φ_o, θ_o). The program offers also the possibility to calculate the least square circle and then to display the magnified out-of-roundness profile graph.

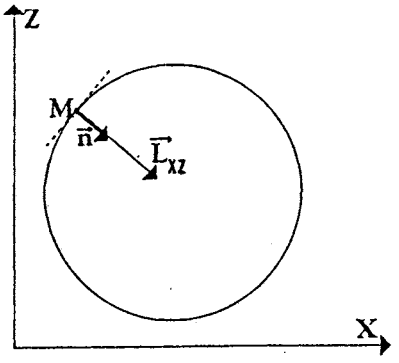


Figure 6 : The slowness vector $\vec{L}_{XZ}$ associated with starting circular section : case of the Y plate.

The theoretical dissolution shapes of X, Y and Z sections are reproduced in Figure 7 at different etching times. From the dissolution shapes obtained we can predict that

(i) On deep etching plane faces are formed quite early on X and Y sections. Faces normal to the Z axis widen rapidly on both X and Y sections inducing the formation of sharp crystal edges.

(ii) Curved faces symmetric about the X axis persist around this crystallographic axis on the etched Y sections.

(iii) The curvature is very slowly modified on the Z section and with prolonged etching we observe a late formation of weak edges about the X_1, X_2 and X_3 crystallographic axes.

In addition the corresponding magnified profile graphs are displayed in Figure 8. A careful examination of these profile graphs reveals that the minima of the out of roundness are directly connected to the formation of plane faces on the etched section.

4 - PROFILE GRAPHS OF SINGLY ROTATED QUARTZ PLATES

To furnish some complementary informations on the orientation dependent dissolution process of quartz crystals some final profile graphs of various singly rotated quartz plates (φ = 0°, θ_o) with a starting circular section are displayed in Figures 9, 10 and 11. All the quartz plates are deeply etched and the depth of etch reaches at least 15 µm. We observe that the values of the angles Ψ_{mi} corresponding to the different minima that exhibit a profile graph vary rather rapidly with the orientation θ_o of the starting circular section.

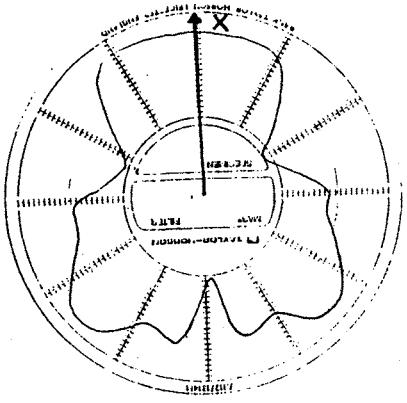


Figure 9 : Magnification (x 200) of the final profile graph of a BT-49 (φ_o = 0, θ_o =-49°) plate.

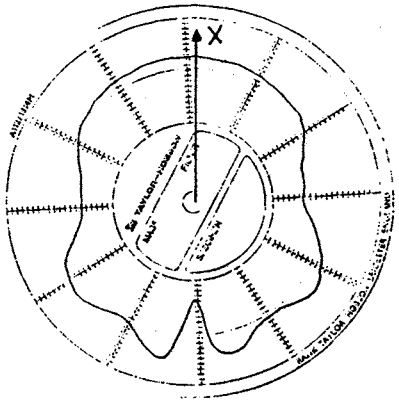


Figure 10 : Magnification (x 500) of the final profile graph of a BT-24 (φ_o = 0, θ_o = - 24°) plate.

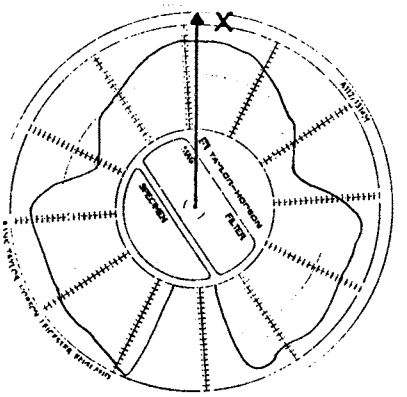


Figure 11 : Magnification (x 500) of the final profile graph of a AT ($\Psi_o = 0, \theta_o = 35°$) plate.

To predict correctly the theoretical shape of a profile graph related to a singly rotated quartz plate the two-dimensional theory of dissolution must be extended effectively into three dimensions. In particular we need to know the exact equation $L_{XYZ}(\Psi, \theta)$ of the dissolution slowness surface. Effectively, each point M lying on the starting circular section is now located by means of the angle Ψ (Figure 12) that the vector $\vec{OM}$ makes with the -Z' axis and the analysis requires to associate at this point the orientations (Ψ, θ) which correspond to the plane tangent to the circular section at this point M. A tentative to determine the constants appearing in the equation $L_{XYZ}(\Psi, \theta)$ from the experimental profile graphs related to differently oriented quartz plates has been made. This procedure is rather difficult since only the dissolution slowness corresponding to crystal orientations present in the dissolution shapes can be evaluated. As a consequence in absence of a great number of reliable data the dissolution slowness representative surface is not completely satisfactorily reproduced. A result is given here (Figure 13) in the case of a singly rotated quartz plate with an angle of cut θ_o of - 49°. The theoretical values of the angles Ψ_{mi}^o corresponding to the minima shown in the profile graph of Figure 13 are indicated in table 1. For a comparison experimental values of Ψ_{mi} are also reported in this table. We observe a departure of about ten degrees between the theoretical and the experimental values. As indicated this departure may be attributed to a lack of reliable data collected from the etched contour of various quartz plates.

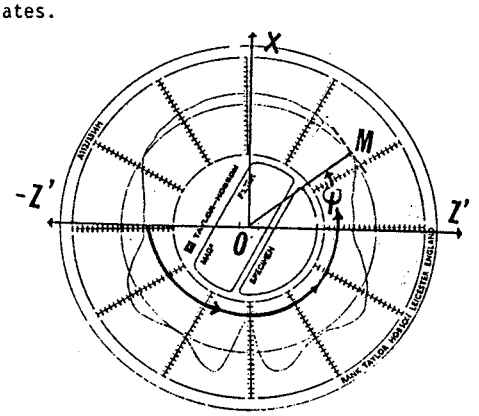


Figure 12 : Definition of the angle Ψ , for singly rotated quartz plates.

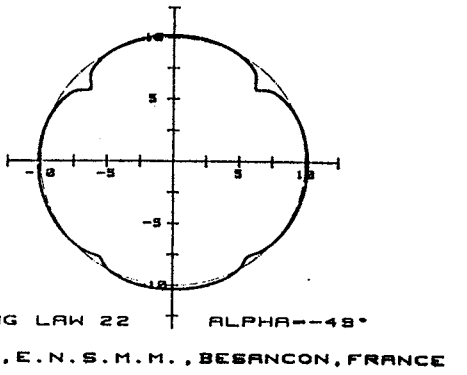


Figure 13 : Magnification of the theoretical profile graph of the BT-49 ($\theta_o = -49°$) plate.

i	Experimental value	Theoretical value
1	116°	128°
2	210°	222°

Table 1 : Comparison of experimental and theoretical values of the angle Ψ_{mi} for an orientation $\theta_o = -49°$.

5 - DISCUSSION AND CONCLUSION

Experimental investigations in the changes in shape, with repeated etchings, of the starting circular section of X-, Y- and Z-cut plates have given evidence for marked orientation effects. In particular the etched contours of X- and Z-cut plates satisfy the crystal symmetry. These observations are in accord with the fact that on the deeply etched surface of X and Z plates develop dissolution figures (Figures 14 and 15) which clearly reflect the twofold and threefold symmetry of quartz crystal. Thus the appearance of both dissolution figures and dissolution profile graphs can be explained in terms of the variation of the dissolution slowness with orientation encouraging us to correlate the experimental observations with the theoretical results derived from the tensorial analysis of the dissolution. Effectively after determining from data the values of the constants appearing in the equation for the dissolution slowness surface it is possible to construct from this equation the dissolution profile graphs corresponding to etched circular X, Y and Z quartz plates and it may be seen that there is a satisfactory agreement between theory and experiments.

To show the correspondence between the development of plane faces during the dissolution and the existence of some local minima of the dissolution slowness, the plots of $L_{XY}(\Psi)$, $L_{XZ}(\theta)$ and $L_{YZ}(\theta)$ vs respectively Ψ and θ are represented in Figures 16 to 18. In addition values of the angle Ψ or θ corresponding to successive minima of the dissolution slowness are collected in table 2 together with values Ψ_{ex} and θ_{ex} related to the bounding faces revealed in the profile graphs of figures 4, 5 and 6.

Z plate		X plate		Y plate	
Ψ	Ψ_{ex}	θ	θ_{ex}	θ	θ_{ex}
30°	≃ 30°	≃ 25° ≃ 41°	≃ 25°	≃ 27° ≃ 42°	≃ 45°
150°	≃ 147°	90° 132°	90°	≃ 70° 90°	≃ 94°
270°	≃ 270°	158°		≃ 123° 150°	≃ 133°

Table 2 : Theoretical and experimental values of orientations correspond to a local minimum in the dissolution slowness.

As expected the dissolution shape tends predominentaly to be formed by faces lying at minima in the dissolution slowness versus orientation plot. Thus as suggested earlier **(24, 25)** on a convex section or surface a plane at a relative minimum in dissolution slowness will limit.

Moreover applying the numerical treatment to singly rotated quartz plates shows that some experimental observations can be partly explained even if as the dissolution slowness varies rapidly with the angles Ψ and θ, the experimental determination of all constants L_o, A_j, B_k and C_ℓ requires a great care because some small departures from the exact values will certainly induce marked deviations in angles Ψ_{mi}.

In conlusion, the above results indicate a clear correspondence between dissolution shapes and orientation and confirm that the general appearence of an etched crystal can certainly be predicted very satisfactory using a three-dimensional analysis of the dissolution process.

REFERENCES

(1) J.R. VIG, J.W. LEBUS and R. FILLER, Report ECOM-4548, 1977 (U.S. Army Electronics Command, Fort Monmouth, N.J.).

(2) J.R. VIG, R.J. BRANSMAYR and R. FILLER, Report DELET-TR 80-5, Fort Monmouth, U.S. Army Electronics Command, 1980, pp 1-8.

(3) R.J. BRANDMAYR and J.R. VIG, Report DELET-TR 81-16, Fort Monmouth, U.S. Army Electronics Command, 1981, pp 1-14.

(4) D. ANG, 32nd Annual Symposium on Frequency Control, Fort Monmouth, N.J., 1978, Electronics Industries Association, Washington, DC, 1978, p 282.

(5) J.R. HUNT and R.C. SMYTHE, Proceedings of the 39th Annual Symposium on Frequency Control, Philadelphia, PA, 1985, (I.E.E.E. New York, 1985), p 292.

(6) C.R. TELLIER, Proceedings of the XIth International Congress of Chronometry, Besançon, France, 1984, Société Française des Microtechniques et de Chronométrie, Besançon (1984), p 115.

(7) C.R. TELLIER and C. BURON, Surface Technology, 22, 1984, p 287.

(8) C.R. TELLIER, Surface Technology, 21, 1984, p 83.

(9) C.R. TELLIER, F. JOUFFROY and C. BURON, Material Chemistry and Physics, 14, 1986, p 25.

(10) C.R. TELLIER, Proceedings of the 39th Annual Symposium on Frequency Control, Philadelphia, Pa, 1985, I.E.E.E., New York, N.Y., (1985), p 276.

(11) M. CASTAGLIOLA, C.R. TELLIER and J-L VATERKOWSKI, Journal of Materials Science, 21, (1986), p 3551.

(12) C.R. TELLIER, N. VIALLE and J.L. VATERKOWSKI, Proceedings of the 40th Annual Symposium on Frequency Control, Philadelphia, Pa, 1986, I.E.E.E. New York, N.Y., (1986), p 76.

(13) C.R. TELLIER, N. VIALLE and J-L VATERKOWSKI, First European Time and Frequency Forum, Besançon, France, 1987 (Imprimerie du Conseil Général du Doubs, Besançon, 1987), p 159.

(14) T. DEGUIN-LEBLOIS and R. BOURQUIN, Proceedings of the First European Time and Frequency Forum, Besançon, France, 1987 (Imprimerie du Conseil Général du Doubs, Besançon, 1987) p 184.

(15) T. LEBLOIS, R. BOURQUIN and C.R. TELLIER, Proceedings of the 2nd European Time and Frequency Forum, Neuchâtel, Switzerland, 1988 (Fondation Suisse pour la recherche en Microtechniques, Neuchâtel), 1988, p 893.

(16) C.R. TELLIER, T. LEBLOIS and P. MAITRE, Journal of Materials Science (to be published).

(17) C.R. TELLIER and J-L VATERKOWSKI, Journal of Materials Science (to be published).

(18) C.R. TELLIER, T. LEBLOIS, 3rd European Time and Frequency Forum, Besançon, France, mars 1989.

(19) F.C. FRANK in R.H. DOREMUS, B.W. ROBERTS and D. TURNBULL (eds.), "Growth and Perfection of Crystals", John Wiley, New York, (1958), pp 411-419.

(20) C.R. TELLIER, N. VIALLE and J.L. VATERKOWSKI, Surface and Coatings Technology, 34 (1988), p 417.

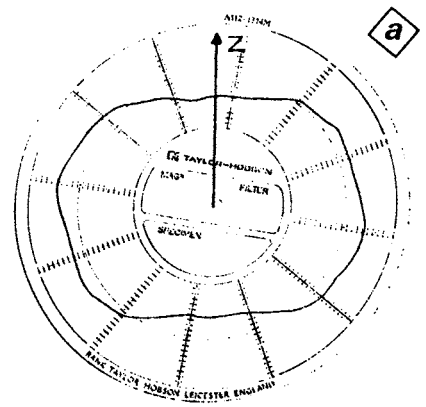
(21) F.M. ERNSBERGER, Journal of Physic and Chemistry of Solids, 13, 1960, pp 347-351.

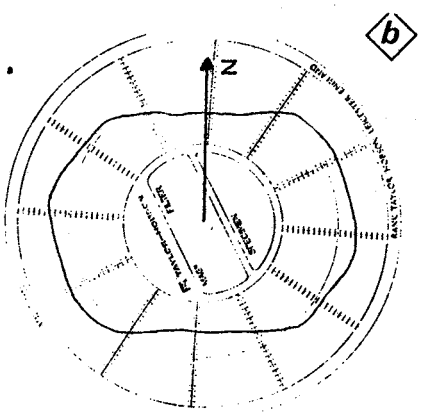
(22) C.R. TELLIER and F. JOUFFROY, Journal of Materials Science, 18, (1983), p 3621.

(23) H.W. WEGNER and J.M CHRISTIE, Physics and Chemistry of Minerals, 9, 1983, p 67.

(24) B.A. IRVING in P.J. HOLMES (eds), "The Electrochemistry of Semiconductors", London, Academic Press, 1960, p 256-289.

(25) K. SANGWAL, Etching of Crystals, North Holland, Amsterdam, 1987.





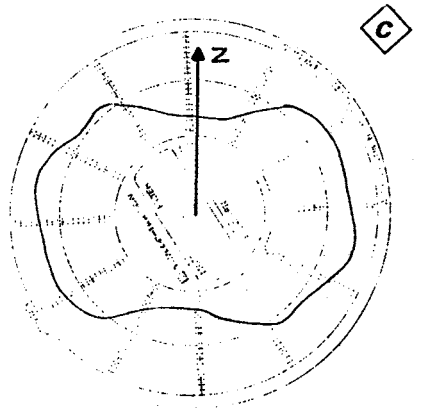
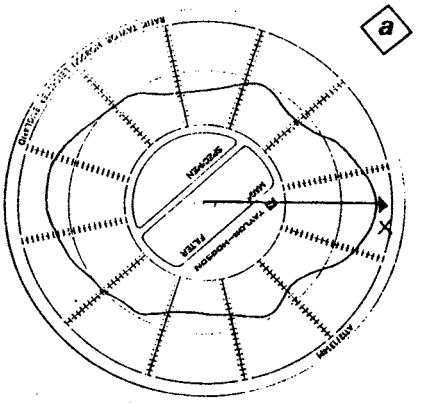
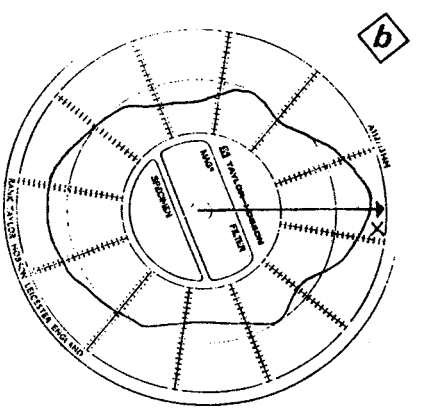


Figure 3 : Magnifications of the experimental profile graphs of an X section with a starting circular shape at three different stages of etching.
a) Magnification is 1000
b) Magnification is 500
c) Magnification is 500





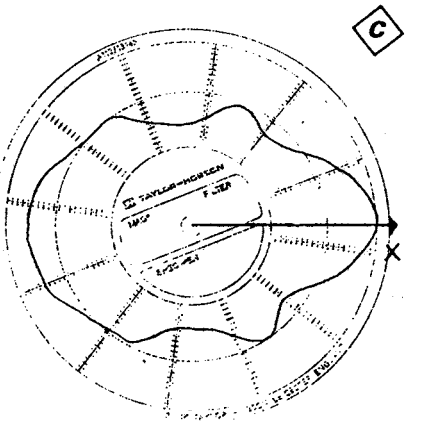
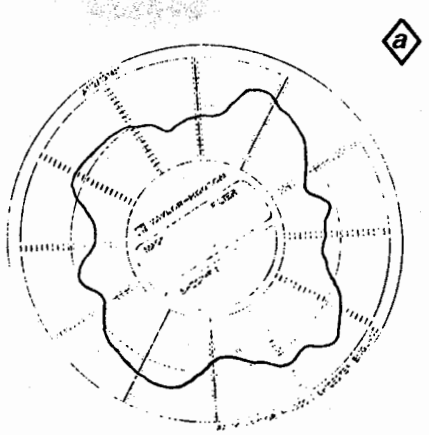


Figure 4 : Magnifications of the experimental profile graphs of an Y section with a starting circular shape at three different stages of etching.
a) Magnification is 1000
b) Magnification is 500
c) Magnification is 500



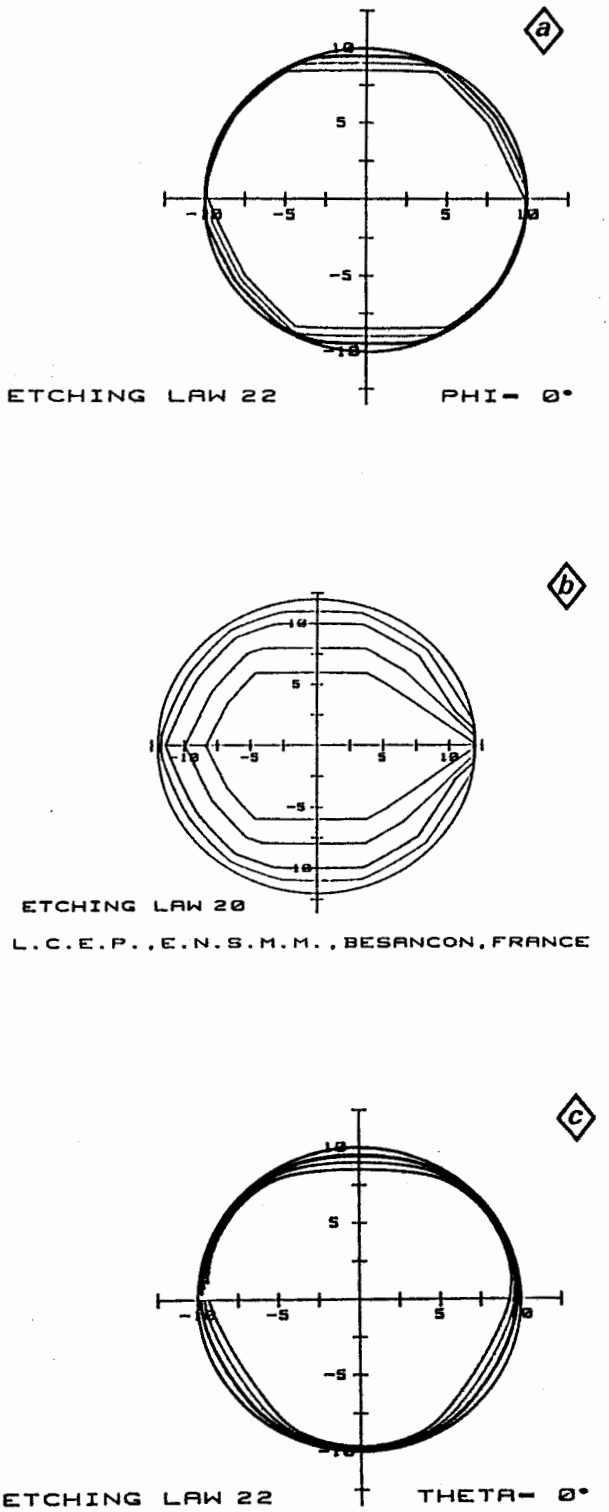
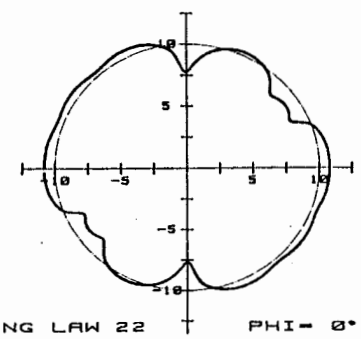
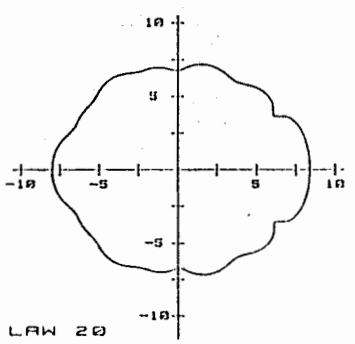


Figure 5 : Magnifications of the experimental profile graphs of a Z section with a starting circular shape at three different stages of etching.
a) Magnification is 5000
b) Magnification is 2000
c) Magnification is 2000

Figure 7 : Theoretical evolution of the shape of quartz sections.
a) X section
b) Y section
c) Z section.





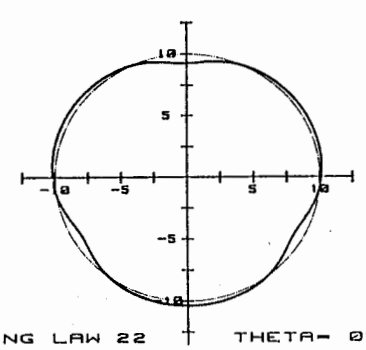
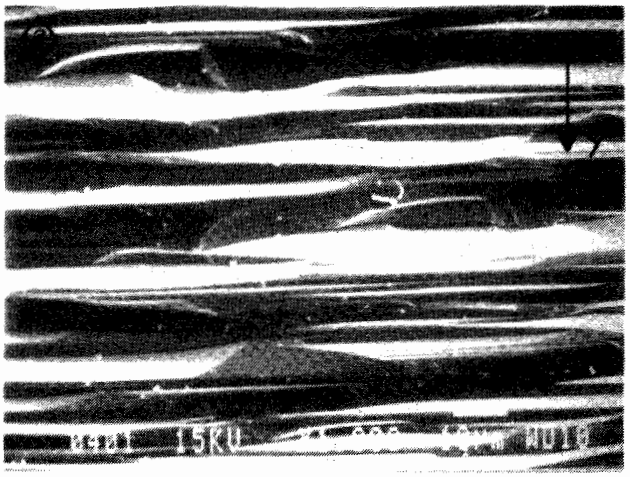


Figure 8 : Theoretical magnification of profile graphs of etched sections.
a) X section
b) Y section
c) Z section



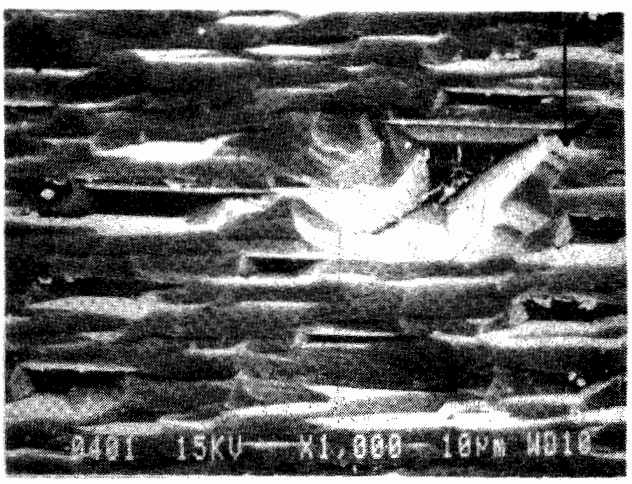


Figure 14 : S.E.M. micrographs of the face of an X plate.
a) top face
b) bottom face.

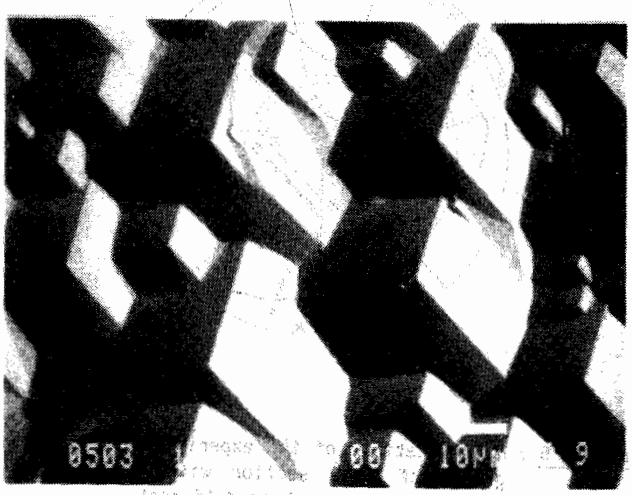


Figure 15 : S.E.M. micrograph of the planar face of a Z plate.

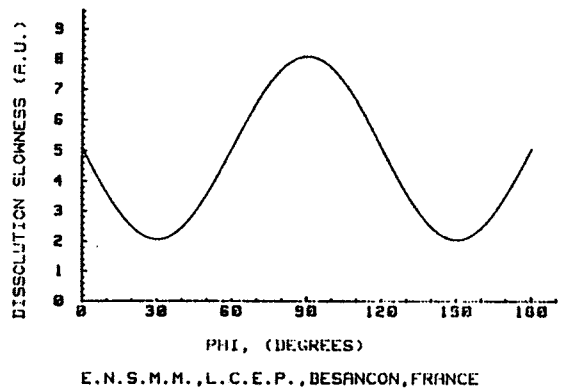


Figure 16 : Theoretical variation of the dissolution slowness L_{XY} against the angle φ, $\theta_o = 0°$.

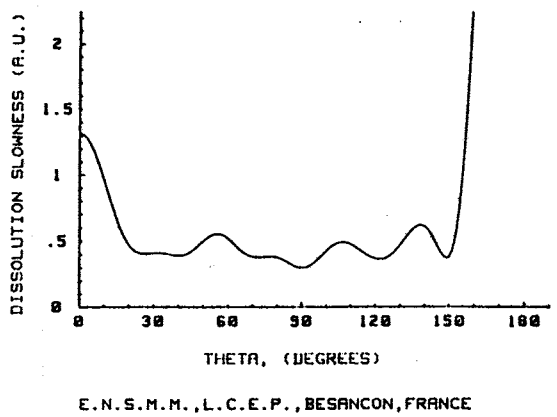


Figure 17 : Theoretical variation of the dissolution slowness L_{XZ} against the angle θ, $\varphi_o = 90°$.

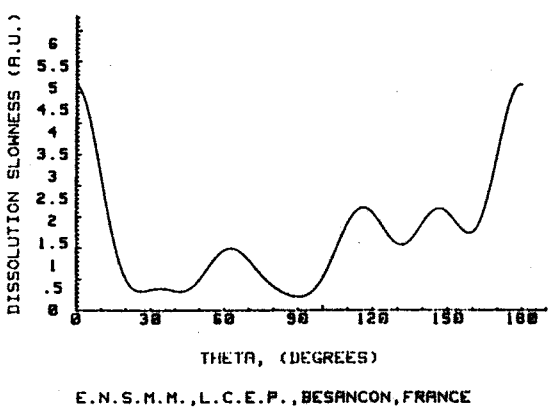


Figure 18 : Theoretical variation of the dissolution slowness L_{YZ} against the angle θ, $\varphi_o = 0°$.

* * * * * *

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the Direction des Recherches Etudes et Techniques which depends on the French Department of Defense.

* * * * * *

TECHNICAL ASPECTS OF $GaPO_4$

G. ENGEL, P. KREMPL, J. STADLER

AVL LIST GmbH, Physics Division, A-8020 GRAZ, AUSTRIA

Abstract

Recently, quartzhomeotypic $GaPO_4$ was proposed as a promising piezoelectric material for high temperature applications, due to its high piezoelectric coefficient d_{11} and the stability of the α-phase up to the α-c (cristobalite-) transition at temperatures > 1200 K. Beyond its attractivity as a transducer material, the question arises, whether it is also useful as a resonator material, and, in particular, whether temperature compensated cuts exist.
In the present paper we show the existence of temperature compensated cuts with high coupling constants. This fact disproves the opinion, that α-β transitions are necessary conditions for temperature compensated cuts in the class of quartzhomeotypic crystals. We further discuss some other technical aspects of $GaPO_4$ concerning its applicability, such as the material's constants, morphology and crystal perfection. We conclude, that $GaPO_4$ is not only a promising material for high temperature sensor applications, but also a candidate for 'temperature compensated' bulk wave resonators.

1. Introduction

Gallium orthophospate $GaPO_4$ belongs to a crystal family with the composition $M^{3+}X^{5+}O_4$ and a structure which can be considered as a derivative of α-quartz (enantiomorhic space groups $P3_121$ and $P3_221$), constructed by the alternate replacement of half of the Si atoms by M^{3+} = Ga and the other half by X^{5+} = P [1]. This results in a doubling of the c lattice parameter as compared with quartz. $GaPO_4$ polymorphs corresponding to the cristobalite phases of SiO_2 also exist [2] (Fig. 1), but contrary to, e.g. the isomorphous crystals $AlPO_4$ and $FePO_4$ no β-phases (enantiomorphic space groups $P6_422$ and $P6_222$) are known. Up till now, gallium orthophosphate was considered as a rather exotic member of the crystal family with quartz-type structures. Therefore, it has mostly been only investigated together with other quartzhomeotypic crystals, e.g. concerning solubility in growth media [3], synthesis [4], crystallography [2,5], photo-electron [6]- and unpolarized phonon [7]-spectra. Beyond a general interest devoted to piezoelectric crystals of said MXO_4 family [8], a special interest in the compounds $GaXO_4$ (X = P,As) emerged as a result of a theoretical modelling, which indicated a very low temperature dependence of physical properties [9]. X-ray refinements up to 750°C confirmed the predicted high structural stability of $GaPO_4$ and have shown that the structural changes with temperature can be

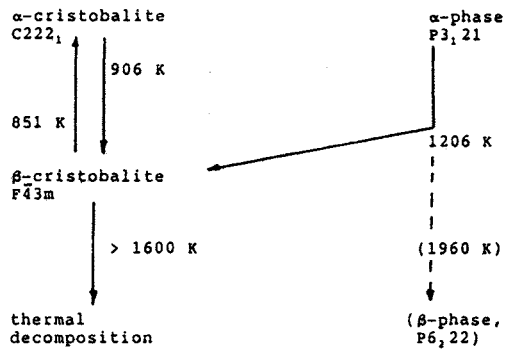


Fig. 1: Phase relations in $GaPO_4$ (see Ref. 2). The transition of the α-phase to the β-cristobalite phase at 1206 K is irreversible. It should be noted, that short-time temperature excursions up to 1400 K are possible for high quality single α-phase crystals (this matches the statements in Ref. 18). Incipient α-β phase transition dynamics are present in the α-phase with an estimated stability limit of nearly 2000 K [10]. This is indicated by the dashed line. These dynamics are abrupted at the alpha to cristobalite transition.

described as the uncompleted 'back-tilting' of the GaO_4 and PO_4 tetrahedra towards higher ($P6_422$ or $P6_222$) symmetry[1]. Further evidence for the high temperature stability of $GaPO_4$ was presented in a recent study on the class of quartz-homeotypic crystals, where the tetrahedral tilting in selected MXO_4 crystals was discussed and a measurement of the temperature dependence of the piezoelectric coefficient d_{11} in $GaPO_4$ reported[10]. The usefulness of gallium orthophosphate as a piezoelectric material for high temperature sensor applications was concluded. The question now posed is, whether it is also useful as a resonator material.

It is the purpose of the present paper to discuss some aspects related to the applicability of gallium orthophosphate as a piezoelectric resonator material. At first, the material constants are discussed including those constants which are important for the search of temperature compensated cuts, i.e. the temperature coefficient of the elastic constant c_{66}. It turns out, that the analogue of the AT-cut in quartz and berlinite must exist in gallium orthophosphate also, since Tc_{66} is positive. Next, an example for a resonator in the vicinity of the compensated cut is given.

It is worth to consider crystal growth and perfection as further aspects, for which the recent literature is shortly reviewed in the sections 4 and 5.

We conclude (section 6) that the combination of the high temperature capability and the existence of compensated cuts offers attractive perspectives for resonator applications of $GaPO_4$.

2. Material constants of $GaPO_4$

$GaPO_4$ has six independent elastic constants (c_{11}, c_{33}, c_{44}, c_{12}, c_{13}, and c_{14}, or, equivalently, the corresponding set of $s_{\mu\nu}$; here and in the following it is understood, that $c_{\mu\nu}$ and $s_{\mu\nu}$ are for constant field). The elastic constants were determined from the measurement of the resonance frequencies of piezoelectrically excited vibrations of thin crystal plates and bars. The principles of the method are outlined e.g. in the Refs. 11 - 14, and the experimental setup was essentially the same as the one used in Ref. 15. X-cuts, rotated X-cuts, and a Y-cut were used as orientations. The size, morphology, quality, and the restricted number of crystals suitable for measurement influenced the choice of the orientations and the size of the samples (maximum dimension was only 7 mm). A redundant set of frequency values resulting from contour and thickness modes was measured, and for each sample an analysis of the secular equations was made. The elastic constants were derived from relating the eigenvalues of the secular equations to the measured frequencies. The results are listed in table 1, where, for comparision, the corresponding values of quartz and berlinite are also included. At ambient temperatures, the elastic behaviour of gallium orthophosphate, quartz and berlinite is very similar. However, due to the higher density of $GaPO_4$, the sound velocities ($v = \sqrt{c_{eff}/\rho}$) and the frequency constants ($N = f.d$) are markedly lower. According to the data in table 1, the rotated Y-cuts with zero c'_{56}, or, equivalently, with a relative extremum of the frequency constant are located at $\vartheta_1 = -37^\circ$ and $\vartheta_2 = 53^\circ$. These cuts correspond to the AC and BC quartz cuts, for which they are located at -32° and 59°, respectively. Further rotated Y-cuts are the temperature compensated (TC) cuts, which are given by a zero first order temperature coefficient of the resonance frequency. For any crystal of class 321 this coefficient is given in its simplest form by the expression[16]

Table 1: Elastic constants $c_{\mu\nu}$ (in GPa) and $s_{\mu\nu}$ (in TPa^{-1}) of $GaPO_4$. For comparision, the corresponding values of $AlPO_4$[32] and of quartz[33] are also given.

	$GaPO_4$ [a]	Quartz	$AlPO_4$
s_{11}	16.8	12.8	17.3
s_{33}	17.5	9.75	12.0
s_{44}	31.3	20.0	26.5
s_{66}	40.8	28.2	40.2
s_{13}	-1.5	-1.3	-1.6
s_{14}	8.7	4.5	5.8
c_{11}	70.7	86.6	64.0
c_{33}	58.3	106.1	85.5
c_{44}	41.9	57.8	43.2
c_{66}	32.1	39.95	28.4
c_{13}	6.6	12.6	9.6
c_{14}	-17.8	-17.8	-12.4

[a] Relative errors were estimated to 3-5% for main axes components, 10% for the others.

$$Tf = \frac{1}{f}\frac{\partial f}{\partial T} = \frac{1}{2}\alpha_v - \alpha' + \frac{1}{2c'_{66}}\,\partial c'_{66}/\partial T, \qquad (1)$$

where α_v denotes the volume expansion coefficient, α' the linear expansion coefficient, and

$$c'_{66} = c_{66}\cos^2\vartheta + c_{44}\sin^2\vartheta - 2c_{14}\sin\vartheta\cos\vartheta$$

the effective thickness shear constant. According to (1), the vanishing of the resonance frequency arises from the cancellation of the effect of thermal expansion against the temperature coefficient of the elastic constants. For most crystals the latter effect is negative and is considerably larger in magnitude than the former. If, however, $Tc_{66}>0$, which means a stiffening with increasing temperature, then angles ϑ exist, where Tc'_{66} is zero. Near these zeros, Tc'_{66} is sufficiently small for the cancellation. In table 2, selected measured values of $Tc_{\mu\nu}$ are listed for $GaPO_4$ and are compared with the corresponding values of quartz and berlinite. It can be seen that $Tc_{66}>0$ for $GaPO_4$ also. Thus $GaPO_4$ is a new member in the rather exclusive family of TC materials.

Table 2: Selected temperature coefficients $Tc_{\mu\nu}$ (in 10^{-4} K^{-1}) of $GaPO_4$ and their comparision with the corresponding values of quartz and $AlPO_4$.

	$GaPO_4$	Quartz[a]	$AlPO_4$[b]
Tc_{11}	-0.37±0.01	-0.47	-0.76
Tc_{33}	-1.18±0.1	-1.6	-2.18
Tc_{66}	+0.31±0.1	+1.7	+1.04

[a] Ref. 32
[b] Ref. 33

3. Search for the Temperature Compensated (TC) Cut of $GaPO_4$

According to the IEEE Standard 176[14], the TC cut vibrating in the slow shear mode in quartz is given by *negative* angles ϑ (the nominal value is $\vartheta = -35\frac{1}{4}°$). The same convention is proposed for gallium orthophosphate, and has already been used above. Any cut angle ϑ can be approximated by the Miller - Bravais indices (h01) of a crystal plane having nearly the same inclination relativ to the c-axis ($\vartheta \cong \vartheta_{h1} = -arctan(a.1.cos30^0/c/h)$; a,c are the unit cell parameters). By symmetry, ($\bar{h}$h1) and (0$\bar{h}$1) are equivalent planes yielding the same cut angle, whereas (0h1), ($\bar{h}$01) and (h$\bar{h}$1) would result in positive values of ϑ. Several crystal resonators with the main faces parallel to (h01) with low values of h and 1 (up to 3) were made, and their frequency-temperature behaviour determined. The result for the sample fabricated from a (302) blank is believed to be close to that expected from a temperature compensated cut, and is shown in Fig. 2. The f-T relation is cubic to a good approximation. This suggests, that the normalized frequency offset can be described by the equation

$$\Delta f/f_0 = a_0(T-T_0) + b_0(T-T_0)^2 + c_0(T-T_0)^3$$

similar to quartz. For the curve shown in Fig. 2, $a_0 = -5.21\ K^{-1}$, $b_0 = 0.022\ K^{-2}$, $c_0 = 0.00433\ K^{-3}$, $f_0 = 5.56746$ MHz, $T_0 = 25^0$ for $\Delta f/f_0$ in ppm. From the order of magnitude of these values, the "true" AT-like cut is expected to be located within $\pm 2^0$ of the crystal plane (302).

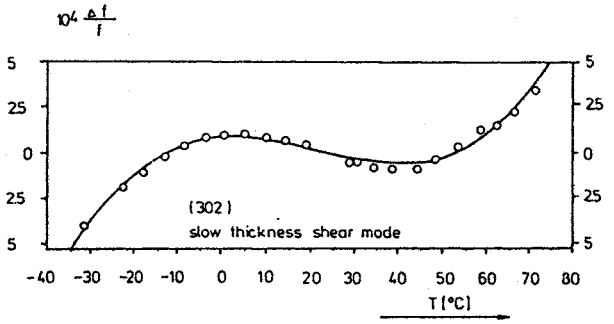


Fig. 2: Frequency-temperature relation in (302) cut $GaPO_4$ grown by the slow heating method. Circles denote measured values. Fitted curve (cubic parabola): full line.

A typical resonance curve of a crystal vibrating in the fundamental thickness shear c-mode is shown in Fig. 3. From the parallel and series resonance frequencies the electromechanical coupling factor is determined to $k = 0.15$, which is larger than the corresponding value for the AT cut of quartz ($k = 0.077$)[16], and of $AlPO_4$ ($k = 0.11$)[17]. The resonance resistance for the sample used for the measurement was in the

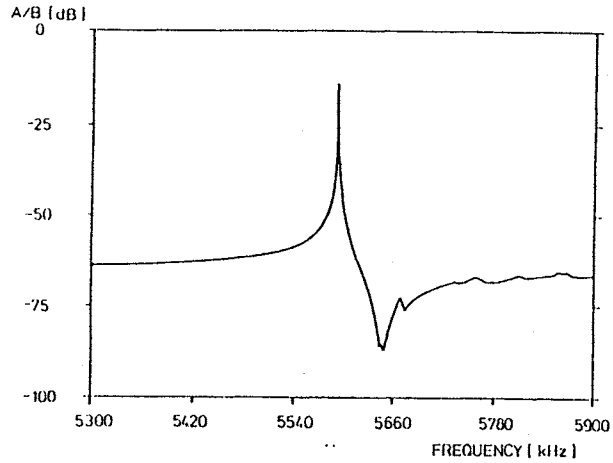


Fig. 3: Piezoelectric resonance peak in a (302) cut $GaPO_4$ resonator demonstrating large electromechanical coupling. Electrodes are evaporated Cr/Au. The frequency constant is N = 1370 KHzmm (uncorrected for electrode mass loading).

range 0.5 to 1 kΩ, depending on temperature. Lower values of R_1 are believed to be achievable by optimized resonator design. It is noted, however, that due to hydroxyl uptake, presumably ≥ 300 ppm, the higher order temperature coefficients are probably not intrinsic values. Furthermore, R_1 might also be at least partly due to anelastic damping (see below). This is consistent with the fact, that anharmonics posed no problems for the evaluation of the results.

4. Crystal Growth

The growth of gallium orthophosphate is usually performed using hydrothermal or related processes with aequous phosphoric acid as the growth solution[18-22]. The solubility of $GaPO_4$ in phosphoric acid has been investigated over a wide range of temperatures and pressures[3,18-20]. The temperature coefficient of the solubility is negative in the range starting from approx. 100°C (the lowest reported growth temperature[21]) up to 350 - 400°C, depending on acid concentration. At higher temperatures, the solubility has a mimimum, and then increases with increasing temperatures[19-20].

Growth temperatures in the realm of the positive temperature coefficient of solubility have the full disadvantage of the high corrosivity of the solvent (the acid solution does not allow the use of the standard steel autoclave technology of the quartz production industry), and as far as we know have not yet been used.

Growth temperatures in the realm of retrograde solubility are therefore preferred, but the problem of corrosive attack is still considerable. It can be solved either by using noble metal liners or quartz glass capsules[18,23], or, at the lower temperatures, Teflon. As an example, Litvin et al[19] recommend 280°C growth zone temperature, 50°C gradient, 7% baffle opening, 5 - 5.9 m H_3PO_4, and a Teflon liner for seeded growth of $GaPO_4$. At temperatures below approx. 200°C, a greater variety of growth methods can be applied, which below 160°C can even be operated at atmospheric pressures[21,24]. Ta and Ag become possible materials for the growth vessels. The vertical gradient autoclave method[21], the slow heating method[21,22,24], the horizontal gradient autoclave method[22], and a method using two vessels heated by infrared radiation[21,24] have been used for growing $GaPO_4$ crystals in this temperature interval.

5. Crystal Perfection

5.1 Twins

One of the problems related to the crystal perfection of gallium orthophosphate is twinning[21]. The types of merohedral twins compatible with the trigonal structure of

Table 3: Effect of merohedral twins in $GaPO_4$.

Type of twin:	Brazil	Dauphine	Combined
Twin element:	m∥{11$\bar{2}$0}	2∥[001]	m∥(001)
Composite symmetry:	$\bar{3}$m1	622	$\bar{6}$2m
Absolute configuration:	reversed	reversed	unchanged
Sign of piezoelectric constants e_{11} and d_{11}:	reversed	reversed	unchanged
Sign of piezoelectric constants e_{14} and d_{14}:	reversed	unchanged	reversed
Frequency constant N:	unchanged	affected	affected
Coupling constant:			
a) Modes driven exclusively by the constant e_{11} or d_{11}:	reduced	reduced	unchanged
b) Other modes:	reduced	reduced	reduced

$GaPO_4$, their twin elements and their effect on selected resonator properties are listed in table 3. Although there are several special cases, where twins have no effect on a specific resonator property, it is very important for a crystal blank to be untwinned. Electrical (Dauphiné) twins are avoided by the use of untwinned seeds, while optical (Brazil) twins are avoided by using high purity growth media and a careful growth process which minimizes the formation of inclusions[21].

5.2 Hydroxyl traces

It is well known for quartz and $AlPO_4$ that the acoustic damping in crystal resonators is related to OH^- defects[26,27]. Correlation curves for quartz were reviewed by, e.g. Brice and Cole[28]. For $AlPO_4$, the relationship between water traces and the "extrinsic" acoustic losses was established on an empirical basis in Refs. 29 - 30.
A low amount of water traces is conjectured to be an important virtue of high quality $GaPO_4$ also.

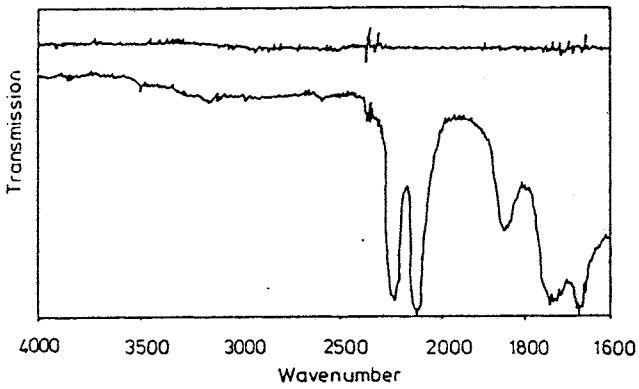


Fig. 4: Infrared spectrum of hydrothermally grown $GaPO_4$. The absorption coefficient at 3400 cm^{-1} is $\alpha \approx 0.3\ cm^{-1}$. Using the procedure described in Ref. 31 and assuming only OH contributing to α, $C(H_2O) = 65\alpha \approx$ 20 ppm weight results. Since, however, intrinsic absorption also contributes to α, the actual water content may even be lower.

The presence of OH^- in $GaPO_4$ crystals was revealed by measuring the absorption of IR radiation at wavelengths of ~ 3 μm. Fig. 2 shows the IR transmission spectrum of a sample with a water trace content of order of magnitude ~ 10 ppm by wt. This order of magnitude seems to be reasonably low for most applications of piezoelectric $GaPO_4$ when compared with $AlPO_4$. However, extensive measurements are needed to establish the quantitative requirements on the maximum water trace content of $GaPO_4$ for resonator applications.

6. Discussion

The elastic constant c_{66} of $GaPO_4$ exhibits a positive first order temperature coefficient, which means a stiffening with increasing temperature. As a result of this unusual feature, gallium orthophosphate possesses crystal orientations with temperature independent resonance frequency. Experimentally realized examples of resonators with orientations in the vicinity of the compensated cut and vibrating in the slow shear mode exhibited a coupling constant as large as ≈ 0.15. This means that, in terms of the squared coupling constant or of the capacitance ratio, a value of nearly four times larger than in quartz can be achieved. This suggests that gallium orthophosphate may be a more suitable piezoelectric material for the use in bulk wave resonant devices than quartz. As an example, the larger coupling factor may lead to larger bandwidth in filter applications. The performance of other singly rotated cuts, e.g. the fast shear mode TC cut, and of doubly rotated cuts on the c-mode branch may also benefit from the large electromechanical coupling. The low frequency constants may also be advantageous for some applications. Furthermore, all constants $Tc_{\mu\nu}$ given in table 3 for $GaPO_4$ are lower in magnitude than the corresponding values in quartz and berlinite. This suggests, that the orientation accuracy requirements may be less severe than in quartz, which could be particularly useful for stress compensated cuts. More freedom in resonator design, e.g. for specific locations of inflection and inversion temperatures, and less sensitivity to thermal gradients may also result.
However, for a substantiation of these conclusions, a more complete and accurate determination of all relevant piezoelectric, elastic and dielectric properties, and their temperature dependencies on large high quality crystals is necessary.
The performance at high frequencies and the $Q\text{-}\alpha_{IR}$ relation are further subjects of interest. The achieved low values of water

uptake in gallium orthophosphate crystals may already be sufficient for good damping characteristics of high frequency devices.

The combination of low values of the first order $Tc_{\mu\nu}$ and the high temperature capability suggests, that $GaPO_4$ may be more suitable than $AlPO_4$ for the use in bulk wave devices specified for large temperature intervals or temperatures above, as an example, 400°C. Sensor technology, ultrasonics, and telecommunications are probable fields of applications, provided that large high quality crystals can be grown.

The question, whether the occurence of α-β transitions in the class of quartzhomeotypic crystals can be considered as necessary conditions for the existence of temperature compensated cuts can now be unambigueously answered: it is at least sufficient that the corresponding *dynamics* are present. In a somewhat simplified manner, the α-β dynamics are described by the 'back-tilting' of the MO_4 and XO_4 tetrahedra towards the symmetry of the β-phase. In quartz and $AlPO_4$, the mechanism is completed as soon as the α-β transition is reached. In the case of $GaPO_4$, the back-tilting is abrupted at the irreversible alpha to cristobalite transition (Fig. 1).

7. Conclusion

The results presented in this paper are summarized as follows:

- A complete set of elastic constants is given, showing that $GaPO_4$ exhibits a similar elastic behaviour like quartz.
- The first order temperature coefficient of the elastic constant c_{66} is positive, indicating the existence of temperature compensated cuts.
- Preliminary results obtained from rotated Y-cuts in the vicinity of the temperature compensated cut indicate a cubic f-T relationship similar to the AT-cuts in quartz and berlinite.
- The coupling of these cuts is ≈ 0.15, which is considerably larger than in quartz.
- The Q capability has not yet been fully explored. However, crystals exhibiting a low content of hydroxyl impurities, which are believed to be a limiting factor for the acoustic losses in resonators, have already been grown.
- A literature survey shows that work on crystal growth is in progress.
- Beyond its usefulness as a temperature compensated piezoelectric material, $GaPO_4$ gains considerable attractivity when keeping in mind its potential use as a high temperature material.

Acknowledgement: Part of this work was supported by the Austrian Science Foundation (Project No.4618), which is gratefully acknowledged.

References:

1) O.Baumgartner, A Preisinger, P.W.Krempl and H.Mang, Z.Krist. 168 (1984) 83.
2) K.Kosten and H.Arnold, Z.Krist. 152 (1980) 119.
3) A.Goiffon, J.C.Jumas, R.Astier and E.Phillipot, J.Crystal Growth 71 (1985) 763.
4) E.C.Shafer and R.Roy, J.Amer.Ceram.Soc. 39 (1956) 330.
5) A.Goiffon, G.Bayle, R.Astier, J.C.Jumas, M.Maurin and E.Phillipot, Rev.Chim.Miner. 20 (1983) 336
6) H.Klingelhoefer and G.Lehmann, J.Electron Spectr.& Rel.Phenom. 16 (1979) 259.
7) W.Dultz, M.Quilichini, J.F.Scott and G.Lehmann, Phys.Rev.B11 (1975) 1648.
8) J.R.Sherman Jr., in: Proc. 40th Ann.Frequency Contr.Symposion, 1986, p.91.
9) G.F.Engel and P.W.Krempl, Ferrolelectrics 54 (1984) 9.
10) G.F.Engel and P.W.Krempl Proc 2nd Europ. Frequency & Time Forum (Neuchatel, 1988) 827.
11) R.Bechmann, Proc. Phys. Soc. B (London) 66 (1951) 323.
12) IEEE (formerly IRE) Standard No.176: Proc. IRE 37 (1949) 1378.
13) IEEE (formerly IRE) Standard No.178: Proc. IRE 46 (1958) 764.
14) IEEE Standard on Piezoelectricity, Std-176, Institute of Electrical and Electronic Engeneers, 1978.
15) G.F.Engel and P.W.Krempl, Phys.Lett. 92A (1982) 469.
16) V.E.Bottom, Introduction to Quartz Crystal Unit Design, (Van Nostrand New York, 1982), Ch.8.

17) J.Detaint, E.Philippot, J.C.Jumas, J.Schwartzel, A.Zarka, B.Capelle and J.C.Doukhan, Proc. 39th Ann. Frequ. Contr. Symp. (1985) 234.
18) S.Hirano, K.Miwa and S.Naka, J.Crystal Growth 79 (1986) 215.
19) B.N.Litvin, V.I.Popolitov, M.A.Simonov, O.V.Yakubovich and I.M.Yaroslavski, Sov.Phys.Crystallogr. 32 (1987) 284.
20) A.A.Adkhamov, B.S.Umarov, A.B.Il'yaev, V.I.Popolitov and B.N.Litvin, Dokl. Akad. Nauk. SSSR 292 (1987) 876.
21) G.Engel, H.Klapper, P.Krempl and H.Mang, J. Crystal Growth (1989) in press.
22) B.H.T.Chai, E.Buehler and J.J.Flynn, Europ.Patent Specification No. 57773 (published 1986).
23) U.Krauss and G.Lehmann, Z.Naturforsch. 30a (1975) 28.
24) H.Mang and G.Engel, German Patent Specification No. 3309598 (published 1986).
25) P.W.Krempl, H.Mang, S.Riegebauer and L.Faschingleitner, German Patent Specification No. 3419722 (published 1988).
26) D.M.Dodd and D.B.Frazer, J. Solid State Chem. 17 (1965) 463.
27) R.F.Steinberg, M.K.Roy, A.K.Estes, B.H.T.Chai and R.C.Morris, Proc. IEEE Ultrasonics Conference (Dallas, 1984), Vol.1, p.279.
28) J.C.Brice and A.M.Cole, J Phys.D. 12 (1979) 459.
29) J.C.Jumas, A.Goiffon, Y.Tudic, J.Schwartzel, J.Detaint and E.Philippot, Mat. Res. Bull. 22 (1987) 1063.
30) E.Philippot, A.Goiffon, J.C.Jumas, X.Buisson, R.Arnaud and Y.Toudic, Proc. 2nd Europ. Frequency & Time Forum (Neuchatel, 1988) 843.
31) R.D.Aines and G.R.Rossmann, J. Geophys. Res. 89 (1984) 4059.
32) Z.P.Chang and G.R.Barsch, IEEE Trans. Sonics Ultrason. SU-23 (1976) 127.
33) S.Hearmon, in: Landolt-Boernstein tables, New Series group III Vol. 11, ed. K.H.Hellwege (Springer, Berlin, 1979).

IMPROVEMENT OF BERLINITE'S CRYSTALLINE QUALITY

* E. PHILIPPOT, A. GOIFFON, C. AVINENS
** X. BUISSON, R. ARNAUD, O. BIGNON
*** A. ZARKA, B. CAPELLE, M.T. SEBASTIAN

* URA D407 CNRS, USTL MONTPELLIER, 34060 MONTPELLIER CEDEX France
**SICN, ANNECY
*** UNIVERSITE DE PARIS VI, CNRS

ABSTRACT

The improvement of the quality of berlinite crystals grown in sulphuric acid solutions was pursued through acting on two parameters :

- growth start up conditions
- growth solution concentration in acid

X-ray topography studies confirm the leading role played by the growth start-up conditions in sulphuric acid solutions. It appears that the residual defects shown by the crystals grown in these solutions can mostly be related to growth start-up troubles ; they are generally linked to the occurence of small berlinite crystallites at the seed grown material interface. These crystallites have been eliminated by changing the preparation of growth solutions.

The second parameter strongly influences the crystals' impurity content, generally but wrongly called "water content".

The works enabled us to obtain, in a reproducible way, berlinite crystals with low impurities content and very good crystalline quality.

INTRODUCTION

Berlinite, $AlPO_4$, is a crystal isostructural to quartz but with superior characteristics. Unfortunately, the efforts made during many years to determine optimal growth conditions for this material did not succeed in producing crystal with a sufficient Q value : this surtension coefficient stayed between 10^5 and 10^6.

Recently, we have shown on an experimental base that this Q value could be better than 10^7 when operating in phosphoric acid at T>250°C instead of 180°C (1). But phosphoric acid shows inconveniences difficult to overcome on an industrial scale and we have sought for other growth conditions, chiefly another solvents ; last year, we thus have presented our first encouraging results with sulphuric acid H_2SO_4 (2).

Since then, we have looked for optimal growth conditions with this acid, particularly growth start-up conditions and acid concentration. The results obtained at the "Laboratory of physicochemistry of solid materials UA 407" of CNRS in MONTPELLIER have been transfered to SICN in ANNECY for pre-industrial scale-up of berlinite growth.

We'll concentrate particularly on detailing the importance of growth startup from the seed and acid concentration in the growing solution.

A) IMPORTANCE OF INITIAL CONDITIONS

Experience shows that the crystal's quality is strongly influenced by growth start-up, especially in sulphuric acid and whatever the concentration. The first trials gave crystals with numerous twinings and cracks particularly in the -X zone of crystals grown on X seeds (photo 1).

Examination with optical microscope and polarized light enabled us to detect small crystallites at the seed-first growth interface (Photo 2).

If the crystallites are numerous, they are responsible for the cracks which are originating from crystallites clusters, especially in -X zone. Otherwise they can cause twinings (Photo 3) and numerous dislocations.

X-ray transmission topography permits a more accurate identification of crystals defects (3). First berlinite plate shown in photo 4 is from a sample grown at 200°C in sulphuric acid.

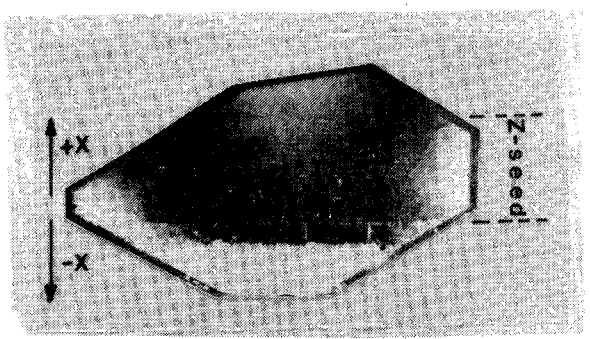


Photo 1 : Twinings and cracks in the -X zone of a Z plate after two bad growth start-up.

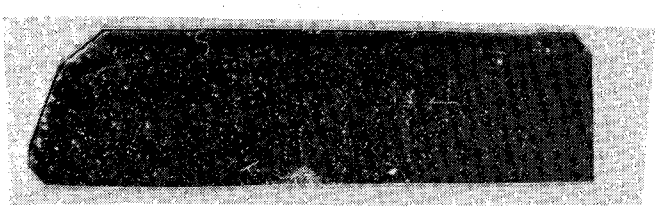


Photo 2 : Crystallites on seed-growth interface.

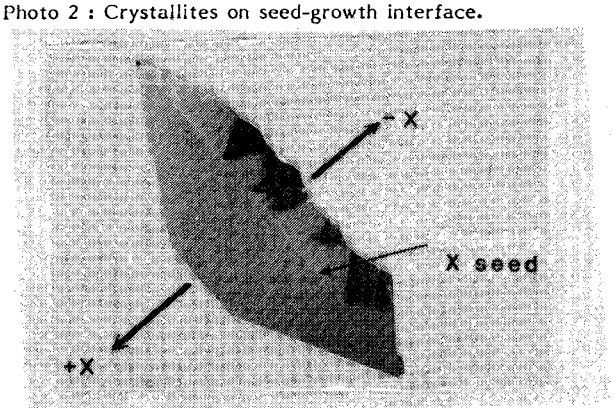


Photo 3 : Twinings in -X zone seen with an optical microscope and polarized light

This plate is among the first sample grown in H_2SO_4 and its quality is not excellent. Nevertheless, many black contrasts are present and they correspond to inclusions, visible in optical microscope, or to stongly strained zones.

With similar growth conditions, a second Y-cut plate shows the same contrasts as in photo 5.

This sample has been also studied in function of temperature. It has been found that the contrasts obtained on X-ray topographs change for different temperatures and this phenomena is directly related to the presence of "water" in the crystal. This water content can be evaluated by means of infrared spectrometry, using a formula derived from the one used for quartz (4).

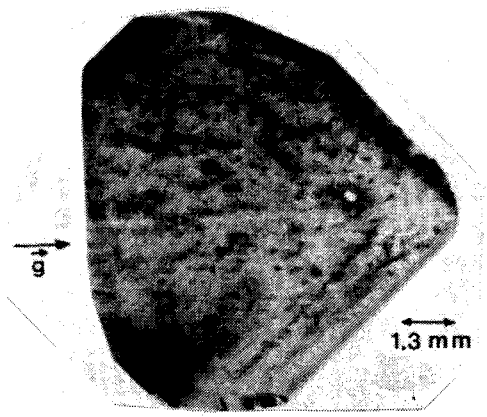


Photo 4 : X-ray topography of a Y-cut plate parallel to (100) plane. 110 reflection.

Photo 5 : X-ray topography of a Y-cut plate.110 reflection. Strong contrasts related to "water" content.

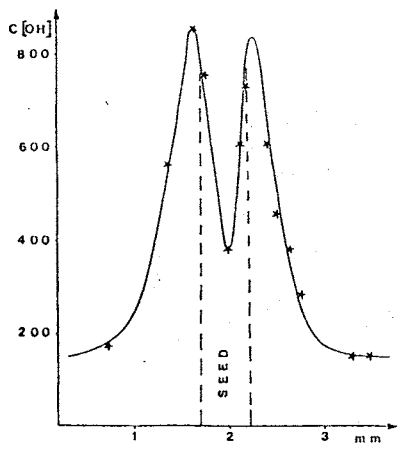


Fig. 1 : "Water" concentration evolution during growth with bad start-up conditions.

Figure 1 shows an example of this "water" concentration's evolution during growth, it has been established after examination of slices cut parallel to seed. We can see that "water" content and crystalline defects vary similarly during all the growth, especially at start-up.

In order to get rid of the defects we must suppress their cause : the presence of crystallites which originate from momentany oversaturation in the solution. We must recall that the solutions commonly used are pre-saturated at temperature which is close to growth temperature i.e. 245°C for instance if growth temperature is 240°C. Since dissolution kinetic of berlinite in sulphuric acid becomes appreciable above 180°C and increases with temperature, there is an important dissolution of nutrient during the time necessary to bring the autoclave to growth temperature. The solution is then over-saturated and micro-crystallites precipitate on the seed at the begining of growth.

In order to check this theory, we have used solutions pre-saturated at temperatures Te more and more elevated with respect to the growth temperature Tc which was held constant. Thus, for a 6M concentration and Tc = 240°C, we raised progressively the value of Te-Tc and we observed, with optical microscope, that the number of crystallites at the interface diminished progressively and disappeared for Te-Tc = 25°C.

Broadly speaking, this value increases with the growth temperature and when the solution is more and more diluted. It means that the dissolution's kinetic increases when the global solubility diminished (Figure 2).

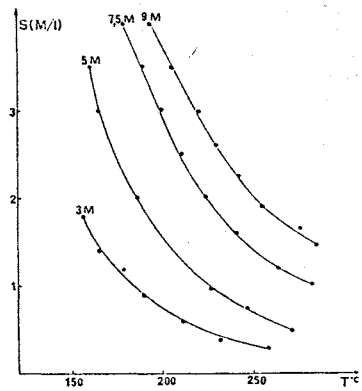


Fig. 2 : Berlinite solubility in sulphuric acid for different acid concentrations.

This is in no way a drawback, on the contrary, it makes more simple growing crystals in dilute solutions : it is possible to use acidic solutions without pre-saturation in $AlPO_4$ when sulphuric acid concentration is below 3M and Tc near 240°C.

The absence of crystallites results in a drastic reduction of the number of defects which can be seen by means of X-ray topography (Photo 6).

Photo 6 : X-ray topography of a Y-cut containing the seed. Better growth conditions but many dislocations from the seed.

In contrary to the previous cases, this crystal presents a relatively low density of inclusions, but it's sufficient to create dislocations initiated from the seed surfaces.

There is another element of tremendous importance concerning start-up conditions of growth, the choice of seed. In order to eliminate dislocations in the seeds, the technique used for quartz crystals has been applied (5). It is based on the selection of seeds in different growth zones. The different results obtained have shown that the quality of crystal is improved especially in the X zone as shown by the sample presented in photo 7. We can see that the +X zone is free of dislocations although the zone around the seed presents many inclusions. The -X zone is not presented here.

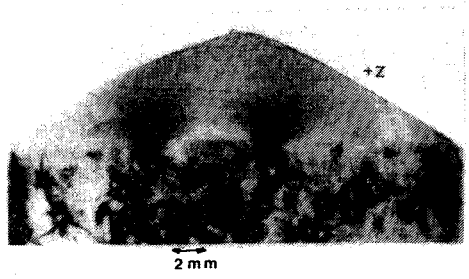


Photo 7 : X-ray topography of a Z-cut (001) from a X plate grown in 6M H_2SO_4 at 220°C.

Photo 8 concerns a Y-cut plate from a sample similar to the one in the previous case where the seed was oriented along X. In this case, growth sectors can be seen and they correspond to the rhomboedrical faces (101) and (102), as well as (001) and (110) faces. The quality of the sample is quite good except in -X zone where dislocations were generated due to the poor growth conditions.

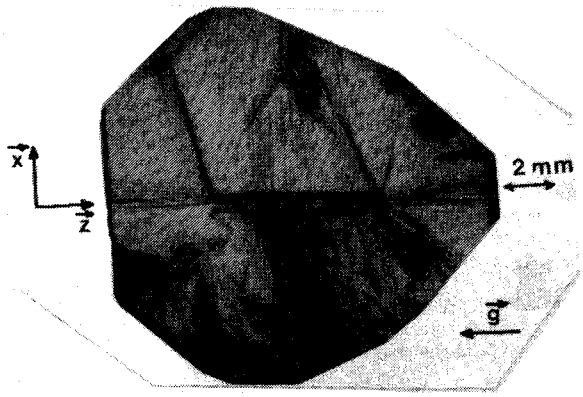


Photo 8 : X-ray topography of a Y-cut plate from a X seed with better growth start-up conditions.

If we associate good growth start-up conditions with good seeds, we can obtain crystals with a crystalline quality comparable with the one of good synthetic quartz crystals (photo 9).

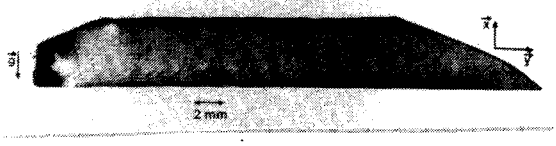


Photo 9 : X-ray topography of a crystal grown from a Z zone seed in 6M H_2SO_4 at 240°C.

In this case, only the +X zone is presented. This sample contains few dislocations and some growth bands parallel to growth front of +X zone.

This improvement on crystalline quality is accompanied by a drastic reduction in "water" content as early as growth start-up, even when using "wet" seeds as shows figure 3.

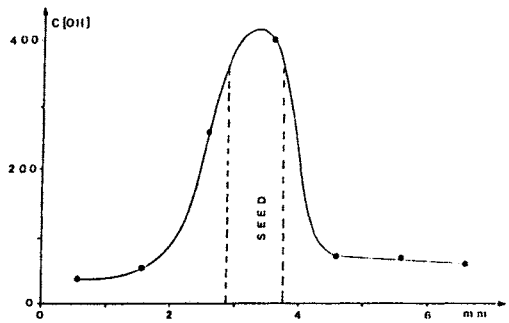


Fig. 3 : "Water" concentration in crystals with good start-up conditions and "wet" seed (crystal grown at 240°C in 4,5M H_2SO_4, Te-Tc ≈ 30°C).

Photo 10 shows an example of this kind of crystals.

Photo 10 : Crystal with good start-up conditions (4,5M H_2SO_4, 240°C).

This systematic study of the growth start-up conditions gave us the capability to obtain, in a reproducible manner, berlinite crystals with crystalline quality comparable to the one of good synthetic quartz crystals. At the same time, we have searched for the optimal concentration in acid, leading all other parameters beeing kept constant ; to crystals as "dry" as possible, i.e. showing a good value for the surtension coefficient, Q near 10^6.

B) INFLUENCE OF CONCENTRATION IN H_2SO_4

Our first trials with sulphuric solutions were made with a 6M concentration which is a value close to those commonly used by most authors using other acids (H_3PO_4, HCl) (6). For this concentration value, we have shown that the growth rate (unit : mm/day/face) rises with the crystallisation temperature Tc and the temperature gradient ΔT (7). The infra-red measurements carried on the crystals thus grown show that the impurities (generally called "water") concentration is proportionnal to growth rate and inversely proportionnal to Tc. Our study was limited to 250°C which is the maximum temperature our present autoclaves can reach. Just as for quartz, we know that the Q value is directly related to the impurities content. Since the main impurity is generally called "water", we first attempted to reduce its value by raising the concentration of the acid solution, i.e. 9M. The result was opposite to what we expected : infra-red spectrometry shows that, for similar growth rate and Tc, the impurity content rises strongly (Figure 4).

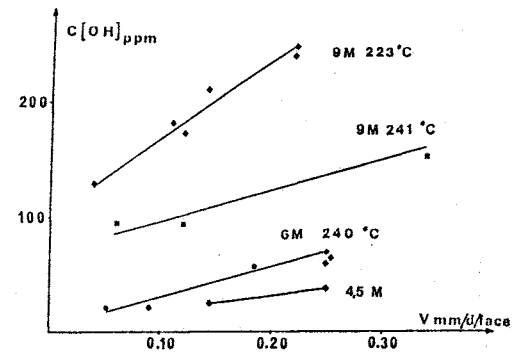


Fig. 4 : "Water" content in terms of sulphuric acid concentration, growth temperature and growth rates.

Hence, the name "water" generally used to give a name to the main impurity is certainly inappropriate ; it could be explained by considering sulphuric acid's dissociations:

$$H_2SO_4 + H_2O \rightleftharpoons H_3O^+ + HSO_4^-$$

$$HSO_4^- + H_2O \rightleftharpoons H_3O^+ + SO_4^=$$

The solvated protons concentration rises with the acid concentration and is certainly the cause for the presence of impurities (O-H bonds) in berlinite crystals.

In order to check this, we have made some growth runs with diluted solutions, first 4,5M. The infra-red spectrometry examination shows, as expected, a strong decrease of the characteristic absorption due to the O-H bonds, this for similar growth rates and growth temperatures (Figures 4 and 5).

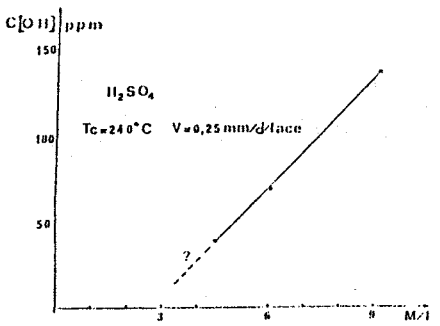


Fig. 5 : "Water" content in terms of sulphuric acid concentration with similar growth rates and growth temperatures.

We thus have started studying more dilute solutions, 3M in particular, the results of which will be soon available. It will help to determine which kind of solution is the most appropriate growing medium for the obtention of berlinite crystals with a surtension coefficient Q as high as possible, for a reasonable growth rate, i.e. 0,2 to 0,3 mm/day/face.

In connection to the lowering of OH concentration, we observe that the infra-red absorption spectra show a better sharpness (Figure 6).

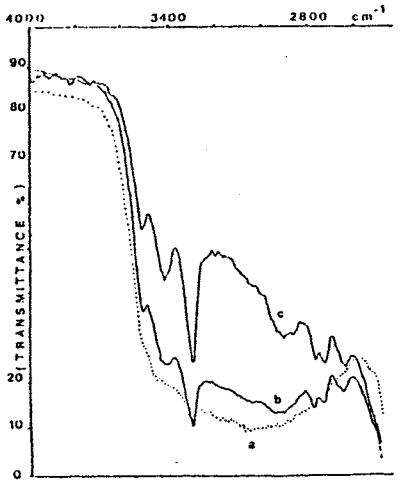


Fig. 6 : Infra-red absorption spectra of :

a) "wet" plate, e = 0,8 mm, C_{OH} > 600 ppm
b) crystal (6M H_2SO_4 Tc = 240°C, V = 0,25 mm/d/face) C_{OH} = 60 ppm
c) crystal (4,5M H_2SO_4 Tc = 240°C, V = 0,15 mm/d/face) C_{OH} = 30 ppm

It will thus be feasible and advisable to assign the different absorption peaks in the range 4000-2500 cm^{-1} and, consequently, modify growth conditions in order to suppress them. In particular, it is now essential to authenticate the peak 3300 cm^{-1} which is used to calculate the OH concentration. It is possible that, similar to quartz for its 3500 cm^{-1} peak (8), the peak at 3300 cm^{-1} should be an intrinsic property of berlinite.

CONCLUSION

This study shows that production of high Q berlinite crystals is no longer an utopia. X-ray topography and infra-red spectrometry enabled us, first to analyse crystalline quality and OH content of the crystals and then to improve progressively upon their quality through seeds selection and improved growth start-up conditions.

In the present state of our works, it appears that OH concentration diminishes when using more and more dilute acidic solutions. The most appropriate sulphuric concentration is to be determined.

The progresses made in laboratory works permitted to develop berlinte's pre-industrialisation in the ANNECY plant of SICN now being capable of synthetizing in a reproducible manner crystals up to some 10 cm in length.

ACKNOWLEDGEMENTS

This study was carried out with the help of CNET, DAII and CNRS.

BIBLIOGRAPHY

1) J.C. JUMAS, A. GOIFFON, Y. TOUDIC, J. SCHWARTZEL, J. DETAINT, E. PHILIPPOT, Mat. Res. Bull., 22, 1063 (1987).

2) E. PHILIPPOT, A. GOIFFON, J.C. JUMAS, X. BUISSON, R. ARNAUD, Y. TOUDIC, 2nd Eur. Freq. and time forum, 843 (1988).

3) A. ZARKA, B. CAPELLE, J. Appl. Cryst., 18, 533 (1985).

4) R.F. STEINBERG, M.K. ROY, A.K. ESTES, B.H.T. CHAI, R.C. MORRIS, Proc. IEEE Ultrasonics Symp., 279 (1984).

5)a) A. ZARKA, LIU LIN, X. BUISSON, J. Cryst. Growth, 54, 394 (1981).
b) A. ZARKA, LIU LIN, X. BUISSON, J. Cryst. Growth, 57, 466 (1982).

6)a) E.D. KOLB, R.A. LAUDISE, Proc. Ann. Freq. Control Symp., 291 (1981).
b) J. YOSHIMURA, S. TAKI, S. MARUYA, Yogyo-kyokai-ski, 89 (8) 433 (1981).
c) J. DETAINT, H. POIGNANT, Y. TOUDIC, Proc. Ann. Freq. Control Symp., 93 (1980).

7) E. PHILIPPOT, A. GOIFFON, J.C. JUMAS, C. AVINENS, J. DETAINT, J. SCHWARTZEL, A. ZARKA, Proc. Ann. Freq. Control Symp., 138(1988).

8) D. CHAKRABORTY, G. LEHMANN, a) J. Solid State Chem., 17, 305 (1976), b) Z. Naturforsch, 33a, 290 (1978).

QUARTZ SWEEPING IN α AND β PHASES

J. P. BACHHEIMER and G. DOLINO
Université Joseph Fourier, Laboratoire de Spectrométrie Physique (associé au CNRS (UA n° 008) - B.P. 87 - 38402 Saint-Martin-d'Hères Cedex - France

Abstract

Sweeping (i.e. electrolysis of quartz at high temperature) is now an industrial process, used to improve the radiation hardness of quartz oscillators. In general the electrolysis is performed for several days at 500°C far away from the α-β transition at 573°C which could induce Dauphiné twinning or fracture. It would however be interesting to operature at higher temperature, either in the α phase nearer to the transition temperature or even in the β phase, if the problems related to the α-β transition can be solved.
We report our first results (S.I.C.N.'s synthetic samples) obtained in the α phase at 550°C an in the β phase at 650°C with sweeping performed in vacuum or air. The electrical current variation during the sweeping was measured as well as during the final cooling giving different activation energies in the α and in the β phases.
The samples were characterized by I.R. measurements at 77K : a complete disparition of the OH lines was observed after sweeping in α phase (650°C - 1500 V/cm) or in β phase (550°C - 3500 V/cm). However all the samples which were heated in β phase were twinned.

I. INTRODUCTION

It is well known (1) that the performance of piezoelectric devices, particularly resonators, are limited by impurities. The major impurities present in synthetic quartz are alkali ions (M^+) associated to aluminium (Al-M^+ center) which exists as a substitution for silicon and hydrogen related as grown OH defects. Sweeping or electrolysis is now a commercial process which is used to dissociate and eliminate impurities by means of their displacements under an applied electric field (1000 to 2500 V/cm along the z direction). Sweeping in an ambiant air atmosphere replaces the alkali ion by an hydrogen to form Al-OH (e_1-e_2 peaks in I.R. spectra) while sweeping in a vacuum atmosphere forms Al-hole center (2). This is explained by the depletion of as grown OH^- which is not compensated by the introduction of external hydrogen (from atmospheric water). Both processes improve radiation hardness but the best is not yet known. Usually sweeping is performed for several days at 500°C below the α-INC-β first order transition (573°C) which could induce Dauphiné twinning or fracture. Due to the thermal activation of the process, it would be interesting to operate at higher temperature, either in the α phase nearer to the transition temperature or even in the β phase, if the problem of electrical twinning induced in cooling during the transition can be solved. Indeed Wooster al al. (3) have shown the possibility of detwinning quartz crystals. Furthermore in β phase one can expect new effects on defect accomodation and diffusion due to the more open structure. Dodd and Fraser (4) found that the narrow hydrogen lines of the IR spectra are eliminated by heating at T > 1000°C and Krefft (5) has demonstrated a strong hydrogen depletion in the anode side region of natural quartz electrolysed at 700°C. Fraser (6) has found a reduction of anelastic losses in β phase, without sweeping, between 573°C and 700°C.
We present the results of our first (sweeping currents and I.R. spectra) investigations in the α-phase at 550°C and in the β phase at 650°C with air and vacuum atmosphere and we compare with 500°C "classical sweeping". Electrical twinning was controled by means of laser light second harmonic generation (7).

II. EXPERIMENTAL

Samples were successively cut from a single cultured (SICN) quartz bar of medium quality. The impurities content was determined by torch-plasma analysis :

Al	Li	Na	K	Ca	Fe	
1,5	0	2,7	1,6	1,9	0,4	ppm

Typical dimensions were ∿ 1//X) x 2(// Y) x (1 // Z) cm^3 for all the samples (pure Z-seed removal) except for the air swept sample at 650° which was ∿ 1 cm(//Y). The pressure of Pt electrodes foils on sample Z faces was about 1,8 N/cm^2 (we have checked in air experiment that multiplying by 10 the pressure has negligeable effect on the intensity of the sweeping current). Electrolysis in air was carried out in an open furnace while the vacuum electrolysis was done in a vacuum furnace operating at 10^{-6} nm Hg. Electrical field was increased slowly, once the sweeping temperature was reached, to limit the current to about 10 $\mu A/cm^2$ (except for 650° in air) accross the sample. A Lambda 9 Perkin Elmer spectrometer was used to get optical transmission spectra (non polarized light).

III. RESULTS :

- Vacuum sweeping : Z faces of the two samples were finely ground.
Fig. 1 shows the sweeping currents in α phase (500°C) (a) and in β phase (650°C) (b).
Fig. 2a shows the as grown spectrum which displays the s_1, s_2, s_3, s_4 peaks of the OH bounds. The absorption coefficient at 3500 cm^{-1} is $\alpha_{3500} = 0{,}060\ cm^{-1}$.
Fig. 2b shows the result of a "classical" 500°C-1500 V/cm sweeping for 18 days : s_1, s_2, s_3, s_4 decreases, e_2 (Al-OH) appears which demonstrate the effectiveness of the sweeping process (M^+ are replaced by H^+).
Fig. 2d shows the spectrum from the second sample swept directly in the β phase at 650°C : all the s peaks and e peaks disappear, furthermore a strong reduction of the broad OH absorption occurs ($\alpha_{3500} = 0{,}024\ cm^{-1}$). However the sample becomes twinned. We have also look for a better sweeping in the α phase by a subsequent electrolysis of the first sample : after the Z faces were optically polished it took about 2 days to reobtain the final current of the fig. 1a, then the temperature was increased to 550°C (1500 V/cm). After 8 days the spectrum was identical to the former one (fig. 2b) but the activation energy became higher (2,0 eV instead of 1,46 eV). The electrolysis was then prolongated for 6 days at 550° with a stronger electric field of 3500 V/cm. We get thus the spectrum of fig. 2c which is practically identical to the spectrum obtained in α phase. Furthermore, no twinning appears as the temperature transition was not reached.

- Air sweeping : The two samples were optically polished.
Fig. 3 shows the sweeping currents in α (500°C) and β (650°C) phases.
Fig. 4 shows the different spectra obtained within

similar conditions as those used in vacuum sweeping. The "classical" process at 500°C - 2000 V/cm gives (fig. 4b) the e_1-e_2 (Al-OH) peaks and a slight increase of the broad OH spectrum. Subsequent sweeping on the same sample at 550°C- 3500 V/cm reduces all the peaks and the integrated spectrum which becomes weaker than the as grown spectrum. In α phase (650°C-1270 V/cm s_3, s_2 nearly vanished, s_4 is completely eliminated while e_2, e_1 decrease.

Table I lists the activation energies E (according to the law $\sigma.T = A \exp(-E/kT)$ measured during the cooling at the end of the different sweepings and the α_{3500} coefficient of absorption.

Vacuum sweeping	E(eV)	α_{3500}	Air sweeping	E(eV)	α_{3500}
as grown		0,060	as grown		0,064
500°-1500 V/cm	1,45	0,060	500-2000	1,87	0,064
+ 550 - 1500	2,00	0,060	+550-3500	1,95	0,026
+ 550 - 3500	1,70	0,020	650-1270		
650 - 1270	2,24 (β)			1,77 (β)	
	1,90 (α)	0,024		1,97 (α)	0,030

IV. DISCUSSION

There is a noticeable difference between the curves (fig. 1 and 3) of swept current as a function of time: It is about ten times higher in air than in vacuum. This fact support recent evidence (8, 9) on the role of hydrogen (profusely supplied by atmospheric water) in the electrical conduction of quartz. In air the two curves at 550 and 650°C are roughly identical after adequate scaling:the 650°C curve is obtained from the 500° curve with a reduction of ∿ 5 on the time scale and a multiplication of ∿ 15 of the currents. The electrodiffusion process is therefore well accelerated as expected. The ratio of the current gives an activation energy of 1,3 eV typical of alkali conduction (10). However the final ratio should be greater to be consistent with an activation energy of ∿ 1,9 eV typical of conduction by proton H^+ (11). In vacuum the curves are much more difficult to interpret. No steady state current was reached but we note a discontinuity after two-days both for 500°C and 650° C sweepings. May be the lack of hydrogen slows down the process which becomes more dependent on the nature of electrode contact and on the existence of space charge. This space charge can vary very much with temperature and electric field (12). However at the end of these different sweepings the activation energy is also typical of hydrogen conduction except in the 500°C- 1500 V/cm process where the smaller value is perhaps due to a part of conduction by remaining alkali ions.
The evolution of infrared spectra may also qualitatively be explained (2) by the role of hydrogen.i) under vacuum there is mainly the as grown H and "classical electrolysis" is not sufficient (fig. 2b) to dissociate and displace all the M^+ and optically active OH. This can be obtain either with the help of high electric field near the transition temperature in α phase (fig. 2c) or in the β phase at middle field by means of thermal activation (fig. 3c) thus we see the complete disparition of the s and e peaks (which transform in Al-hole center). This fact has not been already mentioned to our knowledge for samples of this quality. We find on the whole sample a result found on the anode side by Krefft (5) in β phase.ii) In air in the conditions (500°C-low field) of fig.4b there is enough H available to compensate the loss of alkali ion giving strong e_2-e_1 lines due to Al-OH. When the temperature and the field are increased the rate of supply of hydrogen from the anode is not sufficient and there is a depletion of H in the sample(fig. 4c). This phenomena is more marked in β phase (fig.4d) with low field. May be a still more intense field in air might allow us to reach the results obtained in α phase in vacuum. Thus concerning the infrared spectra we have found sweeping conditions under vacuum which allow to get similar results in α and β phases. However our β swept samples becomes twinned. The study of twinning and detwinning has been the object of much work some decades ago (3, 13). Frondel found that thin samples can pass through the transition without twinning while Wooster et al. had detwinned bar sample by application of adequate stresses (based on the fact that twins have different elastic modulus). Up to now we were not successful in obtaining a β phase swept thin sample (AT cut 1 mm thick) free of twins. Perhaps because sweeping can introduce strains as recently evidenced (14).

CONCLUSION

We have shown that it is possible to obtain samples totally free of as grown OH. This can be obtained either in β phase (due to high temperature) or in α phase (due to the high electric field) with vacuum-sweeping while in air sweeping such effect will be more difficult. In order to compare the efficiency of β and α phase electrolysis, other measurements (dielectric absorption, E.P.R., X controls ...) are necessary. One should also clarify the possibilities of detwinning.

ACKNOWLEDGEMENTS

This work is supported by a 87-126 DRET contract. The authors wish to thank P. Palleau for his assistance and the building of the vacuum furnace, B. Viard (Besançon) for impurities determination, X. Buisson (SICN) and J.P. Aubry (CEPE) for samples, R. Bourquin (Besançon) for communication of his unpublished results.

REFERENCES

(1) J.C. Brice Review of Modern Physics 57 105-146(85)
(2) H.G. Lipson and A. Kahan IEEE Trans. on Nuclear Science NS 31 1223-1229(1984)
(3) W.A. Wooster, N. Wooster, J.L. Rycroft and L.A. Thomas Proc. Inst. Electr. Eng. 94 : 927-937 1947
(4) Dodd, Fraser. J. Phys. Chem. Solids 26 673-86(1965)
(5) G.B. Krefft Radiation Effects (1975) 26 249-259
(6) D.B. Fraser Physical Acoustic W.P. Mason ed. (Academic Press, NY) Vol V, 59 (1968)
(7) G. Dolino, J. P. Bachheimer, M. Vallade Appl. Phys. Lett, 22, 623-625 (1973)
(8) J. G. Gualtieri 42 nd A.F.C. Symposium(1988) 155-161
(9) A.K. Kronenberg, S.H. Kirby American Mineralogist 72 739-47(1987)
(10) J. Verhoogen Am. Mineralogist 37, 637 (1952)
(11) J.J. Martin IEEE Trans. on Ultrasonics Ferroelectrics and Frequency control (1988) 35 288-296
(12) J. Felsche und J. Lietz N. Jb. Miner. Abh. 109 250-258 (1968)
(13) Frondel Am. Mineralogist 30, 447 (1945)
(14) M.T. Sebastian, A. Zarka, B. Capelle J. Appl. Cryst. (1988) 21 326-329

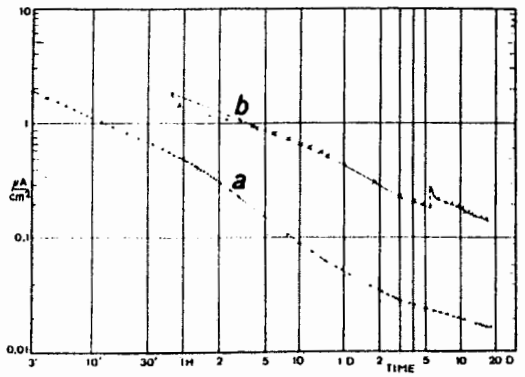


Fig. 1 : Vacuum sweeping :
Sweeping current density vs time for two samples from the same synthetic bar a : α phase 500°C-1500 V/cm
b : β phase 650°C - 1270 - 1500 V/cm

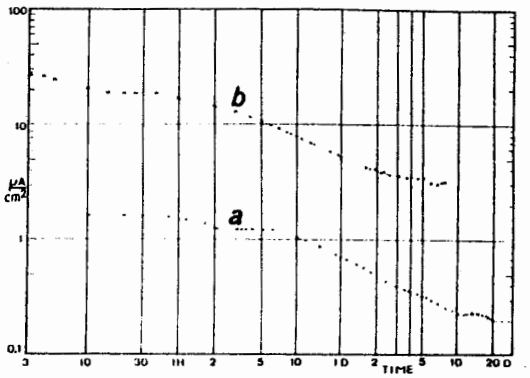


Fig. 3 : air sweeping :
Sweeping current density vs time for two samples from the same synthetic bar as fig. 1. a : α-phase : 500°C - 2000 V/cm b : β phase 650°C - 1270 V/cm.

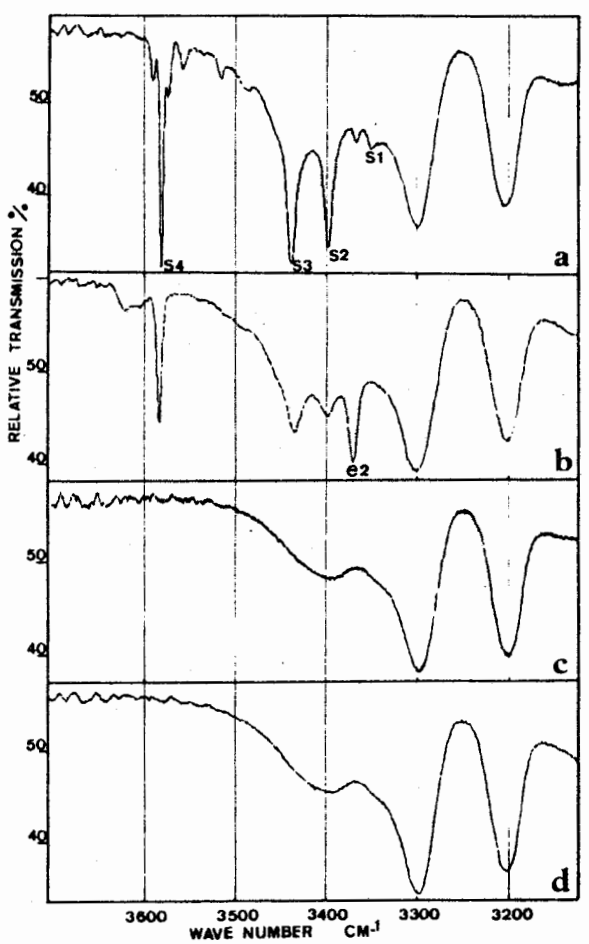


Fig. 2 : vacuum sweeping : IR spectra (beam along Z - 77 K)
a : as grown
b : 500°C - 1500 V/cm 18 days and with 550°C-1500 V/cm 8 days
c : 550°C- 3500 V/cm 6 days (in addition to b)
d : 650°C- 1500 V/cm (fig. 1) 18 days

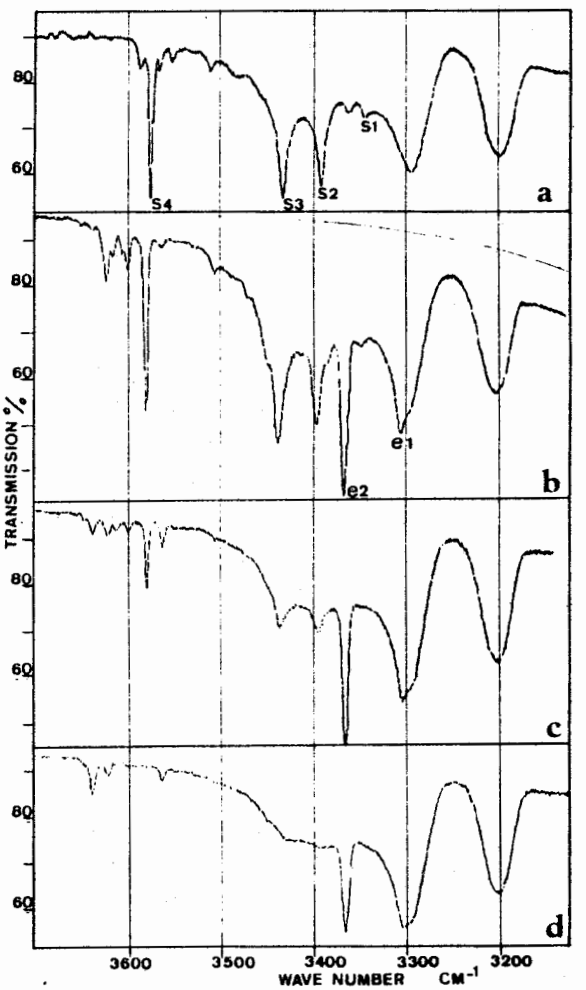


Fig. 4 : air sweeping : IR. spectra (beam along Z - 77 K)
a : as grown
b : 500°C - 2000 V/cm 30 days (— — —) base line
c : idem b + 550°C - 3500 V/cm - 7 days
d : 650°C - 1270 V/cm - 8 days

MILLISECOND PULSARS AND TIME SCALES

Jean-François Lestrade

Bureau des Longitudes
77 avenue Denfert Rochereau
75014 Paris

Pulsars are celestial clocks. However, they have never played any role in the establishment of time scales because the timing of their pulses has not been accurate in the past and because they are complicated physical objects difficult to use for definition purposes. The discovery of millisecond pulsars might change somewhat this situation.

Pulsars:

Pulsars are celestial objects emitting pulses at radio wavelengths which are detected by astronomers' radiotelescopes. These celestial objects are neutron stars, i.e. stars which have reached a very late stage of their evolution (life). This late stage of stellar evolution corresponds to the gravitational collapse of the star core after the chemically-riched outer layers of the star have been expelled into the interstellar medium in producing the supernova phenomena. At this stage, the star is in a highly-condensed state of matter and its diameter is only $\sim$ 10 km (the radius of the middle-aged sun is 700 000 km). The radio pulses observed by astronomers originate from a sharp beam of synchrotron radio emission attached to the neutron star surface and rotating with it. If the geometry is such that this beam crosses the solar system, a radio pulse is detected by radiotelescope once per rotation of the neutron star. So the faster the star rotates the higher the pulse frequency is. This phenomena at radio wavelengths is analagous to the optical blinks from a lighthouse along the sea-shore.

In 1982, the astronomical community was fascinated by the newly discovered pulsar PSR1937+214 which had the unprecedented period of rotation of 1.55.... millisecond. The corresponding angular velocity of the associated neutron star is as high as 640 rotations per second, i.e. 20 times faster than the young pulsar in the Crab nebula and 1000 times faster than ordinary pulsars. Sky surveys were conducted after this exceptional discovery and 7 pulsars with similar properties are known today and are designated millisecond pulsars for the magnitude of their periods. They form a new physical class in pulsar astrophysics characterized by the following phenomenological properties: their periods are a few milliseconds, their period time derivatives are very small ($< 10^{-18}$ s/s), their pulse profiles are very sharp and their timing measurements indicate no phase noise so far. These pulsars are located in the Galaxy (the Milky Way) as well as in globular clusters which form an halo around the Milky Way. Their distances are a few thousands light years from the solar system.

Timing measurements of pulsars:

In radio astronomy, timing observations of pulsars consists in measuring the arrival times of radio pulses at successive positions of the Earth (radiotelescope) along its orbit, i.e. several times per month over several years. These timing measurements are recorded relative to atomic clocks transferred to the radiotelescopes through the LORANC, TV or GPS systems. The precision of timing measurements of millisecond pulsars is at the level of 1 μs or better. Such measurements are done by coherently integrating about one hour of data. This precision is about 1000 times better than for timing done on ordinary pulsars for astrophysical purposes. The properties of millisecond pulsars (no phase noise, no higher order instabilities in the period than the first time derivative) and the precise timing measurements possible with radio astronomical techniques make millisecond pulsars the most "stable celestial clocks" ever discovered by astronomers.

Stability of millisecond pulsar time scale:

The stability of the timing measurements of the best observed millisecond pulsar (PSR1937+214) has been assessed by Rawley et al (1988) from the data acquired since 1982 with the large radiotelescope at Arecibo. Timing of PSR1937+214 is done at Arecibo once a week as an average but more frequent observations (once a day) or long gaps (up to 2 months) have occured. Rawley et al (1988) have analyzed a timing series of 4.2 years with a physical model including the effects of propagation of radio waves in the ionized interstellar medium and in the gravitational field of the solar system and geometrical effects due to the orbital and rotational motions of the Earth. The post-fit residuals between this model and the timing measurements recorded relative to the NBS atomic time scale show rms at a level of 1μs for the first 2 years and of 0.3μs for the last two years after some hardware improvements were made at Arecibo (see Figure 1).

In the adjustment of Rawley *et al*, 8 parameters associated with the pulsar (position, proper motion, period, period time derivative, epoch of first pulse and ionized interstellar dispersion constant) were solved for. The fitting procedure necessarily absorbs some fraction of the measurement uncertainties into biases in the adjusted parameters with the presently limited time interval spanned by the data. For instance, annual systematic error in these 4.2-year timing measurements caused by seasonal variations in the atomic time scale used (or by the Earth orbit ephemerides used) would be fully absorbed in a position bias of PSR1937+214 and the stability inferred from the post-fit timing residuals would appear to be better than it really is. In the future, longer timing series of several millisecond pulsars should significantly reduce such effect.

Nevertheless, bearing that in mind, Figure 2 from Rawley et al (1988) can be interpreted, with some caution, as an indication that the timing series of PSR1937+214 is extraordinarily stable. In this figure, the stability is characterized by the Allan variance of the post-fit residuals between the Arecibo timing measurements and the physical model over various time intervals τ. These authors conclude that for $\tau < 1$ month, the fractional frequency instabilities of PSR1937+214 are smaller than 5×10^{-13} and is dominated by random uncertainties in the timing measurement procedure of the pulsar waveform at the

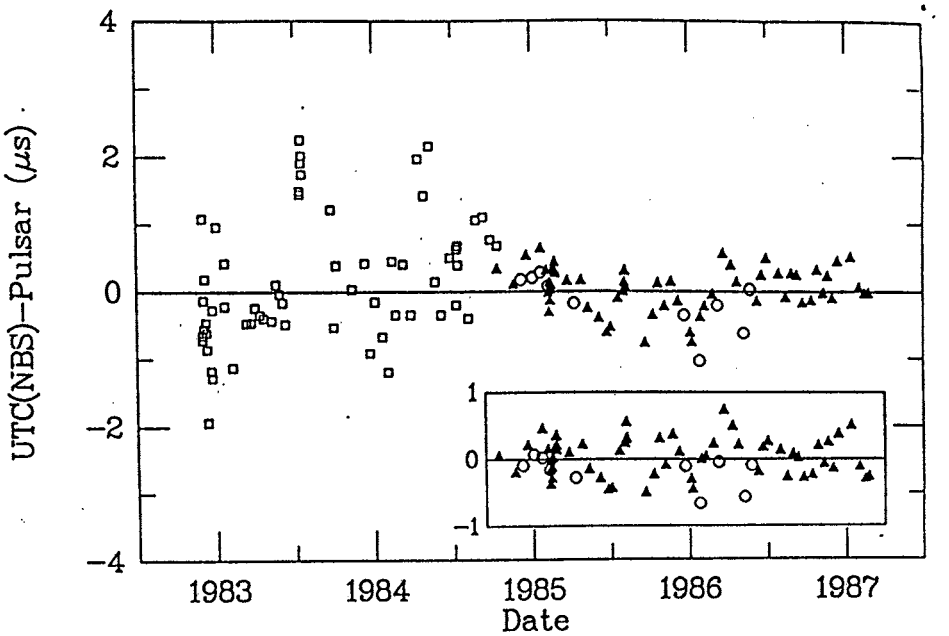


Figure 1 : Post-fit arrival time residuals for the millisecond pulsar PSR1937+214 observed at Arecibo (Rawley et al 1988)

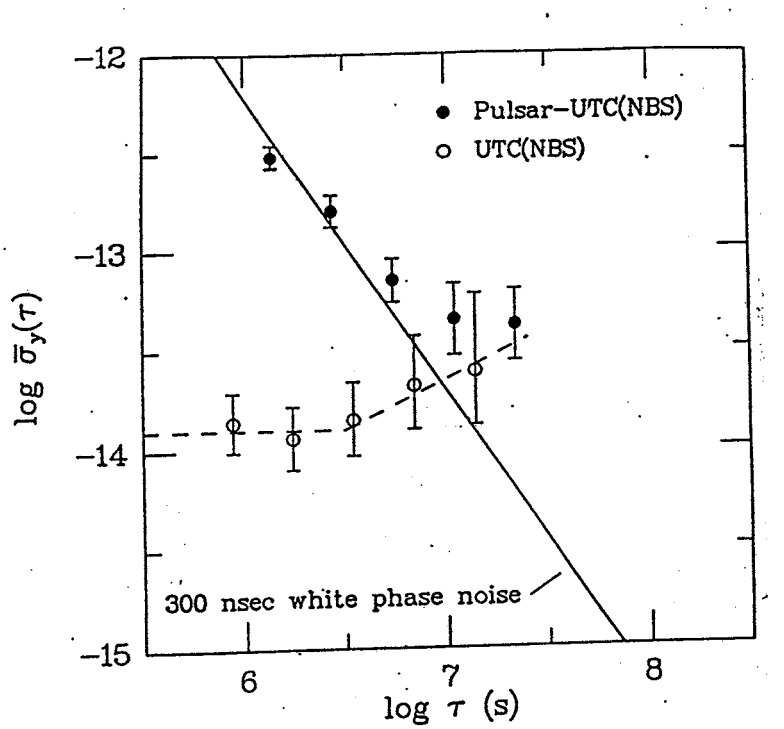


Figure 2 : Fractional frequency stabilities for measurements of PSR1937+214 relative to UTC(NBS) (filled circles) and UTC(NBS) relative to other atomic time standards (open circles) (Rawley et al 1988).

radiotelescope. For $\tau \sim 2-4$ months, the instabilities of PSR1937+214 are at the level of $\sim 10^{-13}$ and are marginally in excess of the instabilities of the reference time standards. But this is likely to be caused by the variable interstellar dispersion which was not properly corrected during the first two years of timing at Arecibo conducted at a single frequency. For $\tau > 6$ months, Figure 2 suggests that the stability of PSR1937+214 is better than 5×10^{-14} and that timing measurements of PSR1937+214 at Arecibo have been limited by the instabilities of the NBS atomic time scale used for the analysis. The PTB and BIPM atomic time scales are known to show fractional frequency stability at the level of 10^{-14} (Guinot 1988). This is smaller than for the NBS atomic time scale used for the analysis of Figure 2 and it is expected that a new reduction of the timing data with one of these two atomic time scales would indicate that the stability of PSR1937+214 for $\tau > 6$ months is at the level of 10^{-14} in reality.

Concluding remarks :

The extraordinary stability of the millisecond pulsar PSR1937+214 over several years open new prospectives for both astronomers and physicists responsible for the establishment of the atomic time scale. From the astronomer point of view, a better atomic time scale would allow monitoring of possible instabilities of millisecond pulsars at a higher level than is presently feasable. Such monitoring might allow astronomers to attribute these possible instabilities to disturbances of the propagation of the radio pulses caused by the gravitational radiation background of the early Universe as hypothesized in theoretical cosmology. This would be a new and major result in observational cosmology and it is eagerly sought by astronomers. They are looking forwards to using a new improved atomic time scale as soon as possible. Improvements in the atomic time scales are expected to come both from the algorithm used to combine primary cesium standards (e.g. BIPM 87 time scale : Guinot (1988)) and from hardware improvements in these primary time and frequency standards. Conversely, or before any significant improvements of the stability of the atomic time scale are acheived, post-fit residuals of Figure 1 strongly suggest than millisecond pulsar timing measurements could be used to enhance the stability of the atomic time scale on time interval longer than 6 months.

It should be noticed that millisecond pulsar timing measurements cannot be used to establish a new time standard which would be more *accurate* than the present atomic standards. Such assertion has been voiced in the astronomical community just after the discovery of millisecond pulsars. A frequency and time standard must be based on an invariable physical experiment that can be repeated identically for any duration at any date. For example, atomic transition is invariable from first principles in physics and the accuracy of any time interval it can defines is only limited by the quality and fidelity of the instrumental procedure used. The simple facts that the periods of millisecond pulsars are variable as shown by timing with existing atomic time scale and that no astrophysical theory can accurately provides the period time derivatives prevent millisecond pulsar timing measurements to produce a new time standard. In other words, pulsar timing series cannot become more *accurate* that the atomic time scales used to calibrate the period time derivative and in that respect cannot become a frequency and time standard. However, as

tentatively shown by Rawley et al (1988) with PSR1937+214, a pulsar timing series can become more *stable* than the atomic time scale used for the observations and advantages should be drawn from this result for the stability of the atomic time scale in the future.

Finally, we want to mention that several millisecond pulsar timing programs are in progress at Arecibo (USA), Greenbank (USA), Jodrell Bank (England) and Nançay (France). The instrumentation in Nançay is based on a new concept for de-dispersion of the pulsar signal based on a frequency swept local oscillator which is controlled by times of arrival of pulses calculated *a priori* (Biraud 1988). Timing campaigns in Nançay have started in the summer 1988 with this new instrumention and the first results are promising (Aubry et al 1988).

Acknowledgements :

I am very grateful to B. Guinot for discussions which have clarified several concepts related to the time and frequency standards for me.

References :

Aubry, D., Biraud, F., Bourgois, G., Darchy, B., Drouhin, J-P., Fairhead, L., Lestrade, J-F, 1988, Journal des Astronomes Français, vol. 33, Juillet 1988, p. 35.

Biraud, F., 1988, Proceedings of a Worshop on Millisecond Pulsar Timing held at Nançay, 1988 June 24, printed by the Observatoire de Paris, p. 1.

Guinot, B., 1988, Astron. Astroph., 192, 370.

Rawley, L.A., Taylor, J.H., Davis, M. M., Allan, D.W., 1988, preprint

DETERMINING THE EFFECTS OF MICROWAVE POWER AND C-FIELD SETTING ON THE FREQUENCY OF A CESIUM ATOMIC FREQUENCY STANDARD

S. K. Karuza, W. A. Johnson, and F. J. Voit

Electronics Research Laboratory
The Aerospace Corporation
El Segundo, California 90245
U.S.A.

Abstract

In experiments at the United States National Bureau of Standards, Andrea De Marchi showed that there exist optimum values of the C-field setting in a commercial dual-beam cesium frequency standard that make the output frequency insensitive to variations in microwave power. Moreover, and most important, De Marchi demonstrated that the long-term stability of the clock is improved if the C-field is set at one of these optimum settings. To see if these results can be obtained with a clock having a different modulation scheme, the authors, using a completely automated measuring system, performed a similar study on another manufacturer's commercial cesium standard that employed single-beam tube optics and a modulation scheme different from that used by the clocks studied by De Marchi. Our results, similar to those of De Marchi, show that optimum C-field settings exist for the standard we investigated.

Introduction

De Marchi has presented data [1,2] on three Hewlett-Packard (H-P) model 1653 cesium (Cs) frequency standards. His data show that one can improve the long-term frequency stability of such standards by selecting a value of Zeeman frequency (C-field) that reduces the frequency sensitivity of these H-P standards to variations in microwave power (P).

For example, Fig. 1, which is calculated from De Marchi's data [1] is a plot of the fractional frequency change for a +1-dB change in microwave power as a function of the Zeeman frequency (f_Z). It can be seen that there are four Zeeman frequencies (corresponding to four C-field settings) where the change in frequency will be zero for a change in microwave power

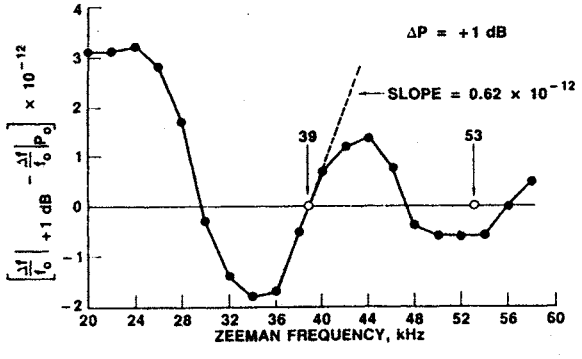


Fig. 1. The difference in residual frequency changes between two power levels (P_o and +1 dB) as a function of C-field in an H-P 1653 cesium frequency standard. Circles show points at a power-insensitive point (39 kHz) and a power-sensitive point (53 kHz). Data are taken from Ref. 1.

of +1 dB. The peak-to-peak fractional frequency change over the range of C-fields plotted is about 5×10^{-12} for this particular H-P standard. De Marchi demonstrated that there was an improvement in the long-term frequency stability of the standard as measured by $\sigma_y(\tau)$, the Allan standard deviation, when the Zeeman frequency was set at these optimum values. Figures 2 and 3 from De Marchi's data [1] show $\sigma_y(\tau)$ as a function of τ for the two Zeeman frequencies of 53 and 39 kHz, respectively, with 39 kHz being an optimum frequency. It is clear that there is an improvement in the long-term frequency stability if the C-field value is set at 39 kHz.

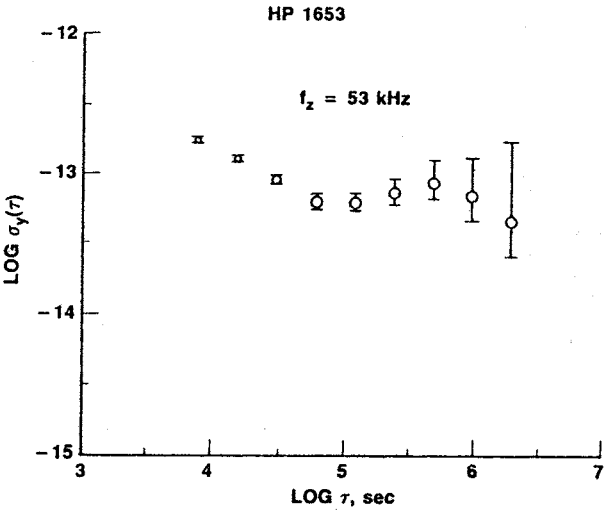


Fig. 2. The measured square root of the Allan variance of an H-P 1653 cesium frequency standard at a power-sensitive C-field setting of 53 kHz. Data are taken from Ref. 1.

De Marchi stated [1] that the results he had obtained on the H-P Cs standards should be " . . . at least typical for all commercial standards . . ." He cautioned, however, that results obtained on Cs frequency standards that used servo loop schemes other than sine-wave or slow square-wave frequency modulation might be somewhat different. Consequently, considerable interest developed to determine if the stability of other Cs frequency standards using different modulation schemes could be improved by this technique of optimum C-field setting. Measurements similar to those of De Marchi were made in our own laboratory on a different manufacturer's Cs standard that used a different modulation scheme. Because the measurements are very time consuming, it was decided to automate them completely in order to minimize the data-taking time. An additional advantage of this automation is that one never has to make and remake microwave power connections.

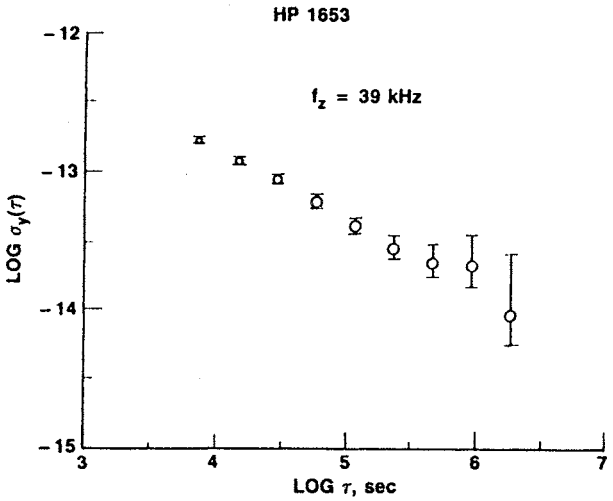


Fig. 3. The measured square root of the Allan variance of an H-P 1653 cesium frequency standard at a power-insensitive (zero crossing) C-field setting of 39 kHz. Data are taken from Ref. 1.

Measurement System

The C-field experiment was performed in our laboratory on a commercial Cs frequency standard that used a square-wave phase-modulation scheme. The standard was modified to allow access to the C-field coil wires and the microwave power source. Figure 4 shows the block diagram of the complete measurement system. Both of the parameters that are varied, namely the C-field current and the microwave power, are computer controlled; the current is set by a precision constant-current generator and the microwave power is changed by a calibrated PIN diode attenuator. The entire system is controlled by an H-P series 300 computer, which also acquires and processes the data.

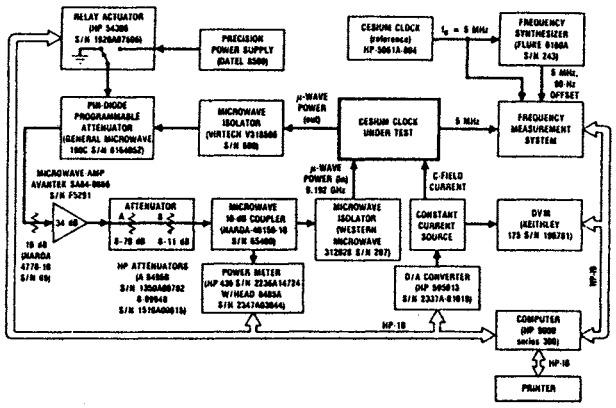


Fig. 4. Block diagram of the C-field measurement system for a cesium frequency standard. The system uses a digital computer as a controller and data gatherer.

Figure 5 is a block diagram of the frequency measurement system. The reference for both the Fluke synthesizer and the H-P counter is an H-P 5061A-004 Cs frequency standard. Before the data are taken, the microwave tuning adjustments in the standard and the microwave power are varied to maximize the output current from the beam tube. The resulting microwave power is called the optimum power (P_o).

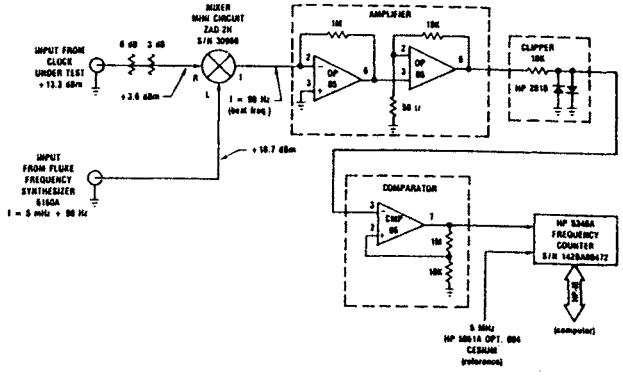


Fig. 5. Circuit diagram of the single-mixer frequency-measurement system used to determine the fractional frequency changes at different C-field settings

A typical data-taking sequence consisted of the following steps:

1. Set the C-field current at some low value (typically 6 to 8 mA) and the microwave power at some value (e.g. at the optimum value P_o).
2. Measure the beat frequency over some long averaging time T (typically 7000 sec).
3. Change the microwave power level (e.g. to P_o + 1 dB).
4. Measure the beat frequency over T again.
5. Increase the C-field current by some programmed amount (typically 0.5 mA).
6. Measure the beat frequency over T again.
7. Change the microwave power back to the initial value.
8. Repeat steps 2 through 7 until the final C-field current (typically 20 to 25 mA) is reached.

Figure 6 shows a plot of the noise floor of the frequency measurement system and a plot of the frequency stability of the commercial Cs frequency standard used in our measurements. The frequency reference standard was an H-P 5061A-004 Cs standard.

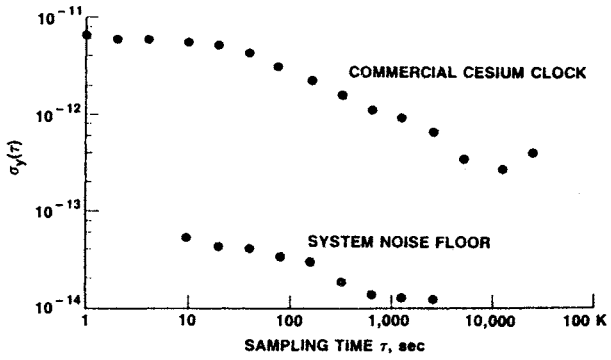


Fig. 6. The measured square root of the Allan variance of the commercial cesium frequency standard and the noise floor of the frequency measurement system

Measurement Results

Figure 7 shows the results of measurements made for changes in microwave power level of -1 dB, +1 dB, and +3 dB. For the -1-dB data, each point represents the difference between two 7000-sec samples; for the +1-dB data, each point represents the difference between two 14,000-sec samples; and for the +3-dB data, each point represents the difference between two 21,000-sec samples. Each data point is calculated as the difference in output frequency between the frequency at the higher power and the frequency at the lower power, both of which are normalized to the nominal output. In other words,

$$\text{ordinate} = (\Delta f_H - \Delta f_L)/5 \text{ MHz}$$

where Δf_H is the change in output frequency for the higher microwave power and Δf_L is the change in output frequency for the lower microwave power. As Fig. 7 shows, for the +1-dB and -1-dB data there is a zero crossing at about 26 kHz, but it is difficult to see if there are any other zero crossings. For the +3-dB data it is clear that there are two zero crossings at about 25 and 37 kHz.

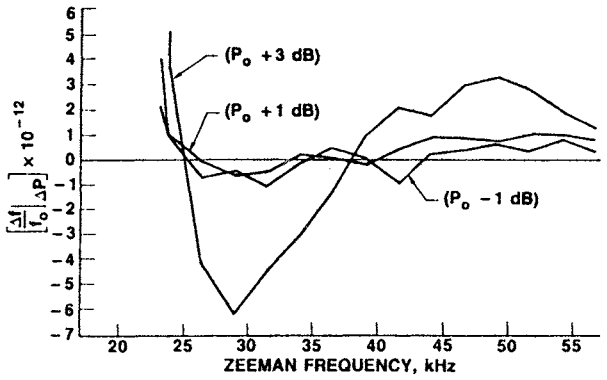


Fig. 7. A plot of the difference in frequency changes, as a function of the C-field of a commercial frequency standard, for three microwave power levels (-1, +1, and +3 dB) with respect to the optimum power level P_o

Because De Marchi, using the H-P Cs frequency standard, had found multiple zero crossings for a +1-dB power change (Fig. 1), we decided to spend the time to make a statistically significant measurement on our commercial Cs frequency standard for the same +1-dB change. Figure 8 shows the results of this measurement, with each data point representing the difference between two 100,000-sec samples (the error bars are ±2 standard deviations). These data show distinct zero crossings at about 25 and 37 kHz, in agreement with the earlier results for the larger power change of +3 dB. Thus it may be possible to shorten greatly the amount of time it takes to determine the location of the 1-dB zero crossing by using data for the 3-dB change. This could reduce the data-taking time to as few as two or three days.

As Fig. 8 shows, the slope at the 37-kHz zero crossing is 0.11×10^{-12}/kHz, compared to the slope of 0.62×10^{-12}/kHz for the H-P standard (Fig. 1). Thus, for a given departure from the optimum Zeeman frequency, the frequency of our commercial Cs frequency standard would be from five to six times less sensitive to power changes than would be that of the H-P standard. We caution that these data are for a small sample of clocks (three H-P clocks in De Marchi's case and one from another manufacturer in our case), and may or may not be typical.

On October 11, 1988, our commercial Cs frequency standard was taken to the National Bureau of Standards (NBS) at Boulder, Colorado, for an evaluation of the standard's long-term frequency stability at C-fields

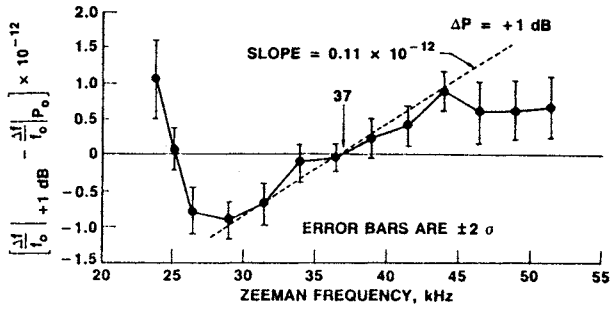


Fig. 8. Final average data of the difference in residual frequency changes, as a function of C-field, for a microwave power change of +1 dB above optimum power

corresponding to Zeeman frequencies of 37 kHz (the optimum found in our measurements) and 44 kHz to determine if the long-term frequency stability is improved by operating at an optimum C-field setting. We are presently in the process of making similar measurements on Cs frequency standards made by other manufacturers.

Inaccuracies in the Measurement System

Our measurement system introduces three sources of error or uncertainty: (1) frequency measurement errors, (2) C-field current setting errors, and (3) power setting errors. The first error, as shown in Fig. 6, has been shown to be almost two orders of magnitude below the measurement data. The uncertainty of the C-field current setting is probably on the order of parts in 10^4 in our laboratory environment over the three months during which data were taken; this stability is largely set by the stability of a precision film resistor. The third source of uncertainty, the measurement of the microwave power, is the most difficult of the three to assess. Figure 9 shows the measured power over 21 days of the C-field measurement time. Separate stability measurements were made on the power meter and head over about two weeks; it was found that the noise, as measured by the standard deviation, in the measurement system was more than two orders of magnitude below the noise in the power measurements, as shown in Fig. 9. The data in this figure will be analyzed statistically and, in

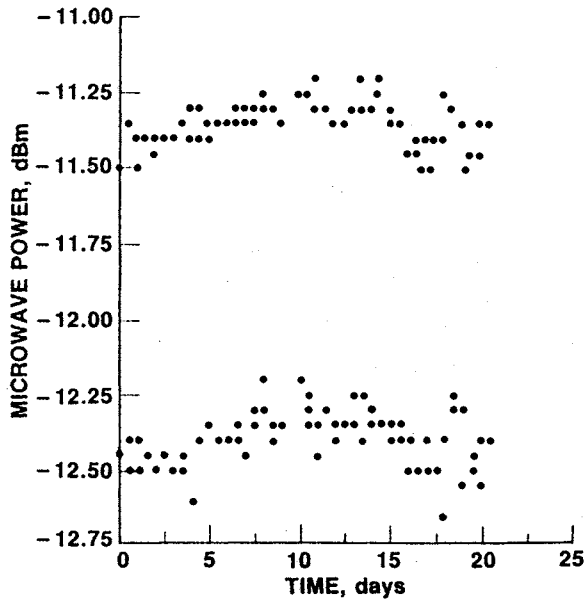


Fig. 9. A plot of power measurements made during the C-field experiment

conjunction with the data in Fig. 8, will be used to calculate the effects of these power variations on the noise floor of the clock.

Conclusions

The C-field experiment of De Marchi was performed on a commercial Cs frequency standard that used a square-wave phase-modulation scheme. The changes in microwave power level were from -1 dB to +3 dB relative to the optimum power level (the power level that gives maximum beam current). The results showed that two optimum C-field levels minimized variations in output frequency that were caused by changes in microwave power.

The Cs frequency standard was sent to NBS to determine if setting the C-field to this optimum value would improve the long-term frequency stability of the frequency standard.

Acknowledgment

The authors are grateful to David Allan and Andrea De Marchi of the National Bureau of Standards, and to John Hurrell of The Aerospace Corporation, for their helpful and stimulating discussions on the subject of optimizing the C-field tuning of cesium frequency standards. This work was supported in part by the U.S. Air Force Space Division under contract F04701-88-C-0089.

References

1. A. De Marchi, "New Insights into Causes and Cures of Frequency Instabilities (Drift and Long-Term Noise) in Cesium Beam Frequency Standards," in Proc. 41st Frequency Control Symposium (Philadelphia, Pa., 1987), pp. 54-58.

2. A. De Marchi, "Rabi Pulling and Long-Term Stability in Cesium Beam Frequency Standards," IEEE Trans. Ultrasonics, Ferroelectronics, and Frequency Control UFFC-34 [6], 598-601 (November 1987).

STUDIES OF COUPLING TO RAMSEY CAVITY IN CESIUM BEAM TUBE

L. JOHNSON P. THOMANN

OSCILLOQUARTZ SA - NEUCHATEL - SUISSE

ABSTRACT.

Use of an E-plane slot to couple the input feed to the Cs beam Ramsey cavity has been studied, along with measurement techniques using cavity arm decoupling at odd $\lambda_g/4$ arm lengths. The slot coupling offers the advantage of being able to displace the feed waveguide slighty to equalize the arm electrical lengths. Position sensitivity depends on slot dimensions. A coaxial - waveguide transition to the cavity input feed using lossy tuning elements was demonstrated as feasible. Such an arrangement allows passage of the coax through magnetic shielding instead of waveguide with an accompanying improvement in shielding.

I INTRODUCTION

The microwave system for interrogation of a cesium beam usually consists of providing two in-phase oscillating fields separated by a space of low magnetic field. This structure is referred to as a "Ramsey Cavity" and is illustrated schematically in Figure 1. In commercial Cesium-beam tubes the cavity is usually formed from WR-90 waveguide with the arms bent 90° in the E-plane. The center feed is an open T-junction, again in the E-plane. The dominant requirement of this structure is in-phase oscillating magnetic fields at the end of the arms where the Cs atoms traverse the structure. To this end, the two arms are of equal electrical length : a multiple of the half guide wavelength $\lambda_g/2$.

II MEASUREMENT METHODS

Since the frequency accuracy of the Cs beam apparatus depends to a first order on the phase difference of the oscillating fields (the arm electrical length with respect to the center feed) it becomes of fundamental importance to measure this length as a part of the fabrication process. The method of accomplishing this uses the effect of decoupling the two arms electrically. This decoupling, measured either mechanically, or electrically, will determine the electrical length. Electrically the junction can be represented by a π network (Figure 3) and it can be readily seen that two arms become decoupled whenever the third (decoupling) arm presents a high (theoretically infinite) impedance. This will occur whenever the decoupling arm is short-circuited and odd multiples of $\lambda_g/4$ long. Evidently the arms will not decouple at the Cs frequency, since they are an even multiple of $\lambda_g/4$ at that frequency.

However, they can be decoupled by a short-circuit on the feed arm adjusted for odd $\lambda_g/4$. This decoupling is observed by connecting a generator to one arm and detecting the transmitted signal to the other arm. If, in the condition of decoupling, one measures the standing wave null on the generator arm, then interchanges the arms, the change in position of the standing wave null measures directly the arm length difference.

This method has the advantage of measurement at the Cs frequency, and the disadvantage that measurements are made without actual short-circuit terminations in place,hence relying on the quality of the electrical connection between the waveguide arm and the generator output arm. An alternate method uses the decoupling of the short-circuited arms at $7/4\,\lambda_g$ and $9/4\,\lambda_g$ and subsequent calculation of the electrical $_g$ lengths at Cs frequency. The method has the advantage of measurements with short circuits in place.Disadvantages are that measurements are not made at the operating frequency, calculations assume a normally dispersive characteristic of the wave-guide, and perturbing effects of the junction are small. These questions are examined in more detail in the following sections.

III SLOT-COUPLED T-JUNCTION

The E-plane slot-coupled T-junction is a junction wherein coupling between the feed arm and cavity arms is accomplished via a transverse slot the full waveguide width (Figure 2). From the figure it can be seen that the branch (feed) arm may be displaced along the direction of the cavity arms and, intuitively, the effective center of coupling the branch to the arms would also be displaced. This situation has the advantage of changing the electrical length of the arms without necessity for machining operations on the arm ends. With waveguide short circuits in place measurements of arm decoupling frequencies using a probe at the Cs beam slot opening (fig.1) at $7/4\,\lambda_g$ and $9/4\,\lambda_g$ would allow calculating the electrical length $_g$difference at Cs frequency, with subsequent adjustement by displacement of the branch arm.

A. Electrical Characteristics.

The equivalent circuit of the T-junction is shown in Figure 3, (REF [1]).Calculated values of (normalized) slot susceptance vs slot width are presented in Figure 4. Limits on accuracy are imposed by the physical characteristics assumed for the slot so that calculations for slot widths up to 6mm are reasonably accurate. Reference [1] discusses these factors in more detail. The schematic presentation in Figure 5 represents the junction and arms as presented in the actual measurement method. The ideal (no junction susceptances) decoupling electrical lengths, $7/4\,\lambda_g$ and $9/4\,\lambda_g$ are modified by the presence of the junction susceptances Ba/Yo. This apparent length change is summarized in table 1.

B. Electrical Length Determination.

The electrical length, $(2\pi/\lambda_g)\ell$, of a wave guide can be determined when λ_g and ℓ are known. λ_g, guide wavelength, is calculated from the relationship given in Eq.(1) Figure 6. Length, ℓ, is unknown due to the fact that the cavity arms are bent 90° making a physical measurement of length impossible. However, an effective length may determined by defining the frequencies at which the arm decouples as being those

for $7/4\,\lambda_g$. With the simultaneous equations formed, the effective λc can be determined, and then λ_g. Calculation of an effective λc is necessary, because the bending process does not necessary leave the guide width, a, constant over the full length of the arm. Once the effective length, ℓ, and cutoff wave length, λc, are specified, the λ_g at any other frequency may be calculated, and hence the electrical length, $\beta\ell$, at that frequency. Considering Fig. 6, it is evident that the decoupling frequencies are modified by the junction susceptances. This would produce a displacement of the actual λ_g vs frequency curve and thus the dispersive characteristics of the waveguide. This effect has been calculated and summarized in Table 2 for our requirements. The calculations assume that both ℓ and λc are independant of frequency. Because of the non-uniformities introduced by the bends, there is no reason to think that this is the case. Evaluation of this can best be accomplished by actual frequency measurements on the Cs beam tube itself, since the frequency is very sensitive to the phase difference between the two arms. To date, the technique has been used on Ramsey cavities with open T-junctions. Results have been satisfactory.

C. Feed Arm Displacement.

As mentioned above, the slot coupled T-junction offers the intrinsic advantage of displacing the branch arm a small amount necessary to achieve electrical length equality of the two arms. The sensitivity of this displacement in compensating for arm difference depends on slot width. Table 3 gives a summary of mesured values of the coefficient of branch (feed) displacement as a function of arm length change as measured at the two decoupling frequencies.

IV COAXIAL - WAVEGUIDE TRANSITION

Our study of coupling to the Ramsey cavity has been extended to the transition from coaxial line to waveguide as being included in the overall waveguide structure. The typical Ramsey cavity is energized by the branch (feed) guide (fig.1). Such an arrangement requires a large opening in the magnetic shielding surrounding the Cs beam interaction region with consequent perturbation in magnetic homogeneity between the cavity arms. A coaxial feed at the Ramsey cavity, by contrast, would require only a small opening in the shielding through which the coax could pass. The concept is illustrated in Figure 7. It is possible to match the transition structure and realize a desired Q of the coupling system by using lossy tuning elements. FIG 8 shows some results of this approach.

Advantages of the method :

1 - Prescribed transmission loss and Q can be obtained and are a permanent part of the beam tube.

2 - The beam tube presents a reasonably matched load for the external microwave interrogating system, simplifying electronics.

3 - The effective temperature dependance of the microwave feed at the beam tube can be characterized.

V CONCLUSIONS

A. Use of slot-coupled T coupling to the Ramsey cavity has been studied showing the possibility of displacing the branch arm to equalize the electrical length of the cavity arms. The sensitivity of this displacement depends upon dimensions of the slot.

B. Calculation of arm electrical lengths by measuring decoupling frequencies other than the Cs frequency are perturbed by the junction susceptances within acceptable limits.

C. A tuned-lossy structure for coupling coaxially to the Ramsey to branch arm was investigated and showed satisfactory results.

VI REFERENCES

[1] N. Marcuvitz. Waveguide Handbook Vol. 10

MIT Radiation Laboratory Series Boston Technical Publishers 1964.

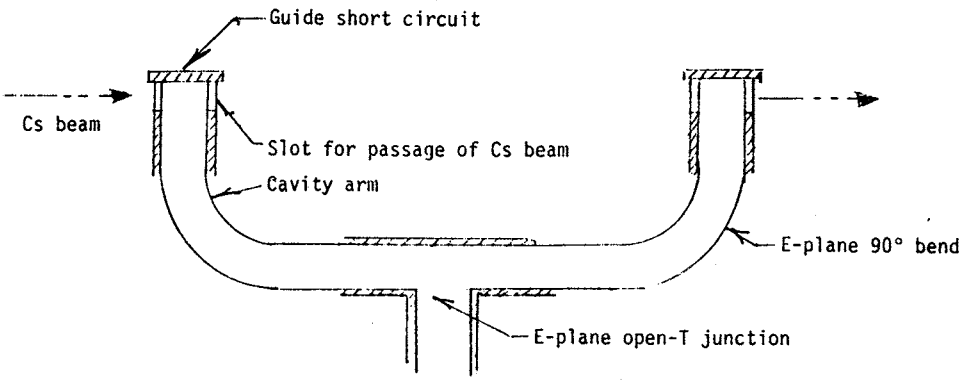
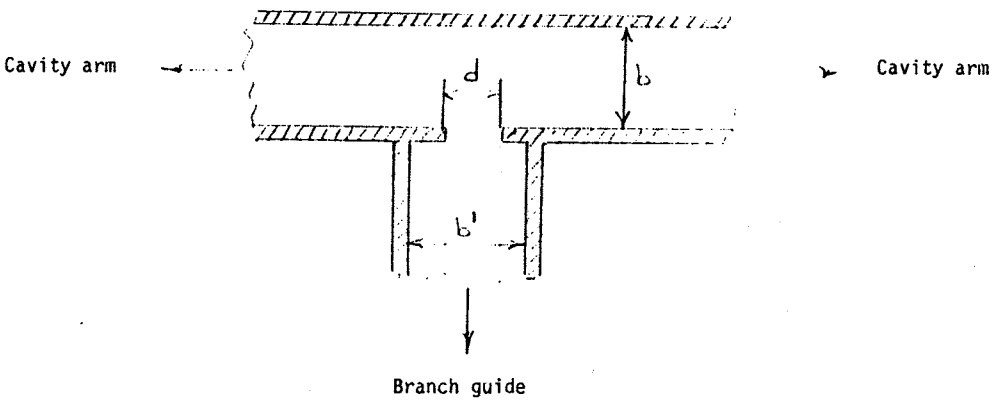
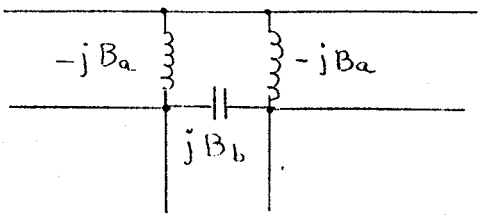


Figure 1. Ramsey cavity cross-section



b = cavity arm guide height
b'= Branch arm guide height
d = slot width
(Note : slot traverses entire guide width)

Figure 2. E-Plane slot-T junction cross section



$$B_a/Y_o = \frac{b/\lambda_g\left(\pi d/4b\right)^2}{1+\frac{1}{6}\left(\pi d/4b\right)^2\left\{\left(\frac{b}{b'}\right)^2+\frac{1}{2}+6\left[1-\sqrt{1-\left(2b/\lambda_g\right)^2}\right]+12\left(\frac{b}{b'}\right)^2\left[\frac{1}{\sqrt{1-\left(2b'/\lambda_g\right)^2}}-1\right]\right.}$$

Where Yo = waveguide characteristic admittance

λg = guide wavelength

b, b', d as defined in FIG.2.

Figure 3. E-plane slot-T equivalent circuit

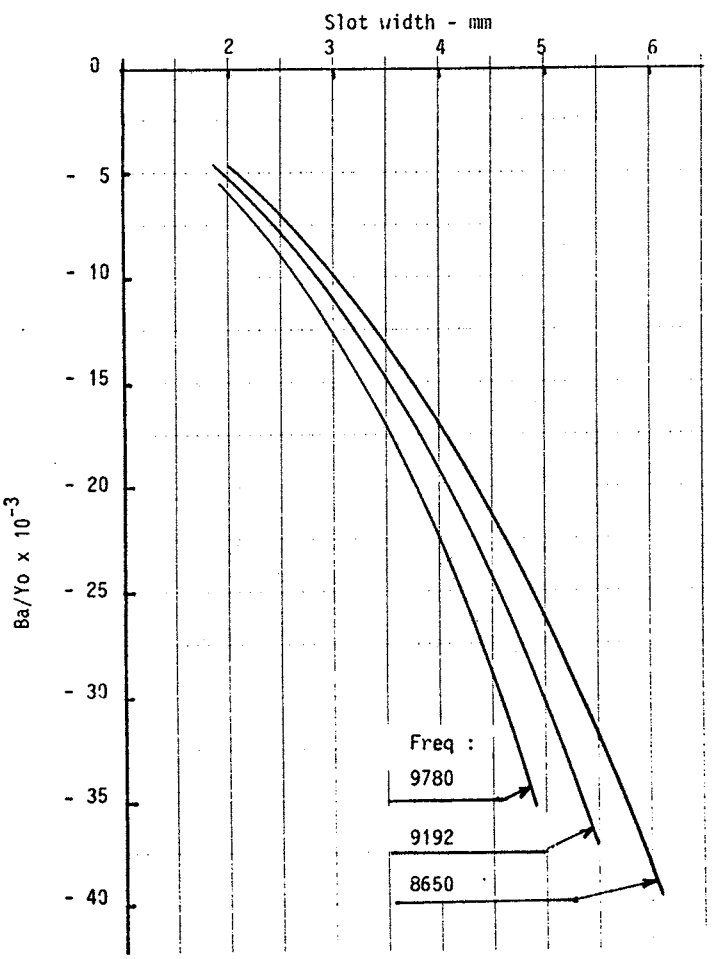
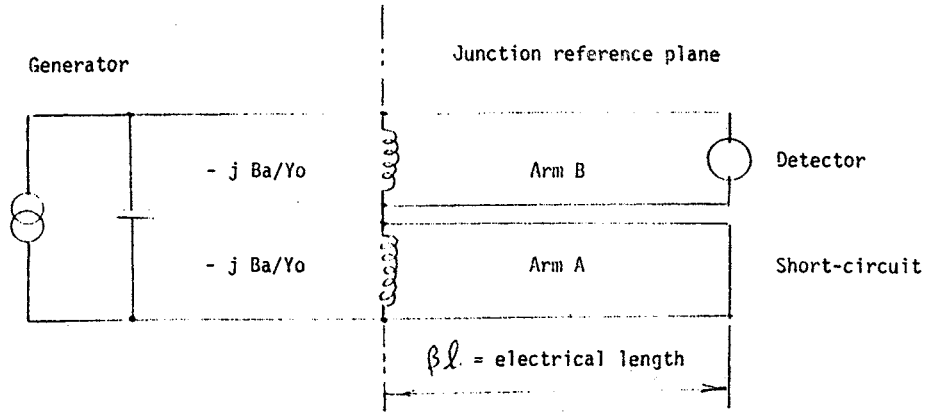


Figure 4. Ba/Yo vs Slot width



Ideal condition : Ba/Yo not present and decoupling of arm occurs for $\beta\ell = 7/4\lambda g, 9/4\lambda g$.
Where $\beta = 2\pi/\lambda g$
ℓ = arm length

Actual condition : Arm A susceptance at reference plane must equal + j Ba/Yo.
For ideal short circuited line :
Cot $\beta\ell$ = Ba/Yo and thus the actual decoupling frequency is determined by the necessary λg to make cot $\beta\ell$ = Ba/Yo

Figure 5. Schematic representation for decoupling frequency.

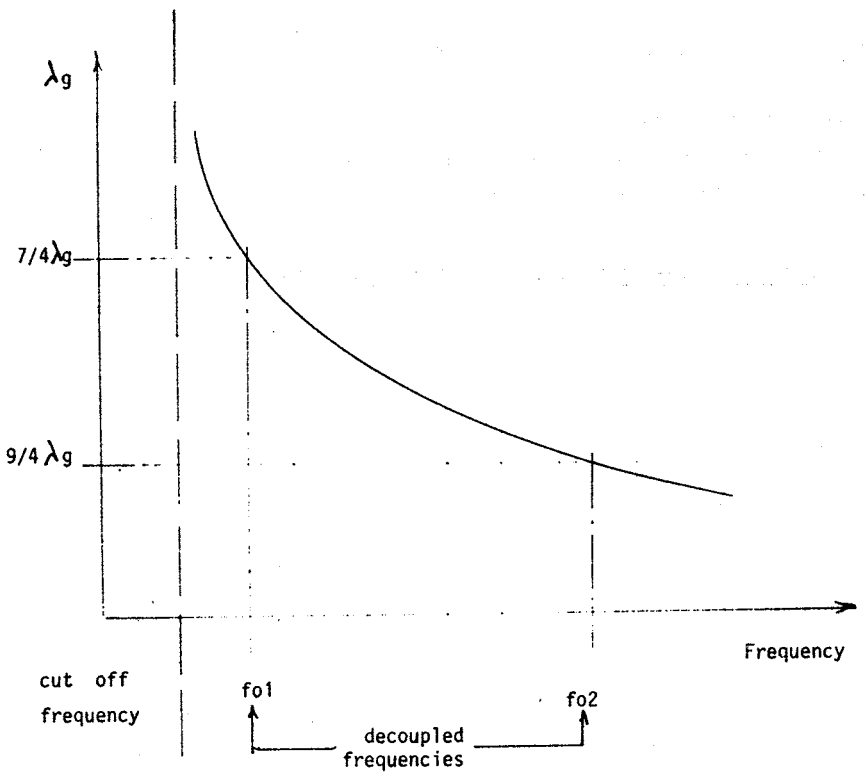
To calculate $\beta\ell$, λ g and ℓ must be known.

λg is given by :

$$\lambda g = \frac{\lambda o}{\sqrt{1 - (\lambda o / \lambda c)^2}} \qquad (1)$$

Where λo= free space wavelength
λc= guide cut off wavelength (= 2 x guide width, a)

This is illustrated graphically :



From decoupling frequency measurements, λc can be determined, hence $\beta\ell$ for any given f.

Figure 6. Determination of λ c from decoupled frequencies.

Slot width	Decoupled Frequency	Arm length difference - mm	Decoupled Frequency	Arm length difference - mm
2	8648.29	3.7879×10^{-2}	9775.37	3.7217×10^{-2}
3	8650.13	8.436×10^{-2}	9778.05	8.3181×10^{-2}
4	8652.61	14.604×10^{-2}	9781.58	14.353×10^{-2}
5	8655.68	22.462×10^{-2}	9785.87	21.874×10^{-2}
6	8659.28	31.474×10^{-2}	9790.73	30.254×10^{-2}

Calculated for ℓ = 93.077 mm ($\beta\ell$ = 8/4λg at 9192 MHz)

Arm length difference at 7/4λg and 9/4λg due to junction susceptance.

TABLE 1

Slot W	Length mm			Calculated Diff.		Difference Ideal Junction
	ARM A	ARM B	DIFF.	Ideal	with junction	
4	93.077	93.0	0.077	0.07669	0.0781	-1.41×10^{-3}
5	93.077	93.0	0.077	0.07669	0.07672	-3×10^{-5}
4	93.077	93.2	0.123	0.12259	0.12099	1.6×10^{-3}
5	93.077	93.2	0.123	0.12259	0.12257	2×10^{-5}
OPEN T	93.0	92.0	1.0	0.99634	0.98728	9.06×10^{-3}

Dimensions in mm.

Effect of junction susceptance on calculated arm length at 9192 MHz

TABLE 2

SLOT W	Freq.	Coeff.	Freq.	Coeff.
9 mm	8687	3.8	9821	3.65
8.5	8663	4.7	9787	4.65
8	8720	5.4	9870	5.6
6	8714	8.5	9860	7.69
4	8485	30	10684 (11/4λg)	27.5

Coefficient of branch line displacement VS arm length change for various slot widths

TABLE 3

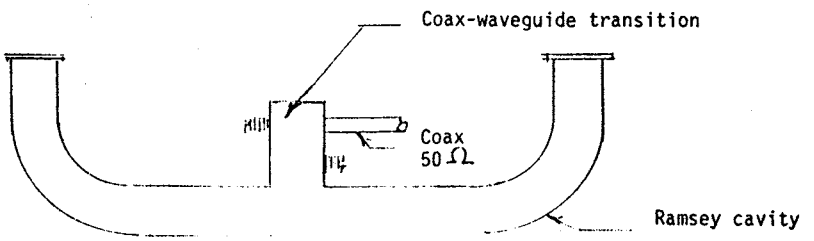
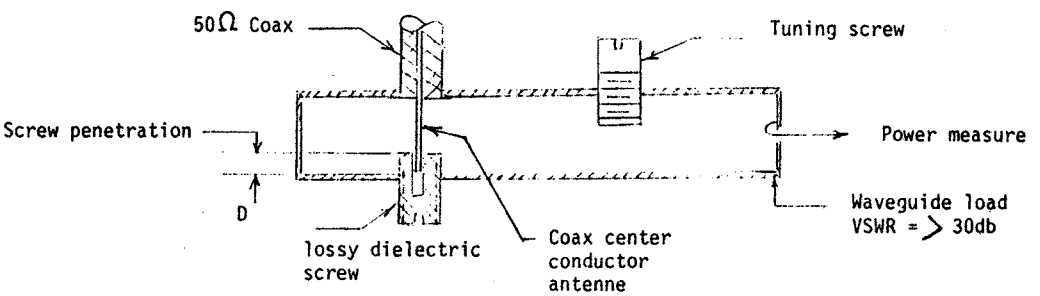


Figure 7. Ramsey with coax-waveguide tuned feed



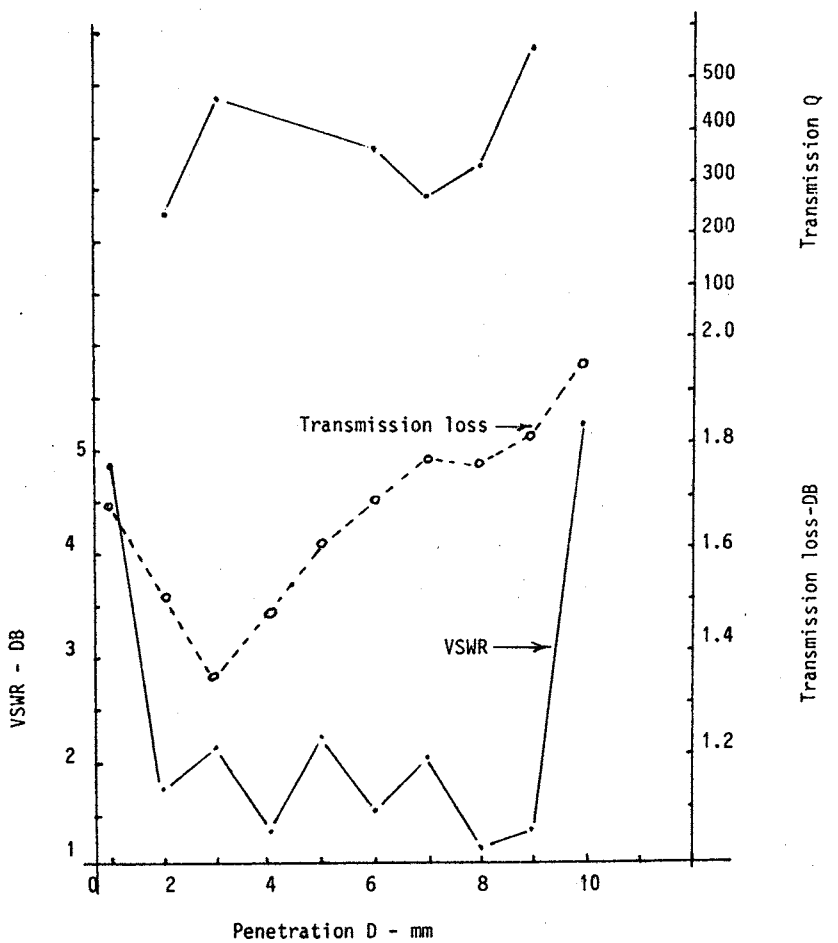


Figure 8. Coaxial feed characteristics

LONG TERM FREQUENCY STABILIZATION OF EXTENDED-CAVITY 1.5 µm SEMICONDUCTOR LASERS .

M. de LABACHELERIE, K.M. DIOMANDE, P. KEMSSU, P. CEREZ

Laboratoire de l'Horloge Atomique
Unité propre de recherche associée à l'université de Paris-Sud.
Bat. 221 - Université de Paris-Sud
91405 Orsay, France.

Abstract:

For coherent optical communications, a low linewidth frequency stabilized 1.5 µm laser would be very valuable. We have designed and operated an extended-cavity 1.5 µm semiconductor laser with a 50 kHz linewidth and a 2.10^{-10} frequency stability floor when locked to the side of an ammonia absorption line.

Introduction:

Extended-cavity semiconductor lasers (ECL) have recently been extensively studied to improve the spectral characteristics of laser diodes [1,2,3] : these lasers have spectral linewidths that are far much better than ordinary laser diodes. The main applications of such low linewidth lasers are connected with coherent optical communications, however, such systems require also long-term stabilization of the laser frequency that can be provided by frequency locking of the laser on a molecular absorption line. While long-term frequency stabilization has already been performed with integrated devices [4], no experiments have been presented for frequency stabilization of an ECL, which is a much more difficult task. After a presentation of our extended-cavity lasers, we will present our latest results concerning the long-term frequency stabilization of such lasers on Ammonia molecular lines around 1,5 µm.

1.Lasers required for coherent optical communications:

The recent interest for coherent optical communications comes from two main advantages of such systems compared to intensity modulated systems:

_ The first one is an increase in the detected signal-to-noise ratio which allows either to lengthen the optical link between repeaters, or to increase the bit rate for a fixed length.

_ The second advantage is the possibility of a very dense frequency multiplexing which allows to group a very large number of channels on the same fiber link.

However, these communication systems require a very high coherence of optical sources: for efficient modulation/demodulation formats, the emittor or local oscillator laser linewidth should be inferior to 0.1% of the bit rate, which imposes practically to attain linewidths on the order of 100 kHz [1]. Up to now, integrated semiconductor laser devices cannot reach such narrow linewitdhs, it is thus necessary to use improved sources such as extended-cavity lasers which have demonstrated linewidths as narrow as 15 kHz [5].

On the other hand, for frequency multiplexing, it is necessary to fix the absolute value of each channel optical frequency; otherwise, the frequency of emitting lasers could drift and would be detected by wrong local oscillators thus leading to a wrong signal detection. Such a fixed optical frequency for emittors and local oscillators can be obtained if one laser is long-term frequency stabilized on an absolute frequency reference, and the other are offset frequency-locked to the long-term stabilized laser, using for instance a Fabry Perot interferometer.

It is now clear that such coherent communication systems require laser sources that are simultaneously very coherent and long-term stabilized on absolute frequency references. In order to make such a laser source, we have choosen to make an extended-cavity 1.5 µm laser (for high coherence), and to stabilize its frequency to a molecular absorption line of ammonia around 1.5 µm (for long-term frequency stability).

2.Extended-cavity lasers:

General configuration :

The linewidth of semiconductor lasers being inversely proportionnal to the square of the laser optical cavity length, the best way to obtain linewidths small enough for coherent optical communications consists in using extended-cavity lasers (ECL) that are made of an anti-reflection coated laser diode coupled to a much longer passive external cavity [2,5].

We have designed and operated such lasers with either GaInAsP diodes (1.5 µm lasing wavelength) or GaAlAs diodes (0.85 µm wavelength): the AR-coated facet output light is collimated by a large aperture microscope objective (0.5 numerical aperture), and directed to a Littrow-mounted diffraction grating which ends the extended cavity (fig.1).

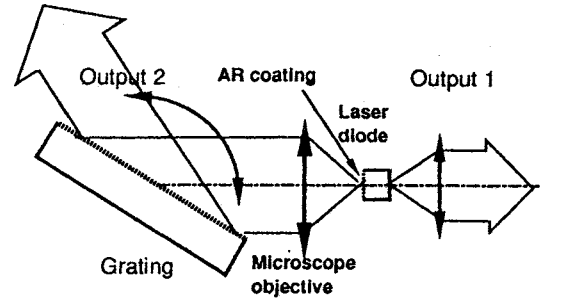


Fig. 1 : Extended-cavity laser

Improvement of spectral properties :

Thanks to the grating spectral selectivity, a broad tuning of the lasing wavelength can be obtained when the grating is rotated about an axis parallel to its grooves .While the tuning range of laser diodes is usually limited to a few nanometers, this external wavelength control allowed us to obtain a tuning range larger than 30 nm with AlGaAs 850 nm lasers, and 100 nm around 1.5 µm for InGaAsP lasers. This large tuning range is very interesting for any atom-light interaction such as laser spectroscopy or optical pumping in Cesium beam frequency standards, because these applications require an accurate spectral tuning of the laser.

It is interesting to notice that most of the buried heterostructure InGaAsP 1.5 µm diodes which usually emit a multi-frequency optical spectrum without external cavity, were always single frequency in an extended-cavity configuration.

The analysis of the RF frequency beat between two identical lasers allowed us to record an instantaneous extended-cavity laser linewidth between 10 and 50 kilohertz depending on the lasers and the operating conditions. However, the lasing frequency was sensitive to acoustical and mechanical perturbations of the cavity length which broadens the optical spectrum for longer observation times; for a 1 second observation time, the optical energy was thus distributed over a 1 MHz frequency band.

For semiconductor lasers, one of the major frequency noise source is related to the frequency perturbations due to reflected optical power in the laser cavity; however, as the sensitivity of a laser source frequency to optical feedback is inversely proportionnal to the laser optical cavity length, extended-cavity lasers are far less perturbed by optical feedback than ordinary diodes, for the same power level reflected in the laser. This is of great practical interest for any light-atom interaction because residual feedback is a major frequency noise source in this kind of experiment.

Mechanical design for a stabilized extended-cavity laser :

Extended-cavity lasers suffer from two main drawbacks:

_The lasing frequency is sensitive to acoustical perturbations of the cavity length.

_ The laser power and frequency are also sensitive to mechanical misalignements (the alignement of the laser cavity must remain stable to about 0.1 µm in 3 directions).

While acoustical perturbations can usually be corrected by an electronic feedback loop, we have not been able to find an accurate error signal to stabilize electronically the diode laser position in the extended-cavity. The key point to obtain a reliable stabilized laser will then be the mechanical and thermal stability of the cavity which has to be better than 0.1 µm to avoid mode hops that cannot be corrected by the frequency stabilization loop.

In order to make such a stable cavity we have choosen to :

_ Eliminate all mechanical movements except one movement which provides a fine frequency tuning for frequency stabilization.

_ Stabilize the temperature of the whole laser cavity within 0.1°C to limit thermal drifts.

Fig.2 shows the mechanical arrangement: the optical components are set in the right position, in order to provide a good alignement and a laser wavelength very close to the desired frequency, and then glued to a cooper plate, temperature stabilized within 0.01°C. A 2 mm thick silica plate inserted in the laser cavity and rotated with a piezoelectric actuator, allows fine tuning of its frequency over several Gigahertz.

In order to test the stability of such a mechanical arrangement, we have monitored the heterodyne beat between two identical free-running lasers on a 1GHz bandwith spectrum analyser and we have noticed that the frequency difference stayed in the ± 1GHz analysed band for several days.

3.Reference lines of NH_3 for frequency stabilization:

As was previously shown [6], new absorption lines that may be used for laser frequency standards should be find in the 1.3-1.6 µm fiber optics band. Unfortunately, avaliable data is not precise enough to choose accurately a new reference line; we have thus first realized a spectroscopic work to explore this wavelength region with a great precision.

Using an ECL we have obtained the first high resolution laser spectrum of Ammonia in the 1.49-1.52 microns range. The 30 strongest lines have been calibrated with a 10^{-7} accuracy using a wavemeter: Fig. 3 shows the recorded spectrum which contains many lines, often very well resolved and significantly absorbing for moderate gas pressures (typically 10 torrs) . However, these lines have a Doppler width of 600 MHz which leads to a practical linewidth of 1GHz if collisionnal broadening is added. Moreover, as no saturated absorption has been observed up to now, the reference lines are too wide to obtain the 10^{-12} stability already recorded [5] with the Cs-D_2 line at 852 nm.

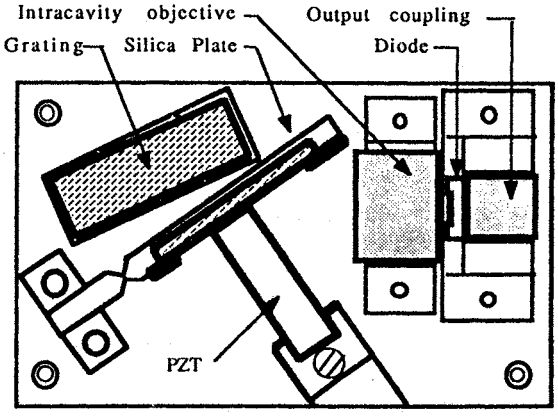


Fig.2 : Mechanical arrangement of the ECL

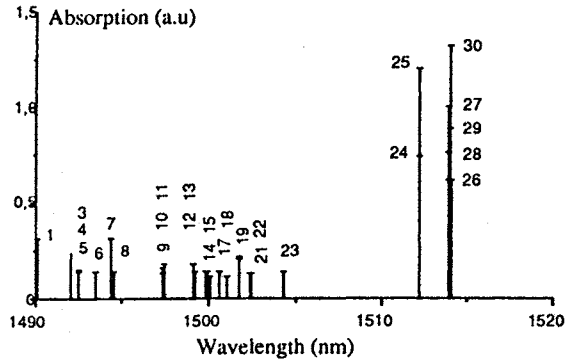


Fig.3 : Main NH3 lines between 1.49 and 1.52 microns

4.Laser long-term frequency stabilization:

One of the strongest lines recorded is a doublet at 1514 nm with a 1.3 GHz separation between each component; it can be used for a long term frequency stabilization of the ECL.

The absorption is high enough to use a 20 cm long absorption length (or a 7 cm cell with 3 pass) filled with 10 torrs of NH_3 to elaborate a precise error signal.

Detection technique :

Stabilization of the laser frequency on the top of the reference line requires a modulation of the laser frequency; however, optimum performances are obtained only when the modulation amplitude attains 30% of the reference linewidth. This condition could not be attained with our laser because its modulation capabilities did not allow to obtain the 300 MHz modulation amplitude that was required in good conditions: the 30 MHz/mA current to frequency conversion was too weak to minimize simultaneous amplitude variations, and the piezoelectric movement could lead to such wide frequency variations only for low modulation frequencies (less than 100 Hz).

We have then decided to use at first a much simpler detection technique to lock the laser to the side of the reference line (see fig.4): the laser beam is divided in two parts with a beamsplitter and the difference between the non absorbed power and the absorbed one is made by a differential photodetector.

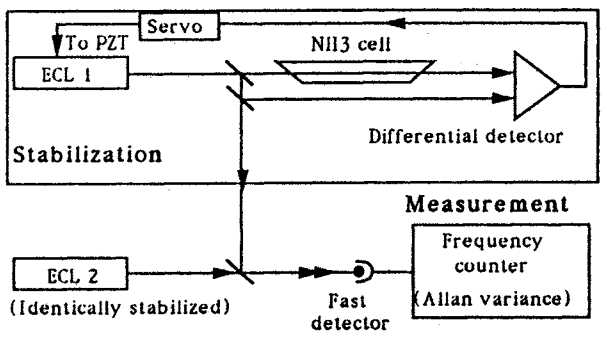


Fig.4 : Stabilization and measurement set-up.

Detected signals :

The absorption signals detected with a 20 cm absorption length and a 10 torr gas pressure usually attain 10 Volts with a detector noise of about 1mV in a 10 kHz band. However, care must be taken to choose a right photodetector/amplifier combination to obtain such performances. At 1.5 µm wavelength, it is necessary to use Germanium photodiodes which are very noisy and have a low shunt resistance : these detectors must be associated with a low voltage noise amplifier to obtain optimum performances. Another possibility consists in using InGaAs photodiode which have much better noise characteristics, and to associate them with a low current noise amplifier; however, InGaAs detector surfaces are usually small which imposes to focus the laser beam and enhances alignement stability tolerances.

Despite a large absorption signal, the sensitivity of the frequency to voltage conversion remains low (20 mV/MHz), because of the large absorption linewidth. However, such signals should logically lead to a 10^{-9}- 10^{-10} stability.

Results :

The frequency beat between two separately stabilized lasers has been recorded and the Allan variance of the frequency fluctuations has been recorded for 0.001 to 10 seconds observation time (Fig. 5). This curve shows that the locked laser stability attains 2.10^{-10} for long observation times.

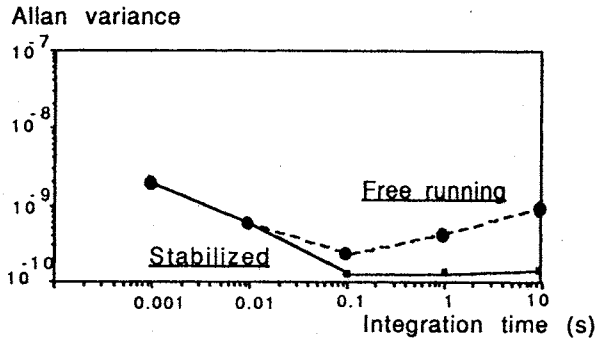


Fig. 5 : Variance of laser frequency fluctuations

Possible improvements :

Such a simple frequency stabilization technique is interesting because of its simplicity and also because the laser frequency is not modulated, however, it is not suitable to keep the stability for many days because the locking point depends on the laser power. It is possible to improve the detection technique by using an heterodyne technique which requires an external phase modulator, and lock the laser on the top of the absorption line.

5.Conclusion:

We have designed and operated a 1.5µm extended-cavity semiconductor laser locked to an absorption line of ammonia. The laser linewith is on the order of 50 kHz and we have shown that it is possible to obtain a 2.10^{-10} long term stability floor by locking it to the ammonia absorption line. Such a laser could be very interesting for coherent optical communications.

References:

1. T.G. Hodgkinson: Br.Telecom Tech.Journ.,3(1985), pp.5-18
2. M.Labachelerie: Optics comm. 55, (1985), pp.174-178
3. M.R.Matthews: Electron. Lett. 21, (1985), pp.113-115
4. T. Yanagawa: Appl. Phys. Lett. 45,(1984), pp.826-828
5. M.Labachelerie: Proc. EFTF 88, Neuchâtel, march 1988, pp.547-554.
6. DJE Knight: Proc. EFTF 88, Neuchâtel, march1988, pp.23-33

APPLICATION OF SQUEEZED DIODE LASER LIGHT IN RUBIDIUM ATOMIC FREQUENCY STANDARD

G.M. SAXENA

National Physical Laboratory,
Dr. K.S. Krishnan Road, New Delhi - 110012 (INDIA)

Abstract

The use of diode laser in atomic frequency standard has found wide acceptability as it has distinct advantages over conventional r.f discharge Rubidium lamp and filter cell combination(1). But the frequency and intensity instabilities of the laser light, due to quantum noise, affect the frequency stability of the atomic frequency standard. The application of the squeezed laser light can reduce the quantum noise to a large extent. In the squeezed light the desired pumping radiation is contained in a quadrature having fluctuations smaller than those in coherent states and other quadrature showing increased fluctuations due to the Heisenberg uncertainty relation. In this paper we have theoretically discussed the frequency stability of the squeezed diode laser pumped Rubidium atomic frequency standard and proposed a scheme for its realization using optical degenerate parameteric amplifier.

INTRODUCTION:

The application of diode laser for Rb standard has distinct advantages vis-a-vis the conventional r.f discharge Rubidium lamp and filter cell in terms of improved S/N and considerable reduction in light shift (2). The performance of the diode laser pumped Rb standard is, however, affected by the laser frequency and intensity fluctuations as these fluctuations are directly transferred to the frequency standard through the light shift effect. The frequency fluctuations in laser light due to quantum noise put an ultimate limit to the frequency stability of the laser pumped Rb standard. The squeezing (3,4) of the electromagnetic field is an effective means of reducing the quantum noise in the laser light. A field in a squeezed state is characterized by fluctuations in one quadrature component smaller than those in a coherent state, the other quadrature component showing increased fluctuations due to Heisenberg's uncertainty relation. The field quadrature components mentioned above are analogous to connonically conjugate position q and momentum p variables of a harmonic oscillator. In this paper we shall describe the application of squeezed laser light in the Rb standard and discuss theoretically the resulting improvement in the performance of the standard.

1. LASER PUMPED Rb STANDARD AND SQUEEZING OF LIGHT

The salient features of the diode laser pumped Rb standard have been described in a theoretical model of the Rb clock by Camparo and Frueholz (5). They have made theoretical calculations to quantify the expected performance improvement upon incorporation of diode laser in Rubidium standard. We shall confine our discussions to the effect of diode laser frequency fluctuations due to quantum noise on the Rb standard. And then show how the squeezing of light improves the clock's stability by suppressing the quantum noise in a quadrature. The fractional frequency fluctuations of the Rb standard due to laser frequency fluctuations are expressed as (Cf. Ref 5, eqn. 8)

$$\frac{\delta f}{f_o} = \frac{1}{f_o}\left(\frac{\partial^2 f}{\partial \nu \partial I}\right) I \delta\nu , \qquad (1.1a)$$

where f, ν, I and f_o denote, respectively the microwave hyperfine transition frequency, the laser frequency, the laser intensity and unperturbed Rb^{87} O-O hyperfine frequency. The above equation has been obtained with the following assumptions:

(i) The laser light and frequency noise are uncorrelated

(ii) The laser frequency is actively stabilized to the portion of the atomic absorption line where $\partial f/\partial I = 0$. Thus, the frequency fluctuations of the Rb standard only depend on the sensitivity of the hyperfine transition frequency to the laser frequency fluctuations $\partial f/\partial\nu$ and the magnitude of the laser's frequency fluctuations $\delta\nu$.] The squeezing of light will reduce the laser frequency fluctuations by suppressing the quantum noise in the quadrature of interest.

The eqn. (1.1a) for the squeezed diode laser light with the field in the desired quadrature can be rewritten as

$$\frac{\Delta f}{f_o} = \frac{1}{f_o}\left(\frac{\partial^2 f}{\partial \nu \partial I}\right) I \Delta\nu , \qquad (1.1b)$$

where Δf and $\Delta\nu$ are reduced fluctuations in the frequencies of the standard and the laser light respectively. Next we shall see how the squeezed laser light in the desired quadrature reduces the frequency fluctuations. The detailed properties of the squeezed light are given in refs. 3 and 4 and also in the refs. cited there. The squeezed state are generated by applying an unitary squeeze operator $S(\xi)$ and a displacement operator $D(Z)$ on the vacuum state $|0\rangle$, i.e.

$$|\xi, Z\rangle = S(\xi)\, D(Z)\, |0\rangle , \qquad (1.2)$$

where

$$S(\xi) = \exp\left\{\tfrac{1}{2}\xi (a^+)^2 - \tfrac{1}{2}\xi^* (a)^2\right\} , \quad \xi = r\, e^{i\theta} , \qquad (1.3)$$

and

$$D(Z) = \exp(Za^{+} - Z^{*}a). \qquad (1.4)$$

The squeezing factor ξ is generally a complex number. The complex number Z is the eigenvalue of the annihilation operator a , in the coherent state $|Z\rangle = D(Z)|0\rangle$. a^{+} is the creation operator. In the definition of the squeezed state $|\xi, Z\rangle$ the order of the operators $S(\xi)$ and $D(Z)$ may be interchanged depending upon the nature of the application with

$$D(Z)\, S(\xi) = S(\xi)\, D(\alpha), \qquad (1.5)$$

where

$$\alpha = Z \cosh r - Z^{*} e^{i\theta} \sinh r. \qquad (1.6)$$

Consider the squeezing of the light from a single mode diode laser. We assume that the laser light is defined by the state $|\Psi\rangle$, which may be a coherent state. The squeezed light is then given by the state

$$|\xi, \Psi\rangle = S(\xi)|\Psi\rangle . \qquad (1.7)$$

To find the frequency fluctuations in the squeezed laser light $|\xi, \Psi\rangle$ we consider the dual nature of the electromagnetic radiation. The frequency σ of the laser radiation and the momentum p of quanta (photons) in the direction of the propagation of the radiation are related by the equation

$$\sigma = \frac{c}{h} p , \qquad (1.8)$$

where C is the velocity of light and h is the Planck's constant.

Using the above equation the uncertainties in the laser frequency and the momentum of the photons are related as

$$\delta\sigma = \frac{c}{h} \delta p. \qquad (1.9a)$$

Similarly for the squeezed laser light the above uncertainties are related as

$$\Delta\sigma = \frac{c}{h} \Delta p, \qquad (1.9b)$$

where

$$\Delta p = \left\{ \langle\Psi|S^{+}(\xi)p^{2}S(\xi)|\Psi\rangle - \left(\langle\Psi|S^{+}(\xi)pS(\xi)|\Psi\rangle \right)^{2} \right\}^{\frac{1}{2}}.$$

For real values of ξ we find that

$$\Delta p = \left\{ \langle\Psi|p^{2}|\Psi\rangle - (\langle\Psi|p|\Psi\rangle)^{2} \right\}^{\frac{1}{2}} e^{-\xi}$$

$$= \delta p\, e^{-\xi}. \qquad (1.9c)$$

From equations (1.9a) - (1.9c) we have

$$\Delta\sigma = \delta\sigma\, e^{-\xi} , \qquad (1.10)$$

thus showing that the squeezing reduces the frequency fluctuations by a factor $e^{-\xi}$. This will be accompanied by an increase in the uncertainty in the time determination of the emission of the photons in the other quadrature because of the Heisenberg's uncertainty relation, and this will lead to increased amplitude (intensity) fluctuations. As it has been mentioned earlier that $\partial f / \partial I = 0$ and this is also valid for the squeezed light, the effect of the increased amplitude fluctuations in the other quadrature would not be of much significance to the Rb clock's performance for small laser light intensity. Thus even for the squeezed light the frequency fluctuations of the laser will determine the limit of the Rb clock's stability.

From the eqn. (1.1b) and (1.10) we obtain

$$\frac{\Delta f}{f_o} = \frac{1}{f_o} \left(\frac{\partial^2 f}{\partial\sigma\partial I}\right) I\, \Delta\sigma\, e^{-\xi}. \qquad (1.11a)$$

On comparing the eqns. (1.1a) and (1.11a) we observe that the squeezed laser light reduces the fractional frequency fluctuations of the Rb standard by a factor $e^{-\xi}$. In terms of Allan Variance eqn. (1.11a) may be written as

$$\sigma_y^{c}(\tau) = \frac{\sigma_o}{f_o}\left(\frac{\partial^2 f}{\partial\sigma\partial I}\right) I\, \sigma_y^{L}(\tau)\, e^{-\xi}. \qquad (1.11b)$$

Here $\sigma_y^{c}(\tau)$ is the Allan Variance of the clock's fractional frequency fluctuations due to the laser frequency fluctuations for a sampling time τ , $\sigma_y^{L}(\tau)$ is the Allan Variance for the laser frequency stability and σ_o is the frequency of the desired optical radiation. The improvement in the frequency stability of the standard resulting from the squeezing of laser light depends on the squeezing factor ξ . Larger the value of ξ , more is the stability of the frequency standard; the first step is to generate the squeezed states. From the nature of the squeeze operator $S(\xi)$ it appears that any phase sensitive non-linear process is a good candidate for the generation of the squeezed state. A optical degenerate parametric amplifier, which is a phase sensitive device, has been successfully used in generating squeezed states experimentally (6,7). In the optical degenerate parametric amplifier (8) (ODPA) a non-linear medium, having non-vanishing second order non-linear susceptibility, is pumped by an electromagnetic wave of frequency denoted by σ_p. The non-linearity of the medium couples this pump wave to two other identical wave modes called signal and idler, whose frequencies σ_s and σ_i satisfy $\sigma_s = \sigma_i = \sigma_{p/2}$. If the wave vectors of the three waves in the medium satisfy the phase matching condition, $\vec{K}_s = \vec{K}_i = \vec{K}_{p/2}$ then signal and idler are propagated through the medium in such a way that one quadrature phase of the signal (idler) mode is amplified and the other is attenuated resulting in the squeezing of light. The squeezing factor ξ here could be made large by choosing high values of the non-linear susceptibility, length of interaction and pump power. In the next section we shall discuss the experimental scheme for the squeezed Rb standard using ODPA as a squeezing device.

2. EXPERIMENTAL SCHEME FOR SQUEEZED RUBIDIUM FREQUENCY STANDARD

The schematic diagram of the experimental set-up of the squeezed Rubidium Standard is shown in the Fig. 1. Here a single mode AlGaAs diode laser tuned to the D_1 absorption line at 794.70 nm is used as a source of the pumping radiation. The laser frequency σ is doubled, using a crystal ($Ba_2\, Na\, Nb_5\, O_{15}$) frequency doubler, and is used as a pump for the optical degenerate parametric amplifier (ODPA). A non-linear optical crystal

$MgO:LiNbO_3$ of an appropriate length is placed inside the cavity formed by the mirrors M and M'. This results in down conversion of the frequency to $\nu_{p/2}$. The mirror M allows laser frequency and its first harmonic into the optical cavity of the ODPA while the mirror M' is coated for higher transmission at ν. The squeezed radiation emerging from M' falls on a Rb^{87} absorption cell containing Rb^{87} isotope and some buffer gas. The absorption cell is kept at a constant temperature inside a microwave cavity tuned to the Rb^{87} ground state hyperfine transition frequency. A small magnetic field is applied along the axis of the cavity to define the reference axis for the quantization and the removal of the degeneracy of the hyperfine levels. The light emerging from the absorption cell falls on a detector of high quantum efficiency and low noise. The output from the detector at the hyperfine reasonance is processed using a servo control loop and the necessary d.c. correction signal so obtained is applied to the VCXO for correcting its frequency.

The above scheme is supposed to improve upon the frequency stability of the standard by one to two orders. The model calculations based on the eqn. 1.11b, using the parameters given in the Table - 1, have been made for the frequency stability $\sigma_y^c(\tau)$ of the squeezed Rb standard for different values of the squeezing factor ξ and plotted in Fig. 2.

CONCLUSION:

We have theoretically discussed the possibility of using the squeezed states in the diode laser pumped Rb atomic frequency standard. As the squeezing has already been achieved experimentally, practical designing of a squeezed Rb standard is possible with the scheme suggested in the paper.

TABLE - I

Parameters used in calculating the Allan Variance, $\sigma_y^c(\tau)$ of the Rubidium Clock based on eqn. (1.11b) for ordinary and squeezed laser light. The graphical presentation of the projected frequency stability based on the following values of the parameters is given in Fig. 2.

Parameter	Value
Squeezed Laser Intensity $(I_s)^*$	1 mW/cm^2
Laser $\sigma_y(\tau)$ (Stabilized laser)	$\frac{9.0 \times 10^{-12}}{\sqrt{\tau}}$ (9)
$(\nu_o/f_o)\frac{\partial^2 f}{\partial \nu \partial I}$	0.68 cm^2/mW (10)
Buffer gas pressure (N_2)	100 Torr

* The laser light will be $I = KI_s$, with I_s = 1 mW/cm^2 taken for the present case. Here K is the combined efficiency of the frequency doubler and the ODPA.

REFERENCES:

1. Motoichi, Ohtsu, Minoru Hashimoto and Hidetaka Ozawa, Proc. 39th Ann. Freq. Contr. Symp., Philadelphia, U.S.A., p 43 (1985).
2. J. Vanier, R. Kunski, A. Brisson and P. Paulin, J. Phy. (Paris) Colloq. 42, Suppl. 12, C8-139 (1981).
3. D.F. Walls, Nature (London) 306, 141 (1983).
4. D. Stoler, Phys. Rev. D1, 3217 (1970), and D4, 1925 (1971).
5. J.C. Camparo and R.P. Frueholz, J. Appl. Phys., 59, 3313 (1986).
6. R.E. Slusher, L.W. Hollberg, B. Yurke, J.C. Mertz and J.F. Valley, Phys. Rev. Lett., 55, 2409 (1985).
7. Ling -An Wu, H.J. Kimble, J.L. Hall and Huifa Wu, Phys. Rev. Lett., 57, 2520 (1986).
8. Laser Handbook, Ed. F.T. Arecchi and E.O. Schulz Dubois (North-Holland Amsterdam) Vol. 1, p 837 (1972).
9. M. Ohtsu, H. Fukada, T. Tako and H. Tsuchida, Jpn. J. Appl. Phys., 22, 1157 (1983).
10. J.C. Camparo, R.P. Frueholz and C.H. Volk Phys. Rev. A27, 1914 (1983).

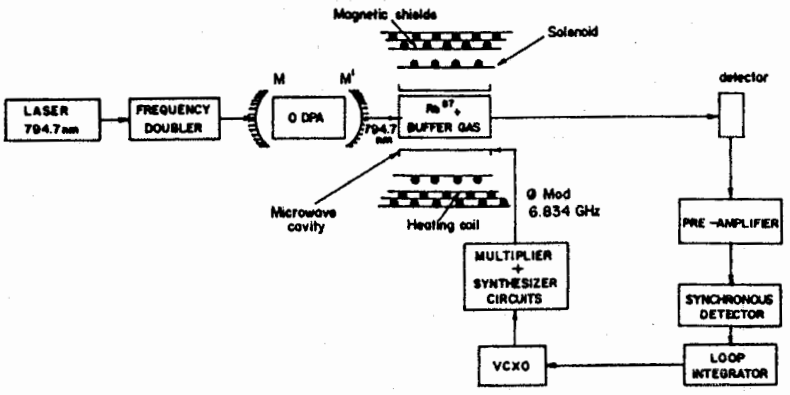


Fig. - 1. Schematic diagram of the proposed squeezed diode Laser pumped RUBIDIUM ATOMIC FREQUENCY STANDARD.

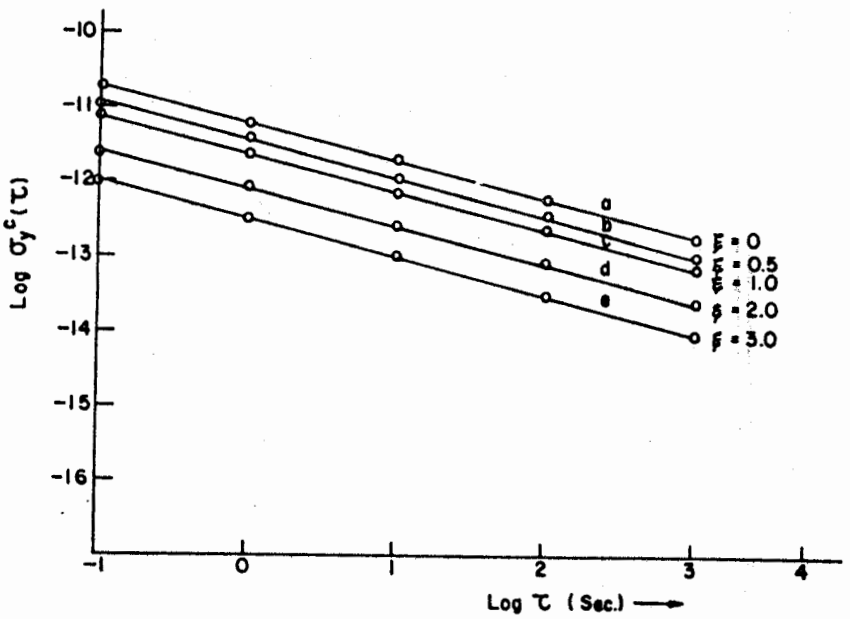


Fig. 2. Estimated frequency stability of the Rb^{87} Clock in terms of Allan variance $\sigma_y^c(\tau)$ with ordinary laser light (a) and sequeezed laser light (b,c,d,e)for different values of sequeezing factor ξ

A CODELESS DUAL FREQUENCY GPS RECEIVER MEASURING IONOSPHERIC EFFECTS AND ITS APPLICATION TO TIME COMPARISONS BETWEEN EUROPE AND USA

M. Imae*, M. Miranian**, W. Lewandowski and C. Thomas

Bureau International des Poids et Mesures
Pavillon de Breteuil
F-92312 Sèvres Cedex, France

*Communications Research Laboratory
Nukui Kita-machi, Koganei-shi, 184 Tokyo, Japan

** United States Naval Observatory, Time Division
Washington D.C. 20390-5100, USA

Abstract

The Global Positioning System is now widely used for daily time comparisons between worldwide spread laboratories, according to the common view method. The obtained precision is of the order of one nanosecond to a few tens of nanoseconds depending upon the distance between stations. In case of long-distance time comparisons, the estimation of ionospheric delays is one of the largest sources of uncertainty.

A new dual-frequency GPS receiver, named GTR2, has been developed at BIPM, for measuring the total electron content (TEC) along the line of sight and then the GPS signal ionospheric delay. GTR2 principle relies upon the properties of the cross-correlation between P-code (Precise code) signals of L_1 (1575,42 MHz) and L_2 (1227,6 MHz) carriers, without demodulation of GPS P-code.

GTR2 preliminary experiments allow TEC determination with 3 minute observation sequences, with an uncertainty of 2×10^{16} m^{-2}, equivalent to L_1 signal delay uncertainty of 1 ns.

Ionospheric delay measurements obtained in Sèvres (Paris area) from GTR2 appear to be larger than the corresponding values given by the model included in GPS message.

Ionospheric delay measurements are also performed by United States Naval Observatory in Washington and in Florida, with dual frequency GPS receivers which demodulate P-code. These results also present notable discrepancies with model values.

At last, the use of measured ionospheric delays largely improves long distance time compartsons as shown with the example of time links between Paris-Observatory and United States Naval Observatory in Washington.

INTRODUCTION

The major errors of time comparisons by simultaneous trackings of GPS satellites (common view method [1]) come from satellite position, receiving antenna position, estimation of ionospheric and tropospheric delays, calibration of receiver differential delays and radio-signal multi-paths.

GPS receiver transportation [2] and adapted means of reception largely reduce some of them. Short-distance comparisons (up to 1000 km) are weakly affected by satellite position error and by ionosphere effects [3], furthermore differential corrections can be determined for the antenna coordinates [4, 5], allowing time transfer with uncertainties of a few nanoseconds.

GTR2 development was carried out at the BIPM by Michito IMAE. He was partly supported, for this research, by the French Government through the Groupe de Recherches de Géodésie Spatiale.

But long-distance time links, like USA-Europe, are performed with a lesser precision, of the order of a few tens of nanoseconds. This is mainly due to satellite position errors and rather bad estimations of ionospheric delays, deduced from the ionospheric compensation model accessible to users of one frequency GPS receivers and providing a relative precision of 50 % [6].

We have developed a new codeless dual frequency GPS receiver, named GTR2, for measuring the total electron content (TEC) of ionosphere along the line of sight and then the GPS signal ionospheric delay [7]. In the following a brief description of GTR2 principle is given. The obtained results are described and analyzed by comparison with values coming from the ionospheric compensation model.

Ionospheric delay measurements are also performed by United States Naval Observatory (USNO) with dual frequency receivers which demodulate GPS signal P-code. Examples of the obtained results for two sites, Florida and Washington [8] are shown. At last GTR2 results, associated with USNO results, are applied to the time comparisons between Paris Observatory (OP, Paris, France) and United States Naval Observatory (USNO, Washington D.C., USA). This original study involves measured ionospheric corrections for the two branches of time link by GPS common view method; an example is given for a thirty days period (November 1988).

1. THE DUAL FREQUENCY METHOD - GTR2 PRINCIPLE

A radio-signal of carrier frequency L which crosses ionosphere is delayed by a quantity expressed, at the first order approximation, as follows:

$$T_{ion}(L) = \alpha \cdot \frac{TEC}{L^2} \qquad (1)$$

T_{ion} is expressed in ns, L in Hz, α is a coefficient equal to 134.36 and TEC is the total electron content (in m^{-2}) of ionosphere along the signal path.

TEC, which is directly linked to the electron density and thickness of ionosphere, largely varies with solar activity, local time, longitude and latitude of the reception station. Examples are given in Table 1 with the corresponding ionospheric delays for both carriers L_1 (1575,42 MHz) and L_2 (1227,6 MHz) of GPS radio-signals.

Table 1: Examples of vertical propagation delays of GPS L_1 and L_2 signals

TEC	$T_{ion}(L_1)$	$T_{ion}(L_2)$	$T_{ion}(L_2)-T_{ion}(L_1)$
2×10^{16} *1	1,08 ns	1,78 ns	0,7 ns
1×10^{17}	5,4 ns	8,9 ns	3,5 ns
1×10^{18} *2	54,0 ns	89,0 ns	35,0 ns

*1 : typical value for night time of solar minimum
*2 : typical value for day time of solar maximum

As the propagation delay depends on frequency, TEC can be estimated using a dual frequency method:

$$TEC = \frac{T_{ion}(L_2) - T_{ion}(L_1)}{\alpha} \left[\frac{1}{L_2^2} - \frac{1}{L_1^2}\right]^{-1} \qquad (2)$$

Recently, for purpose of precise geodesy, several different type of interferometric equipments using GPS dual frequency have been realized [9-12]. These are codeless devices with separate reconstructions of L_1 and L_2 signals.

In the same way, GTR2 does not demodulate GPS P-code. The measurement is based upon the cross-correlation between P-codes carried by L_1 and L_2: because L_1 P-code and L_2 P-code are exactly identical and emitted in phase, the cross-correlation of the received P-codes gives access to the difference of L_1 delay and L_2 delay, as shown on figure 1.

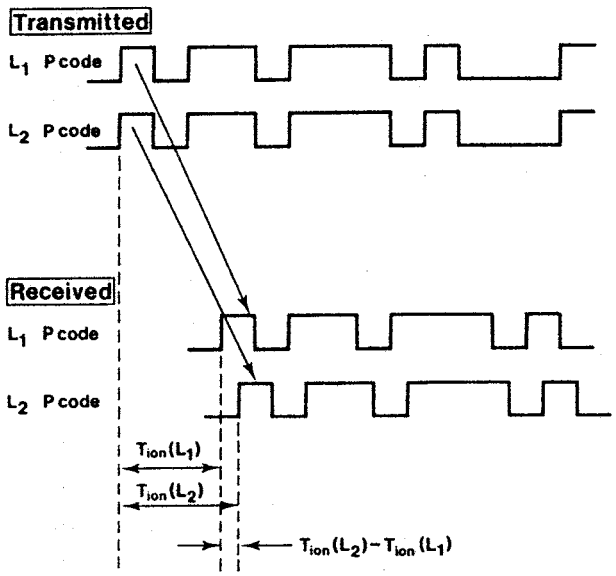


Fig. 1.- Transmitted and received L_1 and L_2 P-codes

TEC is then measured along the line of sight of the GPS satellites which allows, from equation (1), the estimation of the quantity $T_{ion}(L_1)$ and then ionospheric compensation on the real signal path.

GTR2 is very simple, with no need of precise measurement of time interval or high accuracy frequency sources such as atomic frequency standards. To be used for time comparisons by GPS common view method [1], GTR2 needs to be associated to a traditional one frequency GPS receiver.

2. GTR2 RESULTS

GTR2 is located at the Bureau International des Poids et Mesures, Sèvres near Paris, France.

GTR2 operates on a regular basis since October 1988. Its program includes all observable satellites from BIPM but priority is given to the scheduled common views between Paris and Washington (about 25 a day).

GTR2 works with 4 minute sequences: about one minute to point its directive antenna (gain 10 dBi for L_1 and L_2 frequencies) and about 3 minutes to perform the observation. Inside these 3 minutes, the measurement itself needs 3 seconds and 3 other seconds for pretreatment. The resulting TEC is an averaged value over a three minutes sequence, provided with an uncertainty of 2×10^{16} m^{-2} which corresponds to an uncertainty of 1 ns for L_1 ionospheric delay.

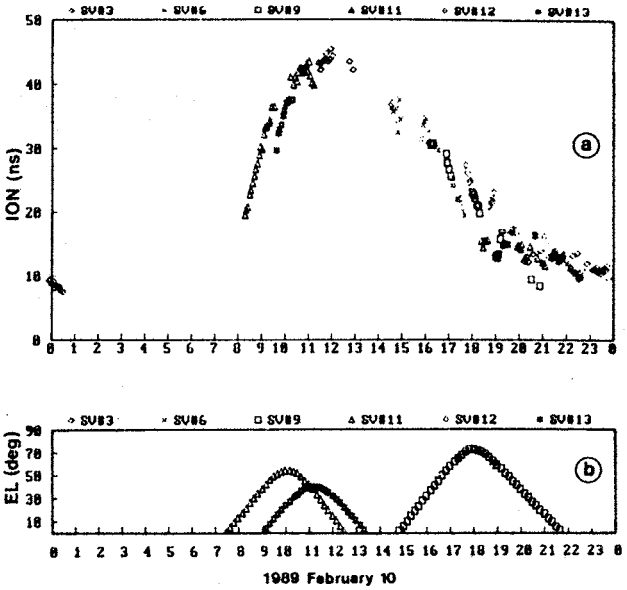


Fig. 2.- Vertical ionospheric delays calculated from GTR2 results obtained on 1889 February 10:

a - Diurnal results
b - Corresponding elevations for satellites 11, 13 and 9.

Figure 2 shows L_1 ionospheric delays obtained from GTR2 on 1989 February 10. The measured values are here converted to vertical estimations using a simple geometric expression based on the assumption that ionosphere is a spherical shell, lying from 200 to 450 km altitude and with uniform electron density. The diurnal effect is evident but discrepancies between values obtained from different satellites are observable. This is mainly due to the vertical conversion model, which is based on too simple hypothesis and which brings an artificial elevation effect, especially degradating for values issued from observations occuring to very small elevations. This can help to interpret discrepancies as between values deduced from satellites 11 and 13 during the morning (about 10 h UTC). Furthermore, GTR2 measurement noise increases for high TEC values occuring for small satellite elevations (< 25°) and the corresponding values must be deleted. For instance, it is the case for measured delays deduced from satellite 9 at about 20° elevation in evening (about 21 h UTC).

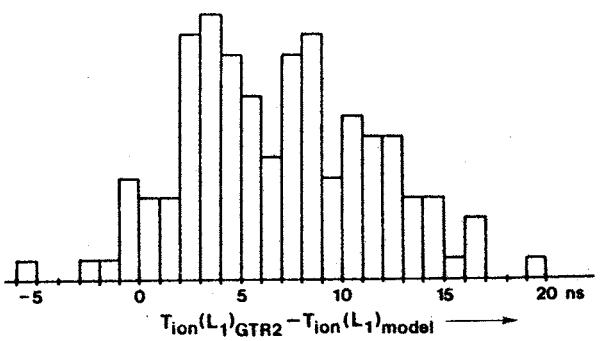


Fig. 3.- Histogram ot the values $T_{ion}(L_1)_{GTR_2} - T_{ion}(L_2)_{model}$, obtained from 130 trackings for a five days period (no vertical conversion is applied)
$T_{ion}(L_1)_{GTR2}$ are GTR2 measurement
$T_{ion}(L_2)_{model}$ are the ionospheric compensation model estimations

GTR2 measurements can be compared to values derived from ionosphere compensation model. These values are accessible from a one frequency GPS receiver located in Paris Observatory and correspond of course to the same scheduled trackings. Figure 3 comes from a five days analysis (1988 October 5 to 9) where about 130 trackings were available. Measured values are very often larger than model ones, the disagreement can even reach 20 ns. The discrepancy exceeds 10 ns for 28 % of the values and 5 ns for 62 % of the values.

Figures 4a and 4b also present measured ionospheric delays compared with model values. These measurements are performed by USNO on two different sites, Florida and Washington, with dual frequency receivers demodulating GPS signal P-code. Raw data obtained with a basic sequence of observation of 6 s and for a single satellite tracked during several consecutive hours, are reported on these figures. On the opposite of what is observed for Paris, model values appear to be larger than measured ones for Florida. The conclusion is not so clear for Washington where the disagreement seems rather linked to the satellite elevation.

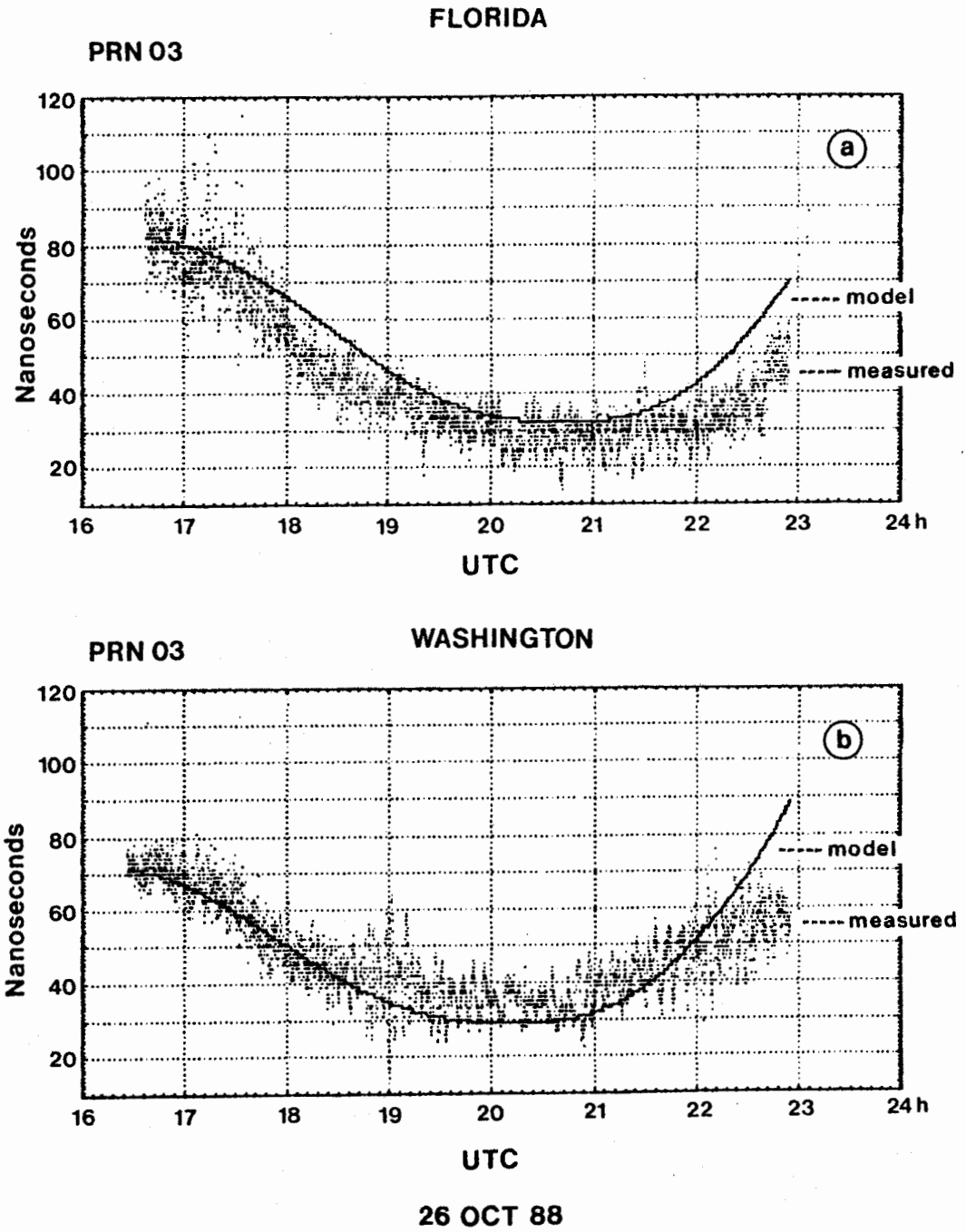
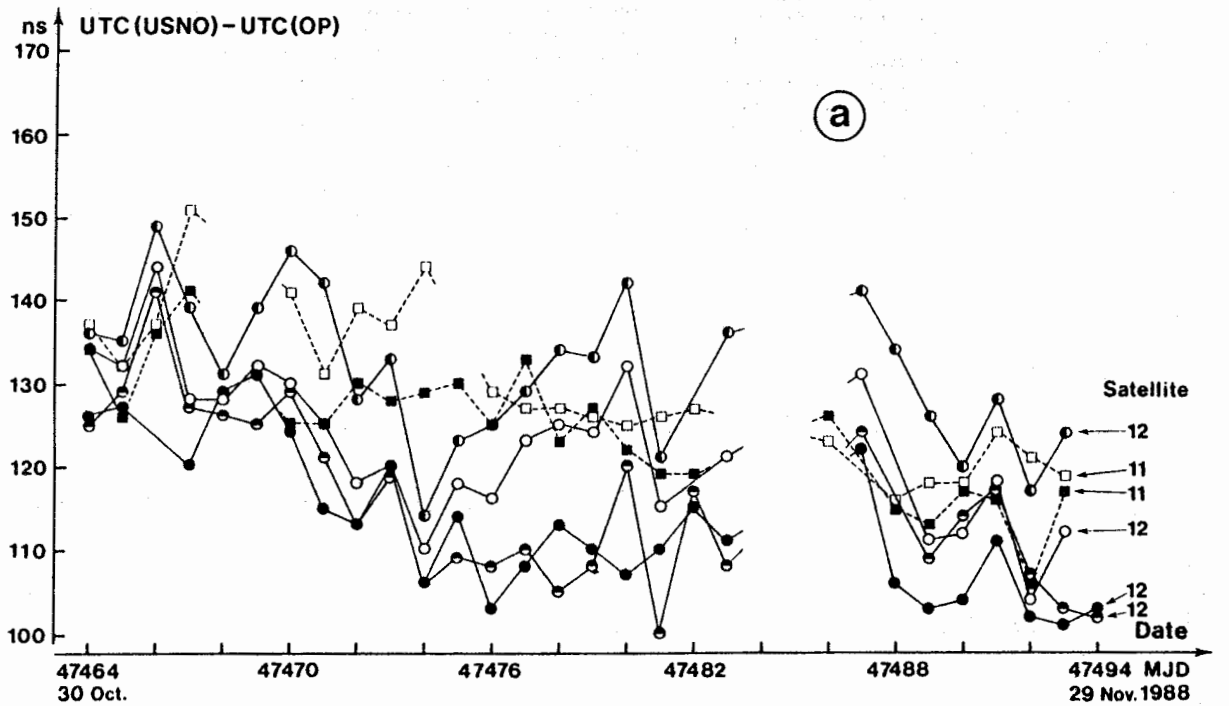


Fig. 4.- Comparison between measured ionospheric delays obtained from USNO decoding GPS receivers and model values for two sites: a - Florida
b - Washington



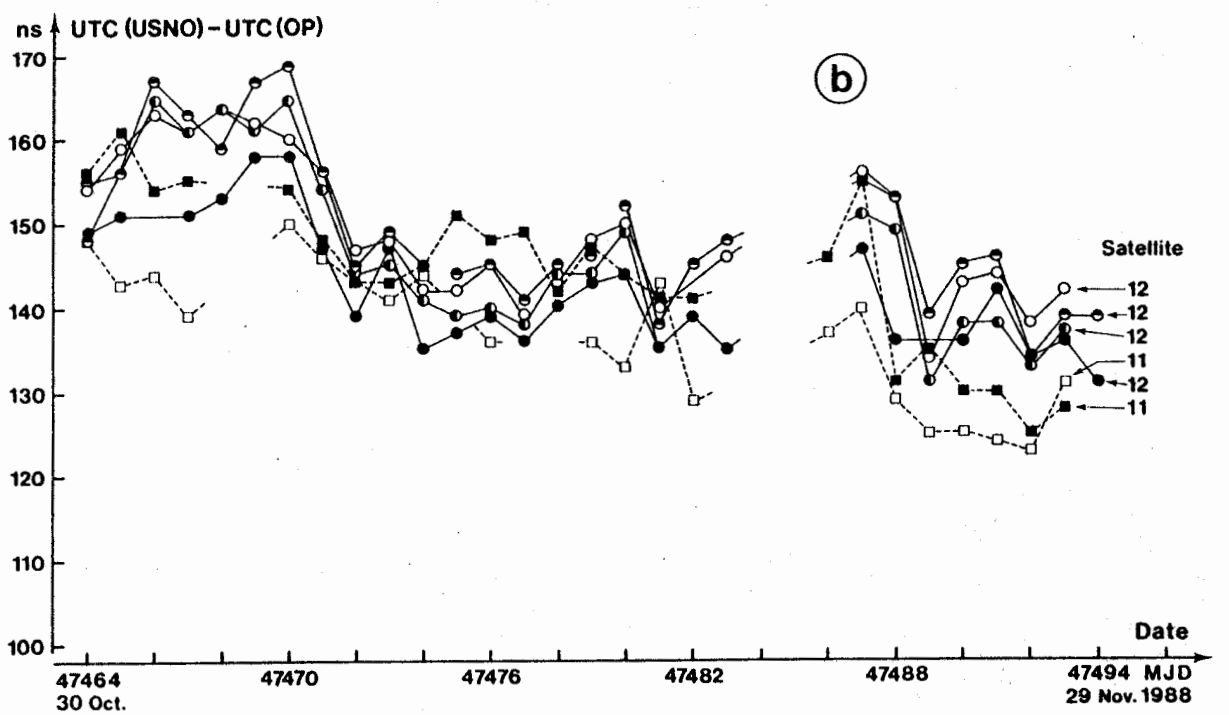


Fig. 5.- Time comparison between Paris Observatory (Paris, France) and United States Naval Observatory (Washington D.C., USA) for a thirty days period (November 1988). The two branches of the satellite simultaneous trackings are corrected with

a - modelized ionospheric values
b - measured ionospheric values

The hours indicated for satellites 11 and 12 correspond to the scheduled common views on MJD 47464 (1988 October 30):

- sat 12, 1h38
- sat 12, 1h51
- sat 12, 2h04
- sat 12, 2h38
- sat 11, 17h19
- sat 11, 18h04

3. APPLICATION TO TIME COMPARISON BETWEEN EUROPE AND USA

Measured ionospheric delays obtained from GTR2 and from USNO dual frequency receiver can be used to improve the daily time comparisons between Paris (Paris Observatory) and Washington (USNO).

A preliminary study has been performed for a thirty days period, from October 30 to November 29 1988. It consists in a comparison of the daily USNO - OP time link where the two branches of simultaneous satellite trackings are both corrected with either modelized or measured ionospheric delays. Data were available for only 6 of the 25 daily scheduled common views and correspond to two different satellites: PRN 11 (2 common views spread over 1 hour) and PRN 12 (4 common views spread over 2 hours). The results are shown on figures 5a and 5b where it appears very clearly that USNO - OP time transfer is performed with a much higher precision when measured ionospheric effects are involved.

The remaining errors are satellite ephemeris errors, which are responsible for daily scattering of UTC(USNO) - UTC(OP) values, and also antenna differential coordinates errors, which are responsible for the global shift of UTC(USNO) - UTC(OP) values (the dual frequency and single frequency receivers of USNO Washington corresponding to two different antenna localizations).

CONCLUSION

A prototype of a codeless dual frequency GPS receiver operates on a regular basis at BIPM since October 1988. This equipment is able to provide measured ionospheric delays along GPS satellite lines of sight with an uncertainty of 1 ns. When associated with USNO similar results, it brings a very interesting gain of precision for transatlantic time links.

Furthermore this original device has a very simple structure and appears to be an efficient complement to traditional one frequency GPS receivers. At last, though it exists other methods to measure ionospheric effects, as Faraday rotation or dual frequency transmitted by Navy Navigation Satellite System (NNSS), the Global Positioning System is becoming more and more an important ionospheric monitoring system. Then our codeless receiver belongs to a very promising generation of equipments which could widely overstep geodesy and time purposes.

REFERENCES

[1] Allan D.W. and Weiss M.A., "Accurate time and frequency transfer during common view of a GPS satellite", Proc. 34th Annual Symp. on Frequency Control, 1980.

[2] Lewandowski W., Weiss M.A. and Davis D., 1986, "A calibration of GPS equipment at time and frequency standard laboratories in the USA and Europe", Proc. 18th PTTI meeting p. 265-279 [also published in Metrologia, 24, p. 184-186, 1987].

[3] Guinot B. and Lewandowski W., "Use of the GPS time transfer at the Bureau International des Poids et Mesures", Proc. 19th PTTI meeting p. 3-12, 1987.

[4] Guinot B. and Lewandowski W., "Nanosecond time comparisons in Europe using the GPS", Actes, 2nd European Frequency and Time Forum, p. 187-193, 1988.

[5] Lewandowski W., Chiu M., Douglas R.J., Klepczynski W.J., Strange W. and Weiss M.A., "The positioning of GPS antennae in time-keeping laboratories of north America", to be published in the Proc. 43th Annual Symp. on Frequency Control, 1989.

[6] Klobuchar J.A., "Ionospheric Correction for the single frequency users of the Global Positioning System", IEEE Trans. NTS. 1982.

[7] Imae M., Lewandowski W., Thomas C. and Miki C., "A dual frequency GPS receiver measuring ionospheric effects without code demodulation and its application to time comparison", Proc. 20th PTTI meeting, p. 77-86, 1988.

[8] Klepczynski W.J., private exchange of data, January 1989.

[9] Stansell T.A. Jr., Chamberlain S.M. and Brunner F.K., "The first Wild-Magnavox GPS Satellite Surveying Equipment: WM-101", Proc. 1st International Symposium on Precise Positioning with the Global Positioning System, 1985.

[10] Ladd J.W., Counselman C.C. and Gourevitch S.A., "The macrometer Dual-Band Interferometric Surveyor", Proc. of the 1st International Symposium on Precise Positioning with the Global Positioning System, 1985.

[11] MacDoran P.F., Miller R.B., Buennagel L.A. and Whitcomb J.H., "Codeless Systems for Positioning with NAVSTAR-GPS", Proc. 1st International Symposium on Precise Positioning with the Global Positioning System, 1985.

[12] Sujimoto Y., Kurihara N., Kiuchi H., Kanedo, A., Kawasa S., Sawada T., Shirado T. and Saburi Y. "Development of GPS Positioning System PRESTAR", Proc. 1988 Conference on Precision Electromagnetic Measurements, p. 101-102.

Acknowledgement

The authors wish to thank Dr. Klepczynski from USNO for helpful discussions and collaboration in sending measured ionospheric delay data.

EXPERIENCE WITH TWO COLLOCATED C/A-CODE GPS-RECEIVERS OF DIFFERENT TYPES

D. Kirchner[1], H. Ressler[2] and S. Fassl[2]

[1] [2] Department of Communications and Wave Propagation, Technical University Graz (TUG)
Space Research Institute, Austrian Academy of Sciences, Graz

ABSTRACT

Over a period of about half a year two collocated C/A-Code GPS receivers of different types were used at the Lustbühel Observatory of the Technical University Graz (TUG) in order to investigate problems encountered with one of the receivers and to study the behaviour of the differential delay. After a short comparison of some essential characteristics of the receivers the results of the delay comparison and ensuing findings are given followed by a more general discussion of the problem of receiver delay calibration.

INTRODUCTION

Since late 1982 at the Lustbühel Observatory of the Technical University Graz (TUG) as on-site GPS-receiver a TTS-502 (Serial Number 004) manufactured by Stanford Telecommunications, Inc. (STI) has been used /1/. Several hardware and software updates were issued by STI, but none to adapt the receiver to the change-over from the WGS-72 to the WGS-84 geodetic system, which became valid for GPS on January 22, 1987, and thus the TTS-502 still uses WGS-72 ellipsoid parameters. From May 5, 1988, a second receiver designed and built at the National Bureau of Standards (NBS), now National Institute of Standards and Technology (NIST), is being operated at TUG. This receiver is designated as NBS#03 and is adapted for the use with the WGS-84 geodetic system.

Both receivers are single frequency receivers making use of the C/A-code only. In the following some user-relevant differences between the receivers and their operation will be given. More details may be found in the respective operating manuals and related publications /2,3/. The NBS-receiver requires connection to the outdoor-unit (omnidirectional antenna, preamplifier and first downconverter) by two coax cables (RG-58) one of which leads the amplified and down-converted signal to the receiver and the other supplies the required LO-frequency to the down-converter. The receiver unit consists of two pieces (one of which contains a CRT to display messages and data) and a detachable keyboard to programme the receiver. Two serial interfaces (RS-232-C) are available which allow the receiver to be connected to a modem and a printer. The modem port is bi-directional which allows the measurement results stored in the internal memory to be read out in the agreed-upon format (see Table 1) as used in the common-view mode of operation /4/. The receiver performs pseudo-range measurements every second and smoothes the data over 15 seconds. For a common-view measurement of 13 minutes duration, called a track, a linear regression over 52 such data points is computed and the result related to the beginning of the track is stored together with other relevant information such as elevation, azimuth and ionospheric correction, but related to the end of the track. The measurement resolution is 0.1 ns. Via the printer port apart from other information, the 1 s measurements (pseudo-ranges) or the 15 s measurements (smoothed pseudo-ranges corrected for the range to the satellite, the Sagnac effect and the ionospheric and tropospheric and receiver delays) together with the ephemeris data, the ionospheric data and the satellite clock data are available. The selection has to be done via the keyboard. After a power interrupt the receiver automatically resumes operation.

The STI-receiver requires only one cable (RG-213) to the outdoor unit (omnidirectional antenna and preamplifier) to lead the amplified signal to the receiver. The receiver provides a bidirectional serial interface (RS-232-C) to connect it to a terminal and the so-called auxiliary port to connect it to a printer or computer. For operation it has to be connected to a terminal. The terminal is used for programming the receiver and to display the data. Via the auxiliary port all measurement data together with additional information as selected via the terminal are available. The receiver performs one pseudo-range measurement every second and smoothes the data over 6 seconds /5/. The measurement resolution is about 50 ns (1/20 C/A-chip) and the resolution of the computed data 1 ns. Every 6 seconds these data (called raw data) or smoothed data (the smoothing period and the kind of smoothing can be selected) together with other information (elevation, azimuth, doppler frequency, satellite status etc.) are output to the CRT of the terminal. The 6 second pseudo-ranges, all computed data, and information contained in the navigation message such as ephemeris data, ionospheric data etc., grouped in so-called data categories may be output via the auxiliary port according to the choice of the user. To obtain the measurement results for a common-view observation one has to collect the appropriate data comming out from the auxiliary port every 6 seconds. Then one has to compute the results from a linear regression over 131 data points and to store the data according to the common-view format. After a power interrupt no operational parameters are lost, but the receiver needs user intervention to resume tracking.

Both receivers can be used for positioning. Both need external 5 MHz and of course a 1 PPS referenced to the local time scale.

Over the years problems appeared with the STI-receiver which manifested themselves in missing 6 second data points or the inability of the receiver to track a satellite at all. By means of the second receiver, generously on loan from NIST, it was possible to separate these errors from general problems caused by GPS-satellites or possibly by the receiver set-up. In addition the collocated operation of the receivers over nearly 200 days (mid November 1988 the STI-receiver was sent back to STI for evaluation and repair) also provided the opportunity to study the differential delay of the two receivers, which is an essential parameter for time comparisons and therefore of special interest. Before this extended comparison the delay of the TUG on-site receiver was compared to other receivers during two international campaigns, but only for periods not more than a few days /6,7/.

MEASUREMENTS AND RESULTS

Fig.1 shows the set-up of the receivers. The antenna of the NBS-receiver was mounted on a separate tripod

at the same height as the one of the STI-receiver at a distance of 1.15 m to the south. The coordinates of the antenna of the STI-receiver were available from Doppler measurements /8/ converted to WGS-72 and later to WGS-84 coordinates /9,10,11/. They were input into the NBS receiver in form of geodetic coordinates (corrected for the antenna displacement in NS-direction) and into the STI-receiver in form of spatial coordinates thus avoiding the problem of internal coordinate conversion with wrong ellipsoid parameters (semimajor axis and flattening) which would cause an error of about 2 m. The use of WGS-72 coordinates instead of the WGS-84 ones would lead to an error of about 12 m. Both receivers were fed by the same clock via distribution amplifiers for the 5 MHz and the 1 PPS (the 1 PPS's had a rise time of less than 5 ns and agreed to better than 300 ps) and were connected to a computer to store both the 15 second data from the printer port of the NBS-receiver and the 6 second data provided by the auxiliary port of the STI-receiver. The common-view data computed by the NBS-receiver, and available via the modem port, were also read by the computer. Both receivers were powered by an uninterruptable power supply.

Common-view measurements with the two collocated receivers were performed from May 5, 1988 to November 9, 1988 (net measurement period 187 days) according to the common-view schedule published by the Bureau des Poids et Mesures (BIPM) valid from December 15, 1987, and updated on June 22, 1988. Due to the above mentioned problems with the STI receiver and to external reasons only data of 4321 tracks from a theoretical number of 4523 tracks were available from this receiver and only 2873 were common view tracks in the true sense, which means same begin time and full number of points. The NBS receiver failed to track 64 times whereof 57 failures were caused by satellite problems and other external reasons and stored data of 45 incomplete tracks (8 caused by external reasons). Furthermore, the printer port was not connected prior to June 15 to the computer, thus 15 seconds data and ephemeris data were only stored from this date on, whereby for some tracks the ephemeris data were not in the proper format or incomplete reducing the number of tracks with all data available for both receivers to a mere 2180. Figs. 2 and 3 show the results of the 2873 common view tracks. Each point presents the common-view result obtained by the STI-receiver minus that of the NBS-receiver. In Fig. 2 the results of the linear regressions for each track, which are used for the computation of the differences, are referred to the beginning of the track (BOT) and in Fig. 3 to the middle of the track (MOT). The solid lines are the results of moving averages using a smoothing period of 10 days (SM 10D). In both figures the characteristics of the linear regressions through the 2873 data points are given. From the different scatter of the data and from the standard deviations of the linear regressions one can see that the use of the middle of the track as reference yields smoother results. Figs. 4 and 5 show the results of single tracks for both receivers after employing a moving average with a smoothing period of 3 minutes (SM 3M). For these figures tracks (given in Table 1) with low RMS (SV#11, Cl 68) and high RMS (SV#13, Cl 84) for both receivers where chosen to show the behaviour of the results obtained by the two receivers for the same tracks in more detail. Table 1 contains common-view data in standard format for an arbitrary date (October 17, 1988), but with complete data sets obtained by STI#004 and NBS#03. In order to compare the noise of both receivers two additional columns are added to the NBS data listing the standard deviation of the linear regression through the 131 data points of the STI-receiver ($RMS^{1)}$) and the standard deviation ($RMS^{2)}$) of a linear regression through the mean values of three adjacent data points of the STI-receiver to better adjust the smoothing of the originally 131 data points, where each points is smoothed over 6 second, to the 52 data points of the NBS-receiver, where each point is smoothed over 15 seconds. From these and other data one can see that the measurement uncertainty of the STI#004-receiver is slightly higher than that of the NBS#03-receiver.

Plotting the data of Fig. 3 separated into tracks of the same class of a single satellite reveals systematic offsets between different classes of a single satellite differing for different satellites. Figs. 6 and 7 show this behaviour for satellite SV#09 and SV#11, respectively. A conceivable cause for this behaviour is the use of different geodetic systems in both receivers which result in the computation of different ranges to the satellites according to the different ellipsoid parameters (angular velocity and gravitational constant) used for the computation of the satellite positions in the Earth fixed coordinate system /12,13/. Because the ranges computed by the receivers are only available from the STI, but not from the NBS-receiver, the ranges in both geodetic systems were computed using the ephemeris data provided by the NBS-receiver together with the 15 s data. Then the results given in Fig. 3 were corrected for the range differences (WGS-84 - WGS-72), thereby reducing the number of data points according to the availability of the ephemeris data as mentioned above. The corrected data are shown in Fig. 8. In Figs. 9 and 10 selected data corresponding to the Figs. 6 and 7 are given.

In Table 2 the results of linear regressions for the reduced data set (RD) are given for all tracks of all satellites, for all tracks of single satellites and for tracks of single satellites of the same class before and after applying the correction for the range differences. For all satellites except SV#11 the correction leads to a significant reduction of the offsets and also of the standard deviation of the residuals from a linear regression. For satellite SV#11 the offsets become larger after applying the correction as can be seen from Figs. 7 and 10, respectively. The 10 day smoothed data (solid line in Fig. 8) show variations between a minimum of 1.3 ns and a maximum of 3.9 ns and the one day smoothed data given in Fig. 11 show variations between a minimum of 0.5 ns and a maximum of 5.4 ns. Using the results of Fig. 11 one should bear in mind that due to the temporary malfunction of the STI-receiver the number of tracks per day is not constant and therefore for both receivers working properly the daily variations can be smaller. Using the data of Fig. 3 the square-root of the Allan variance was computed for all tracks separately (same satellite and same class number) to give the stability of an optimum common view operation (collocated receivers) of both receivers in question. In Fig. 12 the corresponding maximum, minimum and mean value of the two-sample standard deviation are given for sampling times of up to 16 days.

DISCUSSION

In order to discuss the differential receiver delay and the delay calibration the way time is provided by the Global Positioning system is of primary interest. The difference between the local reference time and GPS-time is given by the following relationship /2/:

$$\Delta TU = PR - (R + \tau) + \Delta TSV \qquad (1)$$

PR pseudo-range (measured by the user)

R true time range (computed by the user from the station and satellite position and correction for the earth rotation)

τ propagation delay of the signal consisting of the ionospheric delay τ_I, the tropospheric delay τ_T and the receiver delay τ_R (using a single frequency receiver the ionospheric delay is computed by the user from parameters transmitted in the navigation message, the tropospheric delay by means of a model and the receiver delay should be calibrated)

ΔTSV space vehicle clock error (computed by the user from parameters transmitted in the navigation message)

According to Equ. 1 for two receivers R1 and R2 the difference of the measurement results is given by the following equation:

$$\Delta TU_{R1} - \Delta TU_{R2} = [(PR_{R1} - \tau_{R,R1}) - (PR_{R2} - \tau_{R,R2})] - (R_{R1} - R_{R2}) - (\tau_{J,R1} - \tau_{J,R2}) - (\tau_{T,R1} - \tau_{T,R2}) + (\Delta TSV_{R1} - \Delta TSV_{R2}) \quad (2)$$

Assuming that the receivers are connected to the same reference and that they use correct coordinates of the antennae, correct algorithms and correctly calibrated receiver delays, the difference given by the above equation should be zero (within the measurement uncertainty). A difference unequal to zero could be caused by different systematic effects on the PR measurements and/or wrong coordinates of the antennae and/or erroneous delay calibrations and/or different algorithms. If only the receiver delays are incorrectly calibrated and all other contributions are zero the difference given by by Equ. 2 is the differential delay by which the measurements have to be corrected to yield consistent results. But often these assumptions are not valid and there are contributions, which result in a differential delay depending on the satellite constellation used.

For example in our case the use of different geodetic systems resulted in different R_R computations and therefore in a differential delay depending on the satellite position. For extremely low elevation angles (less than 10°) a contribution can also be due to the different models for the tropospheric delay used in both receivers resulting in different τ_T computations. Slight offsets can also be caused by the different computational resolutions (0.1 ns and 1 ns) used in the receivers. But also the pseudo-range measurements can be influenced to a different degree by various effects such as Doppler frequency, signal strength, multipath, etc. causing differential delays depending on the satellite constellation.

If a GPS receiver makes all the data available which are necessary to evaluate the differences given in Equ. 2 the above mentioned effects causing the establishment of erroneous differential delays can be evaluated by the use of calibrators and receiver comparison by collocation of receivers and appropriate corrections could be applied. One could also use a unified software running on a computer which only requires the pseudorange measurements, the coordinates of the antenna and parts of the navigation message or satellite ephemeris and ionospheric data supplied by other means as an input, a method which is widely used within the geodetic community.

With the present GPS satellites (preoperational satellites, no intentional signal degradation) and user equipment commercially available, accuracies of the order of 10 to 20 ns seem to be no problem but the above considerations are important if one is interested in the exploitation of the full accuracy potential for high precision time transfer provided by the Global Positioning System.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the loan of a GPS receiver by the National Institute of Standards and Technology. The work was supported by various grants of the Austrian Council for Scientific Research, the Austrian Academy of Sciences and the Jubilee Fund of the Austrian National Bank.

REFERENCES

/1/ Kirchner, D.:
Precision Timekeeping at the Observatory Lustbühel, Graz, Austria, Proc. ASFC, 37th, pp. 67-77, 1983.

/2/ Dierendonck, A.J., Q.D. Hua, J.R. McLean and A.R. Denz:
Time Transfer Using Navstar/GPS, Proc. PTTI Conf., 13th, pp. 389-419, 1981.

/3/ Davis, D.D., M. Weiss, A. Clemens and D.W. Allan:
Construction and Performance Characteristics of a Prototype NBS/GPS Receiver, Proc. ASFC, 35th, 1981.

/4/ Allan, D.W. and M.A. Weiss:
Accurate Time and Frequency Transfer during Common-View of a GPS Satellite, Proc. ASFC, 34th, pp. 334-347, 1980.

/5/ Melton, W.:
Personal Communication, Oct. 12, 1981.

/6/ Buisson, J.A., O.J. Oaks and H. Warren:
Absolute Calibration and Precise Positioning Between Major European Time Observatories and the U.S. Naval Observatory Using GPS, 1st Int. Symp. on Precision Positioning with the Global Positioning System, Vol. 1, pp. 433-445, 1985.

/7/ Weiss, M.A. and W. Lewandowski:
A Calibration of GPS Equipment at Timing Laboratories in Europe, Proc. PTTI Conf., 18th, pp. 265-279, 1986.

/8/ Schlüter, W. and L. Amberg:
Final Results of Doppler Point Positioning at Several Timekeeping Laboratories in Europe, Nachrichten aus dem Karten- und Vermessungswesen, Reihe III: Übersetzungen, 40, pp. 99-107, Verlag IFAG, Frankfurt a. Main, 1983.

/9/ Meade, B.K.:
NWL-10F versus WGS-72 Doppler Results and Broadcast versus Precise Ephemeris Coordinates, Proc. of the 3rd Internat. Geodetic Symp. on Doppler Positioning, New Mexico, Feb. 8-12, pp. 151-168, 1982.

/10/ Guinot, B.:
Personal communication, Dec. 12, 1986.

/11/ Daily Time Differences and Phase Values, Series 4:
U.S. Naval Observatory, No 1040, Jan. 7, 1987.

/12/ Decker, B.L.:
World Geodetic System 1984, Proc of the 4th Int. Symp. on Satellite Positioning, Vol 1, Austin, Texas, Apr. 28 - May 2, pp. 69-92, 1986.

/13/ Interface Control Document:
(ICD)-GPS-200 B, Nov. 30, 1987.

STI

SV	CL	MJD'	HH	MM	SS	TRCK SEC	AG HR	EL DG	AZM DEG	ION NS	TUG-SAT NS	SLOPE PS/S	TUG-GPS NS	SLOPE PS/S	RMS NS
6	58	7451	00	14	00	780	11	74	339	5	335926	7	-5852	-11	14
12	08	7451	00	58	00	780	5	14	260	12	-744933	-10	-5839	-6	11
9	08	7451	01	14	00	780	9	54	288	5	117368	0	-5845	-9	18
6	40	7451	01	30	00	780	12	49	50	6	335993	20	-5866	2	15
6	8C	7451	01	56	00	780	12	40	63	7	336025	23	-5863	6	15
9	58	7451	02	12	00	780	6	67	348	5	117382	26	-5857	18	15
12	74	7451	02	30	00	780	8	45	298	6	-744977	7	-5858	13	14
11	A0	7451	02	54	00	780	15	46	104	6	190939	1	-5852	1	13
9	48	7451	03	14	00	780	7	60	63	5	117417	9	-5851	2	16
12	58	7451	03	30	00	780	10	66	337	5	-745003	4	-5861	11	17
13	4C	7451	03	46	00	780	7	51	140	6	-376823	-1	-5856	2	14
11	84	7451	04	02	00	780	18	47	54	6	190945	5	-5849	4	14
3	49	7451	04	22	00	780	7	33	179	9	-397466	3	-5860	6	13
9	4C	7451	04	38	00	780	8	31	112	15	117443	-2	-5854	-6	11
12	48	7451	05	36	00	780	14	50	127	15	-745062	0	-5861	8	13
13	84	7451	06	02	00	780	12	38	41	21	-376859	17	-5871	20	15
11	18	7451	16	22	00	780	17	18	329	64	191142	-4	-5858	-6	13
13	30	7451	17	54	00	780	10	18	331	50	-376966	8	-5864	12	13
11	74	7451	18	10	00	780	7	41	281	32	191170	2	-5848	-4	11
11	68	7451	19	02	00	780	8	28	253	37	191190	7	-5848	-1	10
3	18	7451	19	55	00	780	9	18	323	31	-397434	4	-5859	1	14
6	4C	7451	21	56	00	780	4	29	213	9	337186	21	-5863	6	10
9	4D	7451	23	15	00	780	6	12	234	13	117686	3	-5869	-3	13
9	54	7451	23	50	00	780	6	26	246	9	117700	4	-5869	-5	12

NBS

SV	CL	MJD'	HH	MM	SS	TRCK SEC	AG HR	EL DG	AZM DEG	ION NS	TUG-SAT NS	SLOPE PS/S	TUG-GPS NS	SLOPE PS/S	RMS NS
6	58	7451	00	14	00	780	6	71	360	5	335918	18	-5860	0	6
12	08	7451	00	58	00	780	3	18	264	11	-744929	-14	-5835	-10	7
9	08	7451	01	14	00	780	5	58	298	6	117366	8	-5848	0	8
6	40	7451	01	30	00	780	6	45	57	7	335994	16	-5865	-1	8
6	8C	7451	01	56	00	780	7	36	69	8	336022	14	-5866	-2	10
9	58	7451	02	12	00	780	3	68	5	5	117386	8	-5853	0	6
12	74	7451	02	30	00	780	4	50	305	6	-744976	-3	-5857	2	8
11	A0	7451	02	54	00	780	9	49	95	6	190944	-4	-5847	-4	6
9	48	7451	03	14	00	780	3	56	73	6	117410	10	-5859	3	7
12	58	7451	03	30	00	780	5	69	350	5	-744999	-11	-5856	-4	7
13	4C	7451	03	46	00	780	3	57	133	6	-376826	5	-5859	7	7
11	84	7451	04	02	00	780	10	43	48	7	190945	-8	-5850	-10	9
3	49	7451	04	22	00	780	3	40	179	10	-397463	-4	-5856	-2	7
9	4C	7451	04	38	00	780	4	26	117	21	117433	12	-5865	8	10
12	48	7451	05	36	00	780	7	44	133	19	-745068	-3	-5866	4	6
13	84	7451	06	02	00	780	6	32	41	26	-376858	-2	-5869	0	11
11	18	7451	16	22	00	780	9	23	327	56	191134	8	-5866	6	10
13	30	7451	17	54	00	780	6	22	326	45	-376966	0	-5863	2	10
11	74	7451	18	10	00	780	3	39	273	33	191168	0	-5850	-6	8
11	68	7451	19	02	00	780	4	23	248	40	191189	8	-5849	1	7
3	18	7451	19	55	00	780	6	18	316	30	-397434	-5	-5857	-8	9
6	4C	7451	21	56	00	780	3	35	216	8	337195	10	-5855	-4	7
9	4D	7451	23	15	00	780	3	17	238	12	117689	10	-5866	4	9
9	54	7451	23	50	00	780	4	30	251	9	117701	10	-5868	1	6

STI RMS[1] NS	STI RMS[2] NS
14	10
11	7
18	12
15	9
15	10
15	10
14	10
13	9
16	9
17	12
14	7
14	9
13	9
11	8
13	8
15	11
13	9
13	9
11	6
10	6
14	10
10	6
13	8
12	7

Table 1. Common-view data in standard format obtained by the STI- and NBS-receivers for Oct. 17, 1988. ($RMS^{1)}$ and $RMS^{2)}$ are the standard deviations of the results of the STI-receiver for different smoothing of the raw-data).

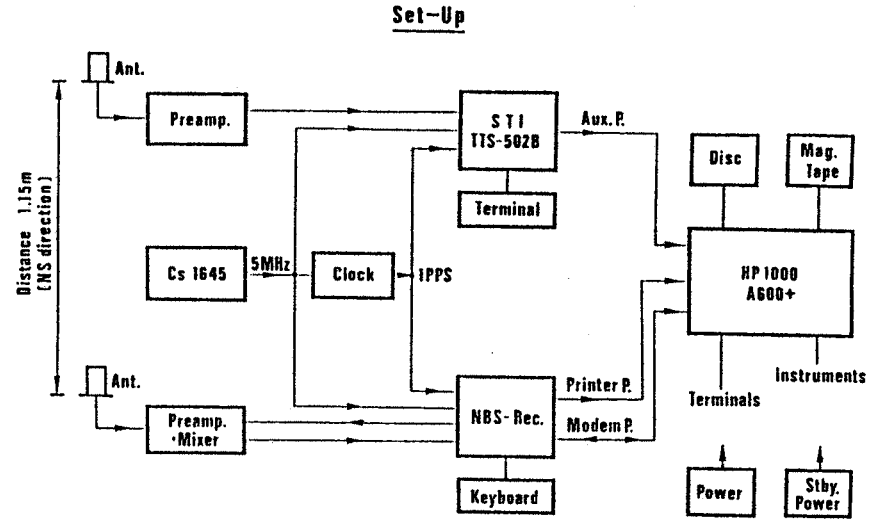


Figure 1. Schematic of the set-up of the GPS-receivers.

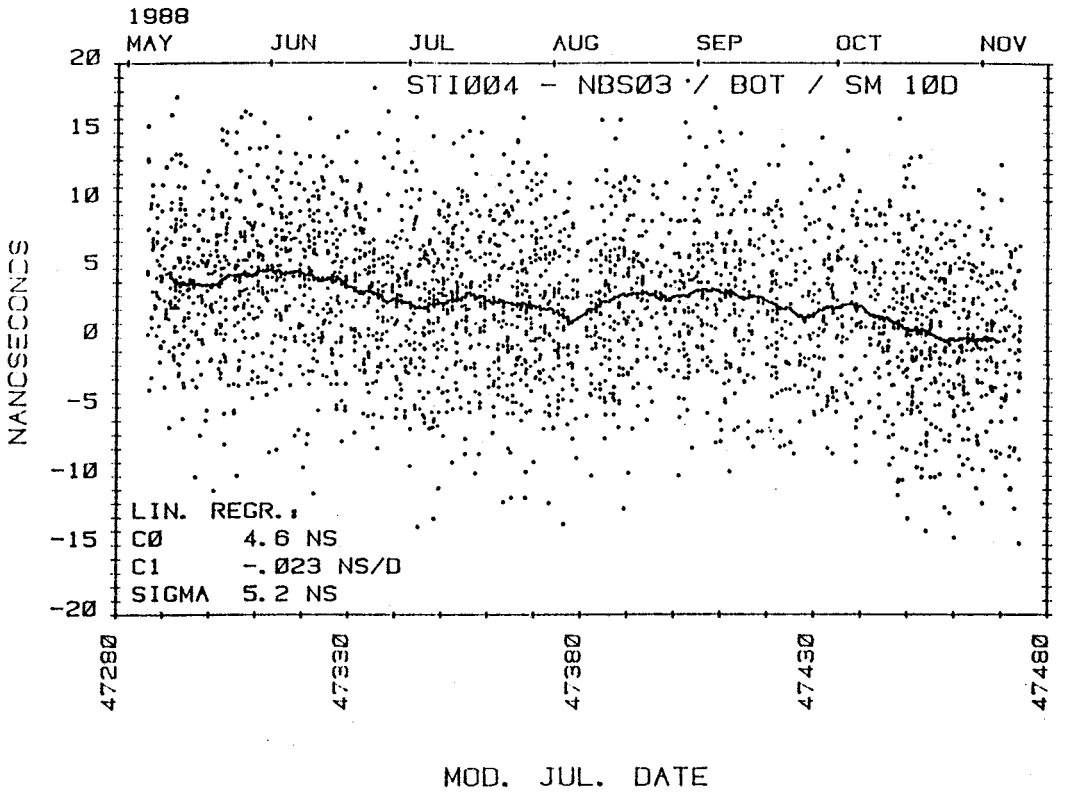


Figure 2. Common-view results of STI-receiver minus NBS-receiver referred to the beginning of the tracks (points: results of single tracks, solid line: smoothed over 10 days).

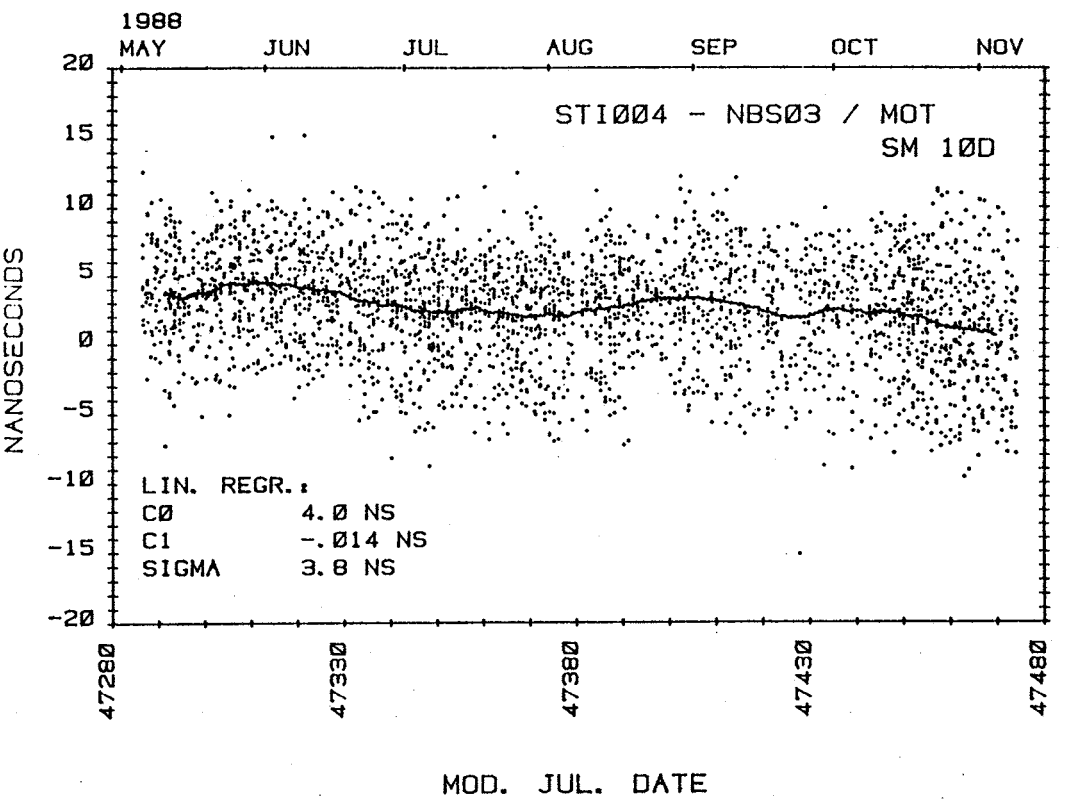


Figure 3. Same as Fig. 2, but referred to the middle of the tracks.

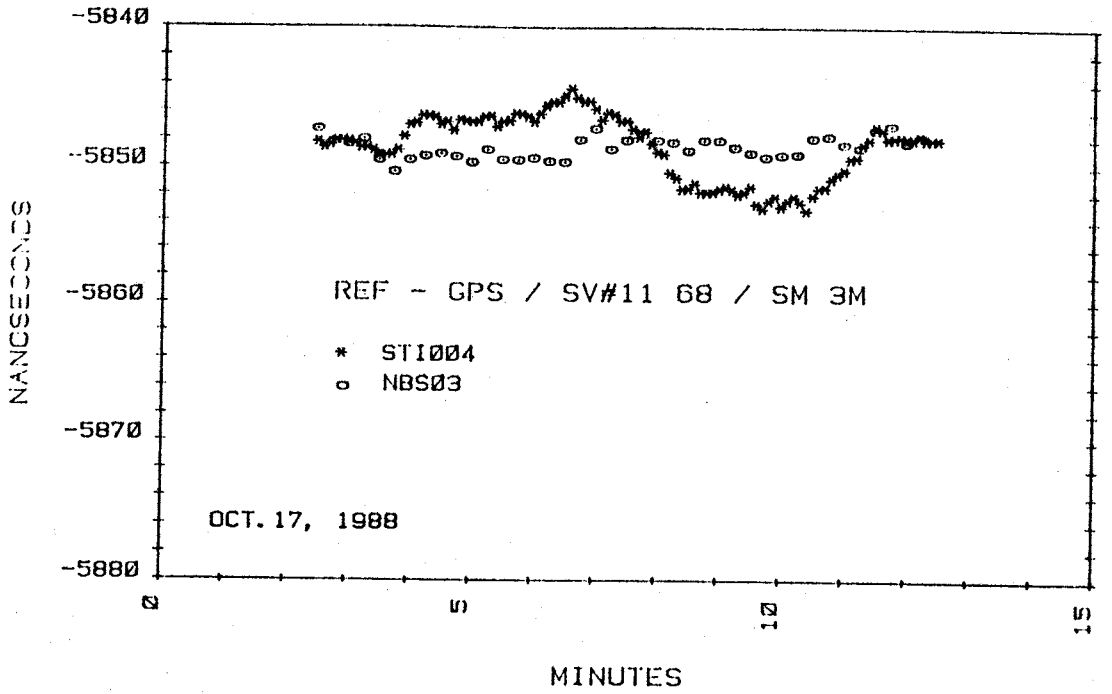


Figure 4. Results of a single track (SV#11 C168) with low RMS, smoothed over 3 minutes.

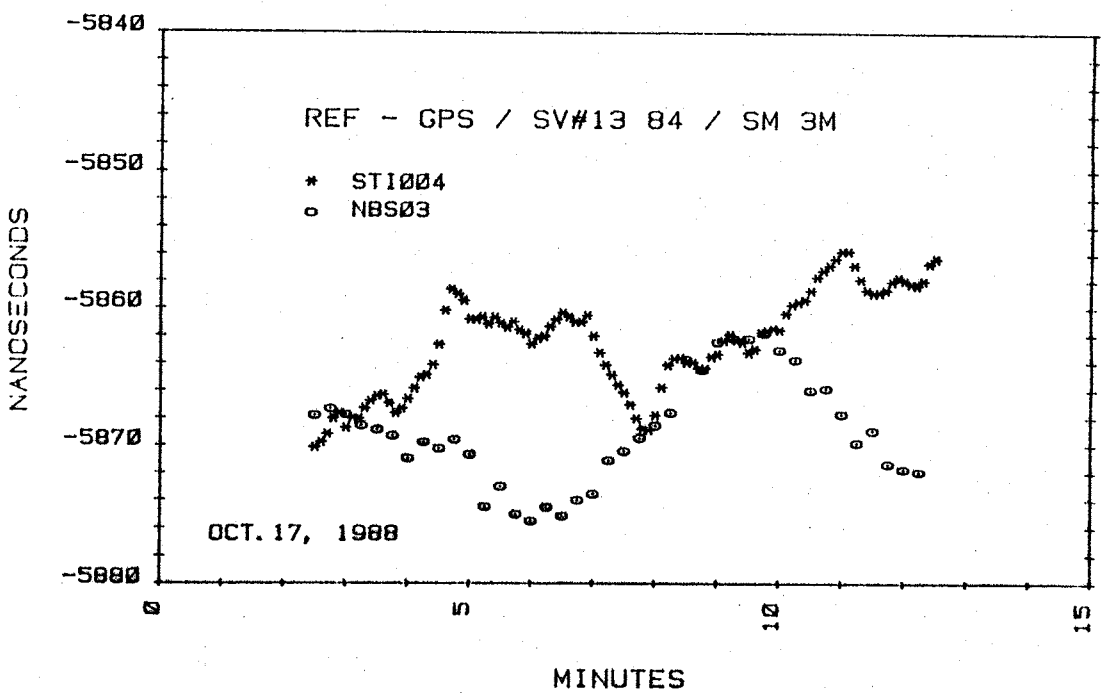


Figure 5. Results of a single track (SV#13 C184) with high RMS smoothed over 3 minutes.

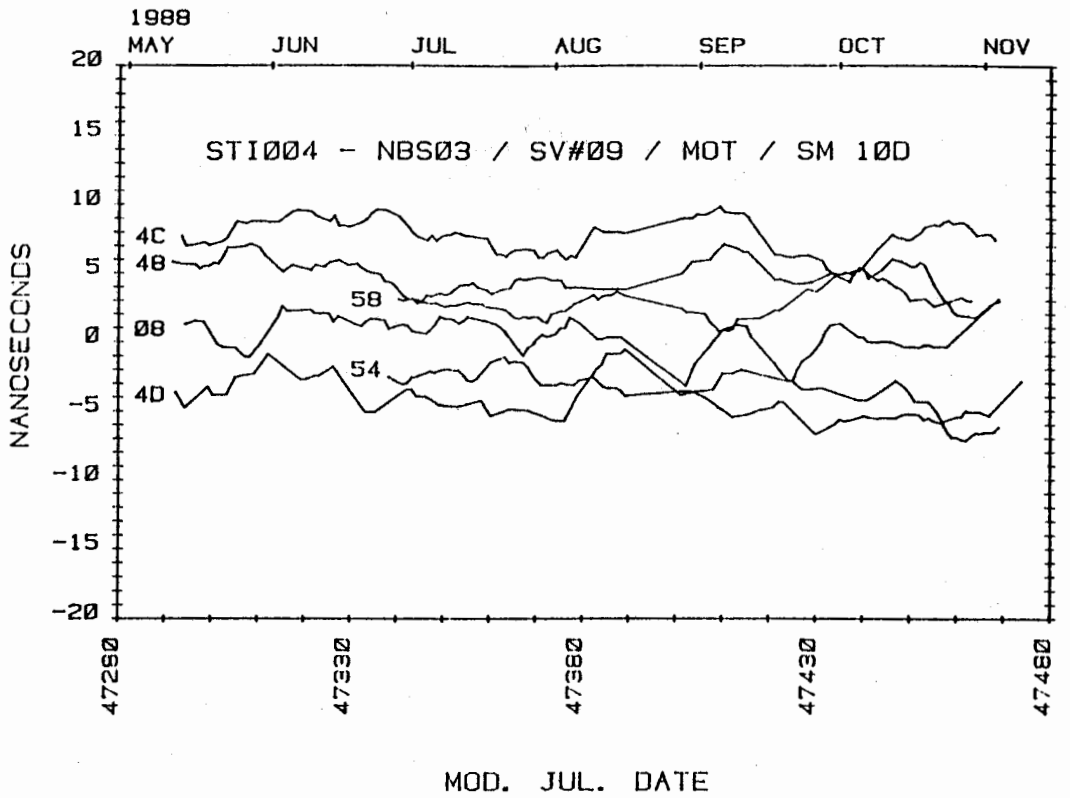


Figure 6. Common-view results of STI-receiver minus NBS-receiver for different classes of satellite SV#09 (referred to middle of the tracks, smoothed over 10 days).

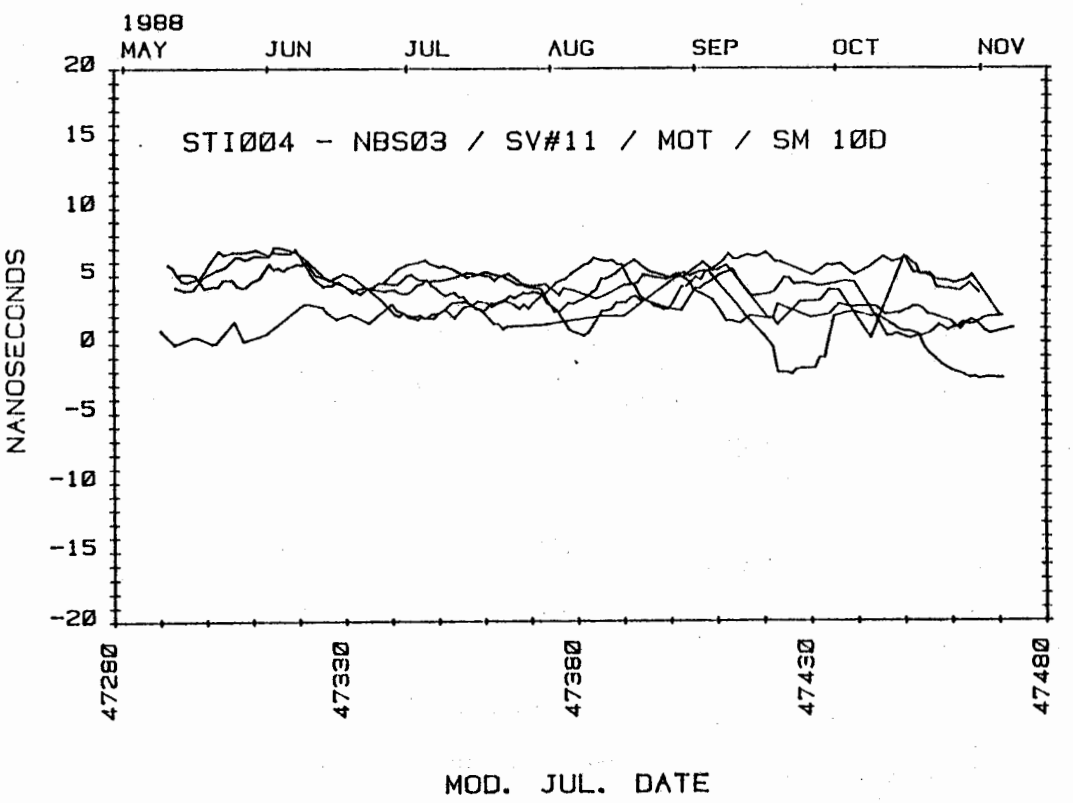


Figure 7. Common-view results of STI-receiver minus NBS-receiver for different classes of satellite SV#11 (referred to middle of the tracks, smoothed over 10 days).

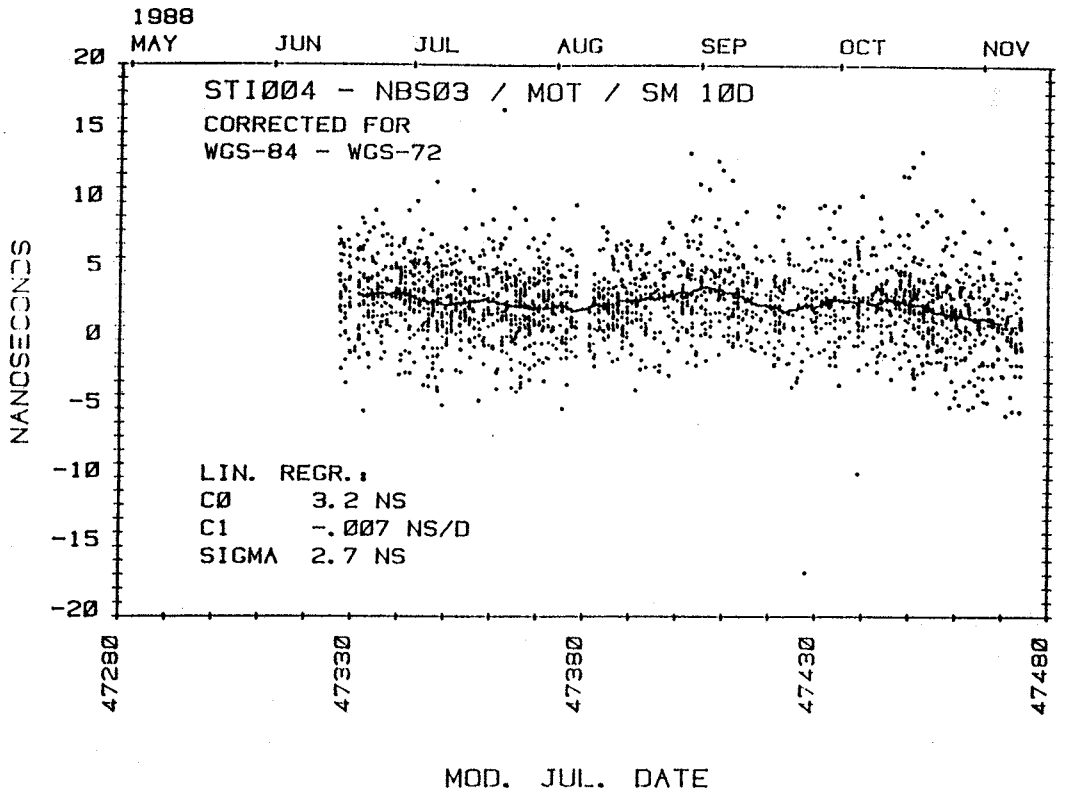


Figure 8. Data of Fig. 3 corrected for range differences resulting from the use of the WGS-72 geodetic system instead of the WGS-84 in the STI-receiver.

SV CL	MOT / RD / UNCORRECTED C0 ns	C1 ns/day	SIGMA ns	MOT / RD / CORRECTED C0 ns	C1 ns/day	SIGMA ns
ALL	2.902	-.010	3.90	3.189	-.008	2.67
3	.925	-.008	3.21	4.206	-.010	2.44
3 18	3.557	-.006	1.99	5.323	-.009	2.12
3 49	-.911	-.012	2.26	3.430	-.013	2.28
6	3.977	-.019	3.79	2.792	-.011	2.02
6 40	5.839	-.018	1.62	2.819	-.022	1.60
6 4C	-2.021	-.006	1.93	3.046	-.007	2.00
6 58	2.294	.003	2.13	2.381	-.002	2.16
6 8C	6.996	-.017	2.08	3.073	-.018	2.09
9	2.434	-.012	4.74	2.852	-.011	2.16
9 08	.859	-.008	2.35	3.511	-.009	2.35
9 48	4.464	-.002	2.04	1.735	-.002	2.07
9 4C	7.740	-.002	2.06	2.486	-.002	2.08
9 4D	-3.758	-.001	2.23	2.778	-.008	2.21
9 54	-1.824	-.037	1.54	3.824	-.038	1.47
9 58	2.461	.006	1.72	2.561	.005	1.73
11	4.158	-.017	2.58	3.946	-.017	3.41
11 18	1.893	.020	2.31	7.074	.021	2.46
11 68	6.517	-.062	2.62	4.926	-.062	2.59
11 74	4.709	-.023	1.56	5.576	-.025	1.59
11 84	3.276	.017	2.19	.155	.019	2.12
11 A0	3.849	-.017	2.09	4.027	-.016	2.09
12	2.543	-.010	4.03	3.711	-.010	2.28
12 08	-2.239	-.018	2.22	2.603	-.014	2.21
12 48	6.679	-.003	2.23	3.077	-.006	2.17
12 58	3.294	-.015	1.57	4.086	-.012	1.61
12 74	2.165	-.013	2.13	4.983	-.009	2.21
13	2.321	.011	3.45	1.876	.017	2.99
13 30	.336	.013	3.24	4.205	.014	3.17
13 4C	1.542	-.002	2.07	3.493	-.006	2.04
13 84	3.932	.034	2.30	-.835	.034	2.33

Table 2. Linear regressions referred to the same begin time (June 28, 1988) for all tracks of all satellites, for all tracks of single satellites and for tracks of the same class of single satellites before and after applying of range corrections.

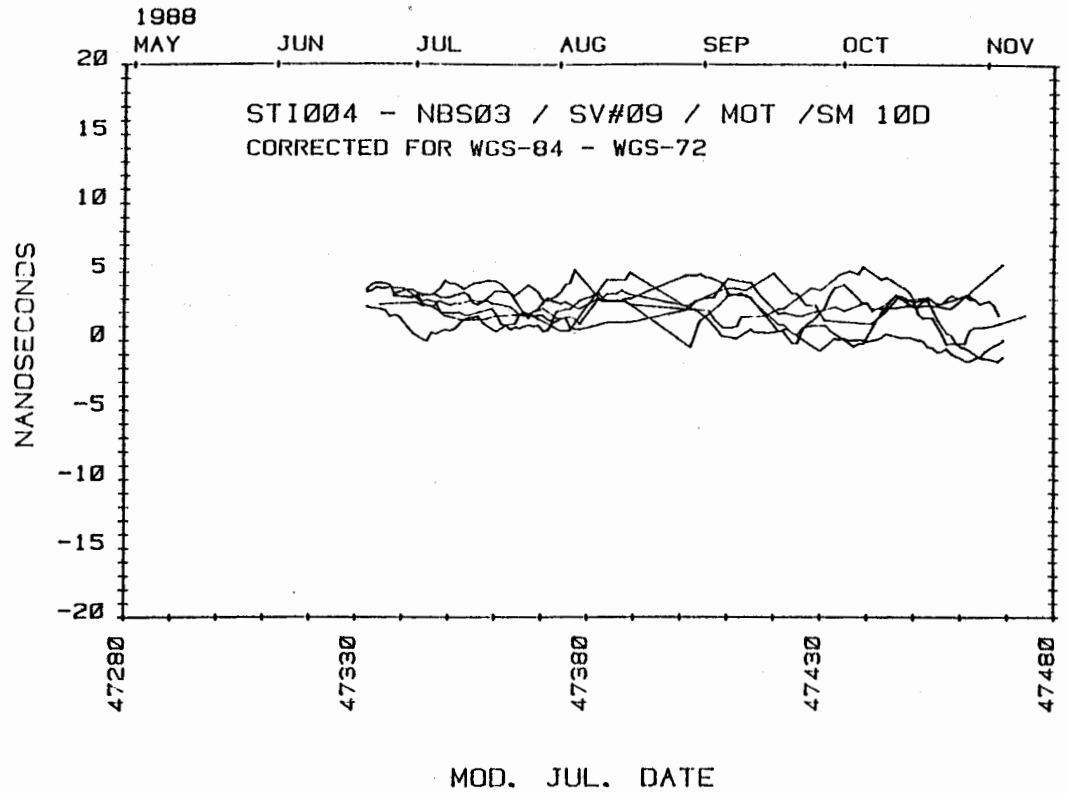


Figure 9. Data of Fig. 6 with the same corrections as applied in Fig. 8.

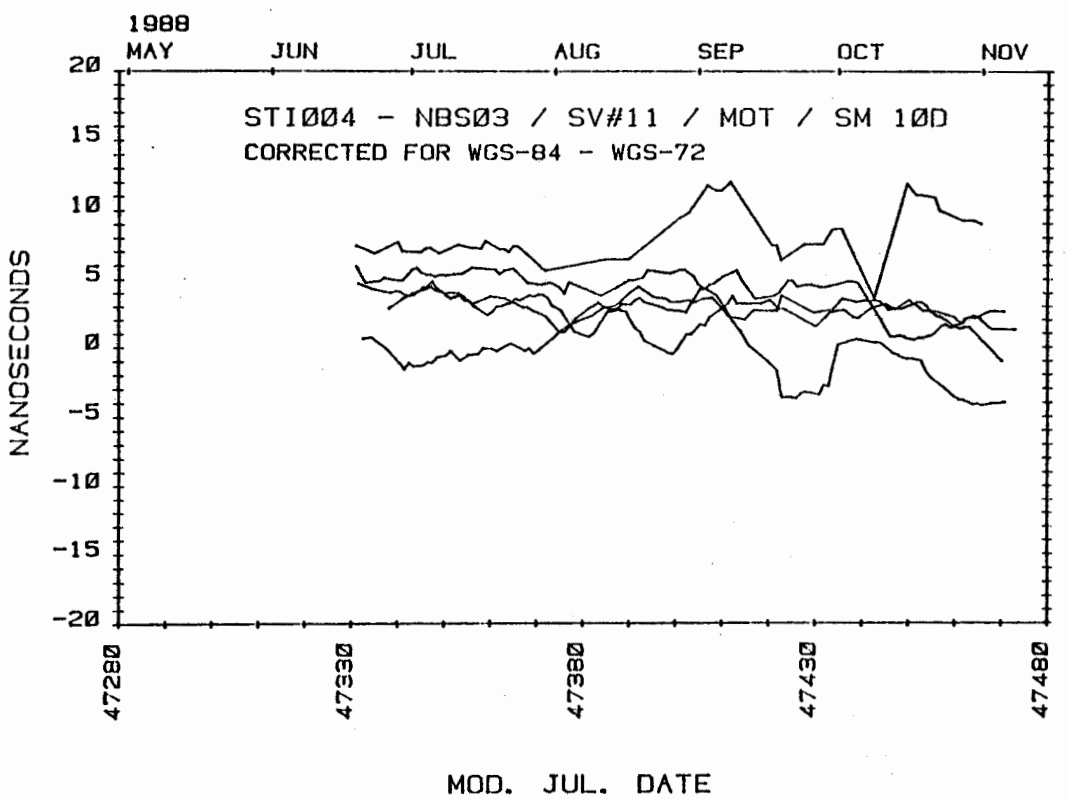


Figure 10. Data of Fig. 7 with the same corrections as applied in Fig. 8.

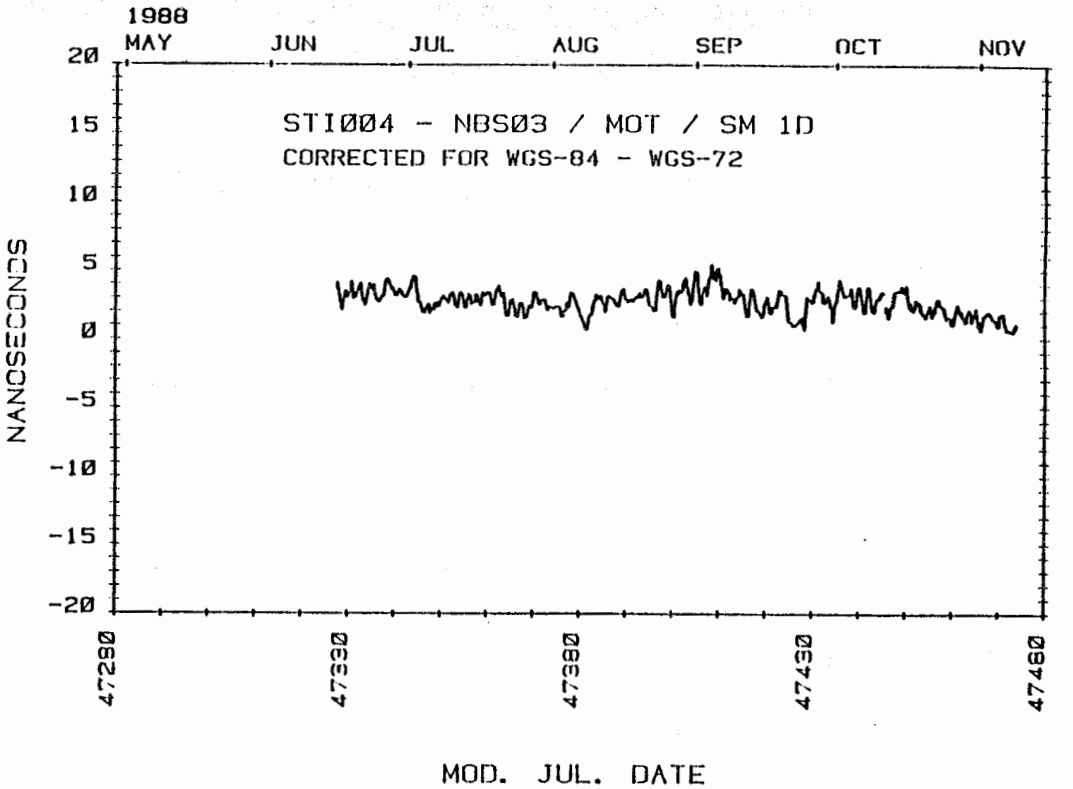


Figure 11. Common-view results of STI-receiver minus NBS-receiver referred to the middle of the tracks and smoothed over 1 day.

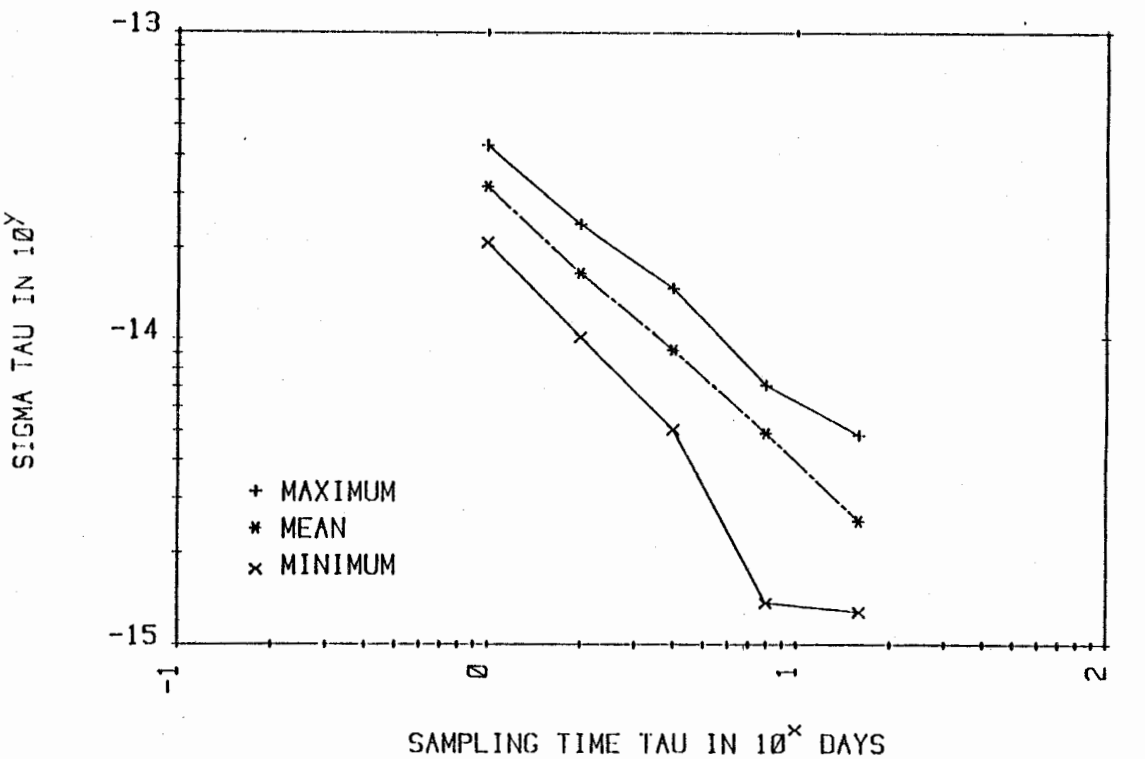


Figure 12. Maximum, minimum and mean of the square-root of the Allan variance computed from the data shown in Fig. 3.

COMPARISON OF GPS TIME RECEIVERS : A USER'S POINT OF VIEW

M. GRANVEAUD[1], W. LEWANDOWSKI[2], P. UHRICH[1], R. TOURDE[1]

[1] *Laboratoire Primaire du Temps et des Fréquences, Observatoire de Paris, Paris*

[2] *Bureau International des Poids et Mesures, Sèvres*

INTRODUCTION

Most of the operational GPS time intercomparisons between time/frequency laboratories are currently carried out through receivers designed by the National Institute of Standards and Technology (formerly National Bureau of Standards) ; various versions of this design have been built up : the most common one has been the Allen Osborne TTR5 device . One of them has been running at the Laboratoire Primaire du Temps et des Fréquences (LPTF, Observatoire de Paris) for several years thanks to a cooperation between this laboratory, the National Institute of Standards and Technology and the Bureau International des Poids et Mesures (BIPM). Furthermore, the BIPM bought in 1988 a GPS time receiver designed by the Société d'Etudes, Recherches et Constructions Electroniques (SERCEL).

Since July 1988, this unit has been running at the LPTF ; it has been regularly compared to the master clock of the Paris Observatory UTC(OP). So, since the above date, the time measurements of two differently designed GPS receivers have been referred to UTC(OP).

At the European Frequency and Time Forum 1988, characteristics and results obtained with the SERCEL NRT type receiver were shown by the company representatives [1]. It is intended to give insight into the SERCEL NRT possibilities from the user's point of view and more precisely to compare the NRT and the TTR5 units [2, 3].

I - GENERAL CHARACTERISTICS

Tables 1 and 2 exhibit characteristics of the two receivers which have been selected as important by the authors.

Table 1 refers to the hardware part. Clearly, the NRT unit is an open system. Electronic cards can be added to the basic device to enlarge its possibilities ; if equipped with four signal acquisition cards, simultaneous measurements from four satellites give access to the GPS time via different ways which leads to a better knowledge of this time. Furthermore, a navigation solution can be obtained in real time. Another electronic card allows the steering of a clock (internal or external) with respect to the GPS time at the level of resolution of 100 ns or less, depending essentially on the available local clock.

Another difference between the two receivers is concerned with the data processing unit. It is fully integrated with the TTR5 device ; it has to be added to the NRT receiver. Advantages of both processing systems can be appreciated through various criteria - hardware as well as software- : memory possibilities, data safeguard, data availability, computing and/or displaying modifications,...

It is also indicated on Table 1 that the NRT antenna unit is delivered with a protection plane against multipaths and wave reflections. The TTR5 antennas are not systematically equipped with this plane.

Table 2 presents the software characteristics. A general overview of the table shows that the NRT unit has been designed in order that various parameters be chosen by the user ; it is a different approach as the TTR5 one.

Let us focus on the computation of the final result UTC(lab)-GPS and its associate RMS by the two receivers. Both of them work through three steps : first, the raw measurements are stored according to a basic periodicity ; then it is operated a computation of UTC(lab)-GPS over an intermediate period ; the final result of UTC(lab)-GPS is obtained through an adjustment of the intermediate values.

In the TTR5 receiver, the intermediate values are obtained from a quadratic adjustment of 15 pseudo-range code measurements. In the NRT device, the phase information is taken into account beside the code one ; the intermediate values are mean values of the code and phase measurements. Computation of the final result is carried out through linear adjustments of the intermediate values in both receivers and the

RMS values are in both cases the standard deviation of the n intermediate results ; but the NRT computing process divides the standard deviation by the square root of n.

TTR5 receiver	NRT receiver
GENERALITIES	
. one channel navigation solution has to take sequential tracks (possible drift of the local reference clock)	. possibility of four channels real time navigation solution . possibility of steering one clock on GPS time
ANTENNA	
. Intermediate Frequency synthesized in the antenna box	. Intermediate Frequency synthesized in the receiver box . protection against wave reflection (microwave absorber)
DATA PROCESSING	
. integrated into the receiver EPROM + RAM + microcomputer	. through one PC type microcomputer which is independent from the receiver
. memory of 150 tracks data	. memory of the tracks data on hard disk
. if power is switched off, the last 19 tracks data are kept	. if power is switched off, all the recorded data are kept

Table 1 - *Hardware selected characteristics of the TTR5 and NRT GPS receivers.*

TTR5 receiver	NRT receiver
TRACKS POSSIBILITIES	
per day	
48	100
MEASUREMENT RESULTS	
display	
1 or 0.1 ns (choice)	1 or 0.1 ns (choice)
basic period	
1 s	0.6 s
intermediate period	
15 s	≥ 0.6 s (choice)
measurement time interval	
120 to 2400 s (choice)	120 to 3600s (choice)
results	
intermediate value : quadratic adjustment (least square method) of 15 pseudo-range code measurements. It refers to the middle of the intermediate period.	intermediate value : average of pseudo-range code and phase measurements over the intermediate period. It refers to the end of the intermediate period.
final value : linear adjustment (least square method) of the intermediate values. It refers to the beginning date of the track.	final value : linear adjustment (least square method) of the intermediate values. It refers to the beginning or middle or end date of the pass (choice)
RMS value : std deviation of the intermediate values (nominal value : 10 ns)	RMS value : std deviation of the n intermediate values divided by √n (nominal value : 1 to 5 ns in common view)
DATA	
azimuth, elevation	
refer to the end date of the track	refer to the date of the final value
ionospheric correction	
available (modelization)	available (modelization)
tropospheric correction	
not available	available

Table 2 : *Software selected characteristics of the TTR5 and NRT GPS receiver.*

II - COMPARISON OF GPS TIME RECEIVERS

Since September 1988, the NRT1 unit of the BIPM has been running continuously beside the TTR5 device at the LPTF. Their antenna have been set up at two different poles as indicated on figure 1. Their coordinates are the following ones in the WGS84 system :

TTR5 antenna	X = 4 202 780.7	m
	Y = 171 367.7	m
	Z = 4 778 658.8	m

as obtained from campaigns of the Institut Géographique National (IGN) and transformations in the WGS84 system. For the NRT,

NRT1 antenna	X = 4 202 783.7	m
	Y = 171 348.2	m
	Z = 4 778 677.9	m

as a result of the IGN and SERCEL measurements. Work carried out by W. Lewandowski on this topics has shown that coordinates of both antennas are wrong (up to 3m for the Y coordinate of the TTR5). The NRT1 antenna has been set up 12 meters higher than the TTR5 one in such a way that there is no screen to all direct GPS signals ; which is not the case with the TTR5 antenna. Nevertheless, it is worth mentioning that a TV antenna is fixed quite near the NRT one.

Both receivers measure the time synchronization UTC(OP)-GPS via the international time comparison schedule : from September to mid December 88, 20 passes were taken (first schedule) - 14 with the standard length of 13 minutes and 6 with 12 minutes -. A new schedule (second schedule) was introduced mid December according to the BIPM recommendation and 17 tracks are common to the two receivers. The measurement results refer at the beginning of the track for the two receivers. Table 3 details results obtained for each pass of the first schedule; it presents the averages m and the standard deviations σ computed from the raw data $D(t) = [UTC(OP)-GPS]_{TTR5} - [UTC(OP)-GPS]_{NRT1}$ over the interval September 20 - December 19. If the measurement results are referred to the middle date of the tracks, the averages are practically not changed but the standard deviations are decreased.

II-1 - Mean values

The mean values m of several passes are significantly different from zero ; which indicates an inconsistency between the coordinates of the two antenna as mentioned in [4]. Careful consideration of the D(t) variations around m indicates possible systematic effects : a drift of the order of 5 ns over 3 months is detected with 5 passes ; furthermore, important and long-term variations appear in 4 cases. In a general way, groups of homogeneous results differing by about 10 ns from surrounding results are relatively frequent. Figures 2, 3, 4 and 5 illustrate three types of results : time bias (fig. 2), time drift (fig. 3), long term variations (fig. 4 and 5).

Pass N°	sat	tracking time	El (°)	Az (°)	m (ns)	σ (ns)	Remarks
1 *	3	21h43	26	318	1.4	4.1	apparent time drift
2	3	6h10	29	162	- 4.6	6.2	apparent time drift
3 *	6	2h02	78	13	- 8.7	4.4	long term variations
4	6	23h44	30	195	- 2.3	4.8	
5 *	9	3h02	66	278	- 2.7	4.2	
6	9	5h02	51	63	- 7.9	3.7	
7	9	1h42	33	230	2.2	4.6	
8	9	4h00	70	7	- 5.3	5.6	
9	9	6h26	21	101	- 8.7	6.6	long term variations
10*	9	1h07	17	220	- 2.4	4.0	
11*	11	20h50	36	237	- 4.6	4.4	apparent time drift
12	11	19h58	53	269	- 3.6	3.5	apparent time drift
13	11	5h50	39	53	- 6.2	4.5	
14	12	2h46	23	247	0.1	5.0	long term variations
15	12	7h24	39	113	- 9.7	4.7	
16	12	5h18	71	354	- 4.9	5.1	long term variations
17*	12	4h18	56	291	- 2.1	3.5	
18	13	19h42	24	326	3.2	3.8	apparent time drift
19	13	7h50	34	40	- 7.3	4.3	
20	13	5h34	41	124	- 6.9	3.7	

* 12 minutes long

Table 3 : *Averages m and standard deviations σ of the time differences $D(t) = [UTC(OP)-GPS]_{TTR5} - [UTC(OP)-GPS]_{NRT1}$ and other data relative to 20 common view GPS trackings from September to December 1988.*

The tracking time, the elevation and the azimuth refer to September 20 (MJD = 47424).

II-2 - Standard deviations

The values of the standard deviations of the D(t) raw data are in the range 3.5 ; 6.6 ns (table 3) for the first schedule and 4.3 ; 9.0 ns for the second one, expressing a slight degradation.

Let us call R the ratio between the TTR5 RMS values and the NRT1 corrected* RMS values for each pass. The average value of R is 1.7 with a division of the R values into two classes : in the first one, they are equal or superior to 2 ; in the second one, the ratio is near to 1, generally due to an increase of the NRT1 RMS which goes up from about 5 to 10 ns. Figures 6 and 7 show for two tracks the intermediate 15s results $[UTC(OP)-GPS]_{NRT1}$ and illustrate the RMS degradation ; clearly a pseudo periodic signal coming from the environmental conditions perturbs the GPS signal reception (figure 7). Figure 8 presents the space repartition of the above two classes among the 20 passes which were taken at the end of 1988 ; the first class is indicated by a square and the second one by a circle. Some tracks cannot be classified because their ratios R change from one day to the other between 1 and 2. They are denoted by a triangle.

Considering various plots, as this one of figure 8, from September 88 to March 89, it appears that the azimuthal angle of about 45° gives rise to ratios generally near to 1, i.e., it is a specifically bad azimuth angle for the NRT1 unit. Other azimuthal directions indicate also degradations : for example, the 315° angle azimuth.

II-3 - Two-sample variances

The two-sample variances of the D(t) raw data have been computed for time intervals of 1, 2 and 4 days. Taking into account the TTR5 stability as estimated in 1988 [5], a rough estimate of the NRT1 unit stability has been computed. For 1 day, the dispersion of the results obtained out of 15 time series is very large and does not allow one to get a satisfactory estimate. For time intervals of 2 and 4 days, it is found that the NRT1 device stability is similar to the TTR5 one, i.e., $\sigma_y(2, 2d) = 3 \times 10^{-14}$ and $\sigma_y(2, 4d) = 1.5 \times 10^{-14}$.

---

* The RMS displayed by the NRT1 are multiplied by $\sqrt{n}$ (n being the number of 15s samples) to be coherent with the TTR5 RMS.

CONCLUSION

Two points appear clearly. First, the NRT1 and the TTR5 GPS time receivers have got different designs ; the NRT unit, as an open system, offers more possibilities and choices than the TTR5. This one emphasizes the operational aspect. Second, the time comparison results obtained with two units at the LPTF, in a particular environment, are not sufficient to get a significant and precise view of their characteristics. The NRT1 RMS values as obtained at the LPTF are smaller than the TTR5 ones but they are of the same order as the RMS values displayed by one TTR5 unit running at the LASSO station of Grasse ; which would indicate that the Paris Observatory environment is responsible for a not negligible part of the observed perturbations. Furthermore, questions arise with respect to the results of figure 3. Are the receivers, one of them or both, working uncorrectly ?

In August 1988, another NRT receiver was compared with the two devices presented in this paper. Unfortunately, the numbers of data were not sufficient to get significant results. The metrological qualification of the NRT receiver would need a long-term comparison of three identical units running in the same environmental conditions.

REFERENCES

[1] GOUNON, R., LELONG, J.N. : GPS high precision time transfer methods and related equipment, Proc. 2d EFTF, 195 (1988).

[2] NARD, G., RABIAN, J., GOUNON, R., 1987, "Utilisation des signaux du GPS en mode differentiel instantané pour des applications temps-fréquence de haute précision", Proc. 1st EFTF, 237 (1987).

[3] LEWANDOWSKI, W., WEISS, M.A. and DAVIS, D., 1986, "A calibration of GPS equipment at time and frequency standard laboratories in the USA and Europe", Metrologia, 24, p. 181-186 (1987).

[4] GUINOT, B., LEWANDOWSKI, W. : Nanosecond time comparisons in Europe using the GPS, Proc. 2d EFTF, 187 (1988).

[5] GRANVEAUD, M., PARCELIER, P., FREON, G., TOURDE, R., VINCENT, M., GRUDLER, P. : Operational highly precise time transfers in France by GPS and by the TV method, Proc. 2d EFTF, 213 (1988).

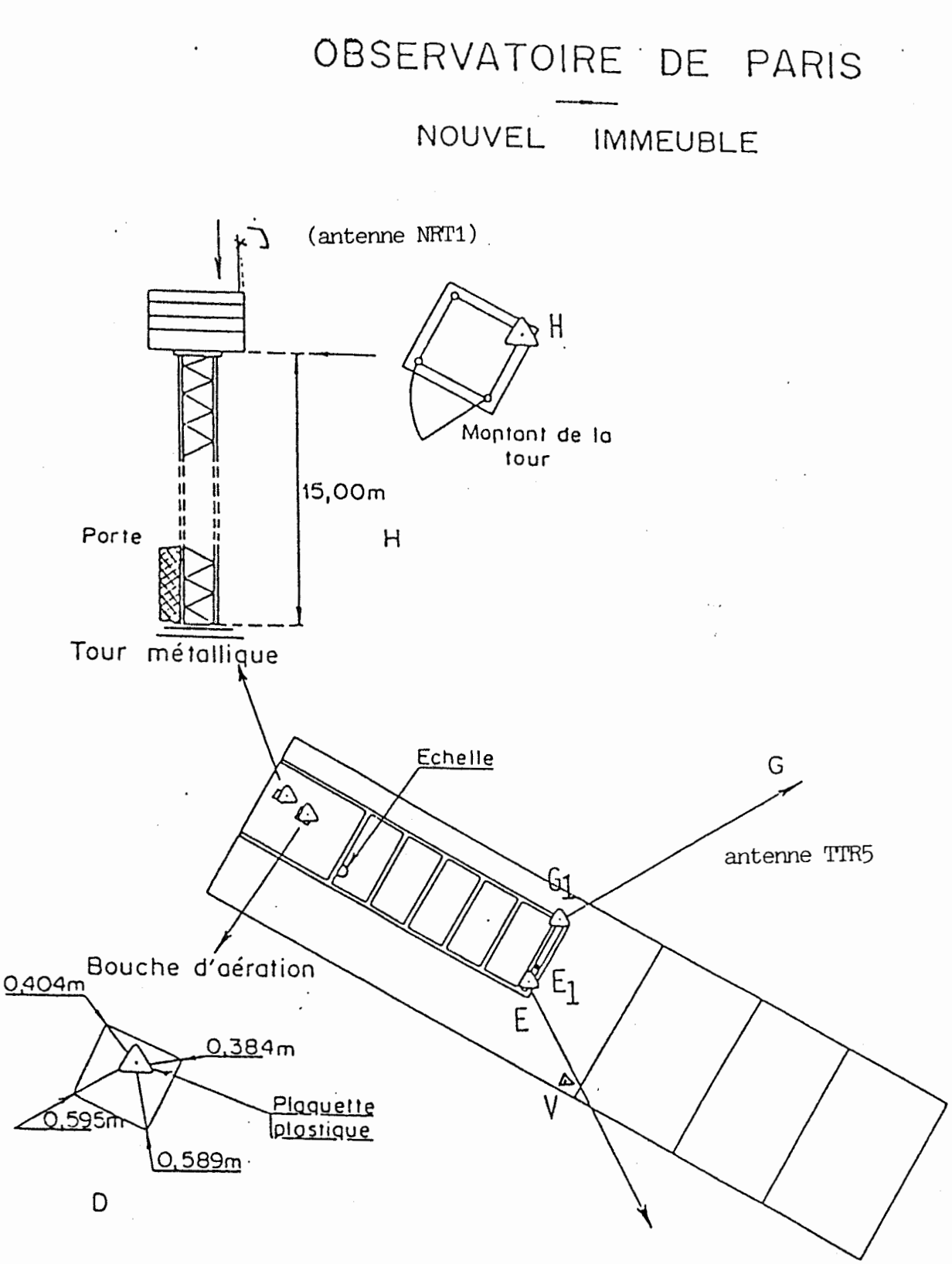


Figure 1 - Location of the NRT1 and TTR5 antennas at the LPTF (Paris Observatory)

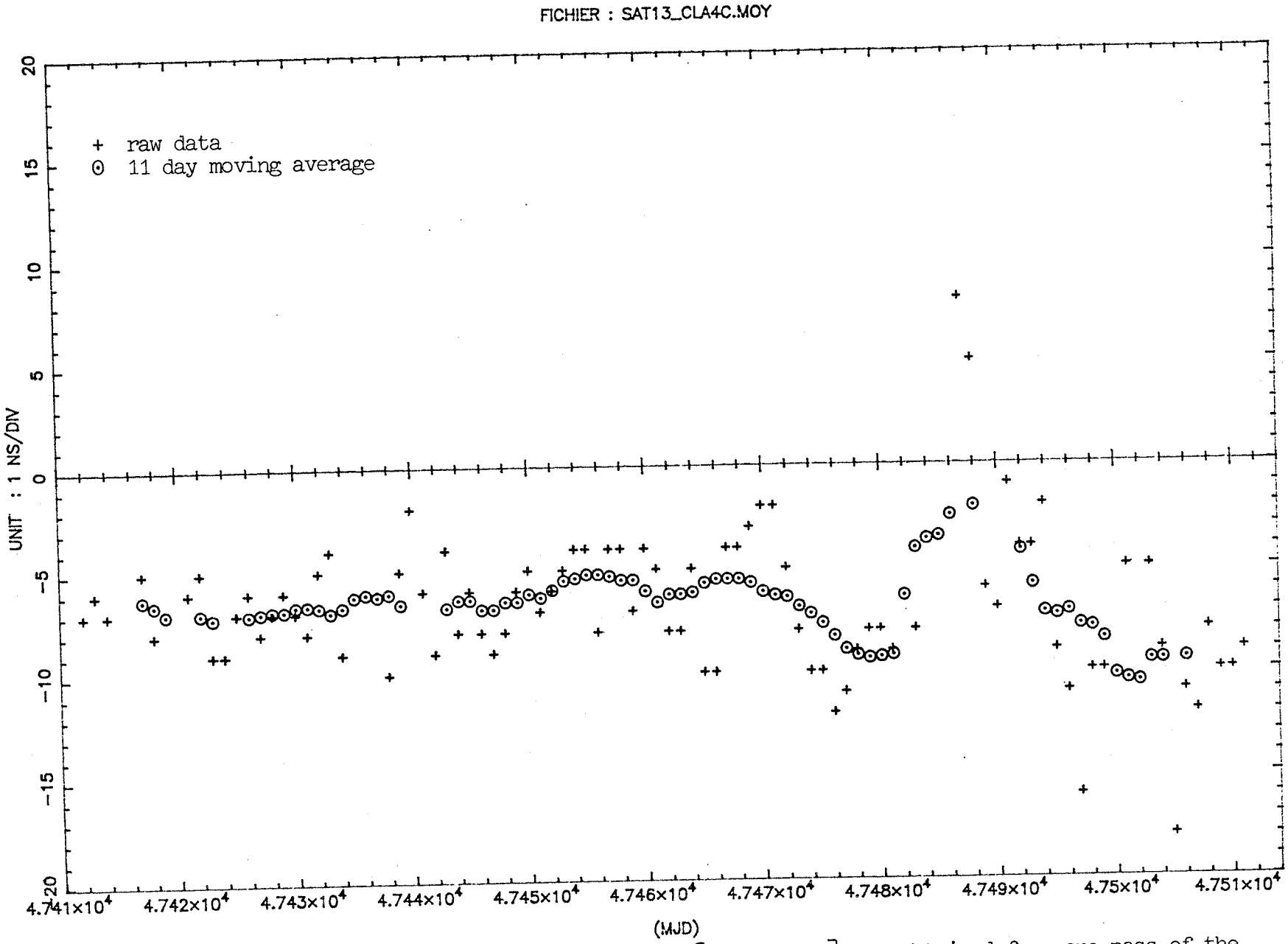


Fig. 2 - Time differences D(t) = $[UTC(OP)-GPS]_{NRT1}$ - $[UTC(OP)-GPS]_{TTR5}$ obtained from one pass of the satellite 13 (pass number 20).

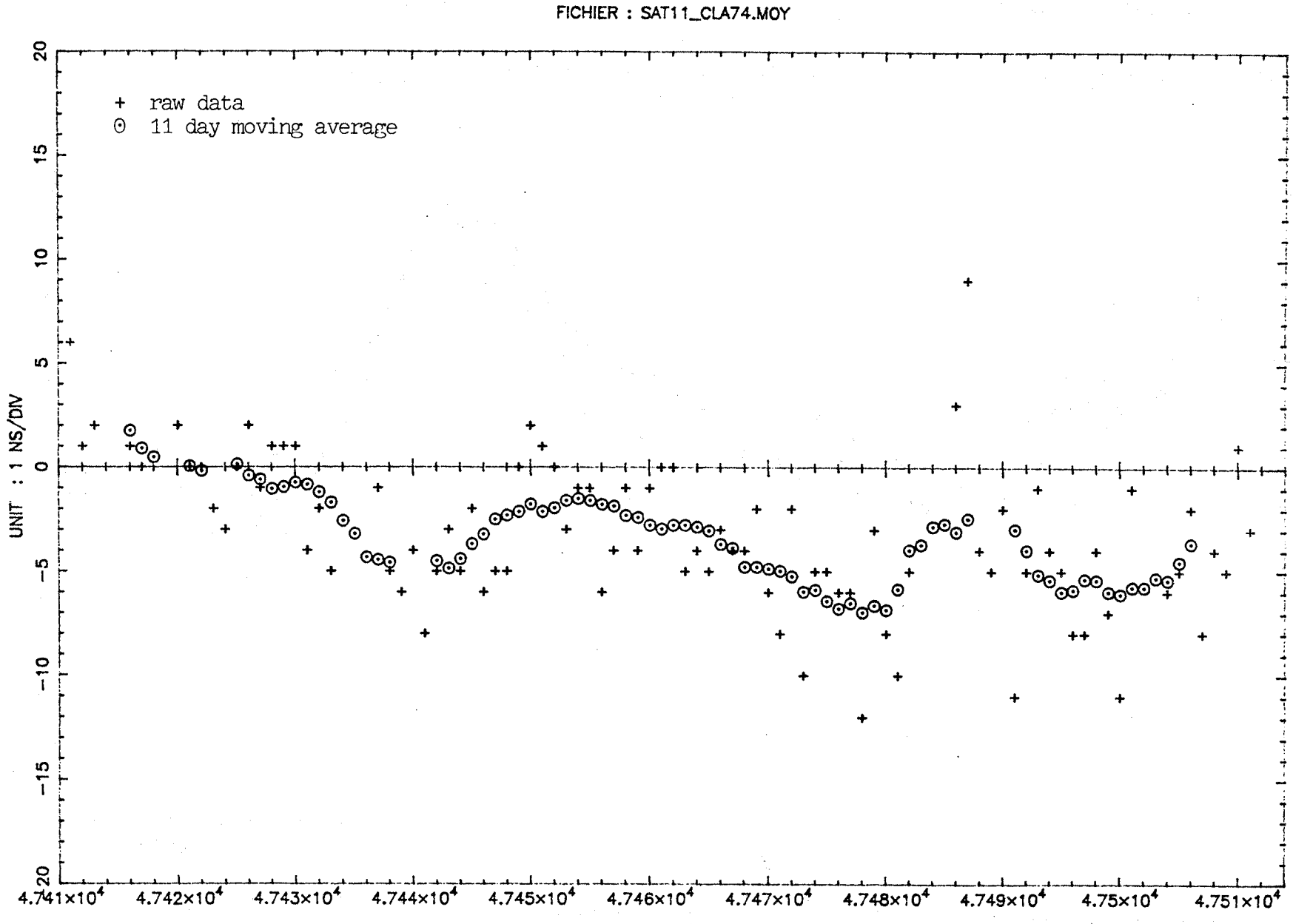


Fig. 3 - Time differences D(t) = $[UTC(OP)-GPS]_{NRT1}$ - $[UTC(OP)-GPS]_{TTR5}$ obtained from one pass of the satellite 11 (pass number 12).

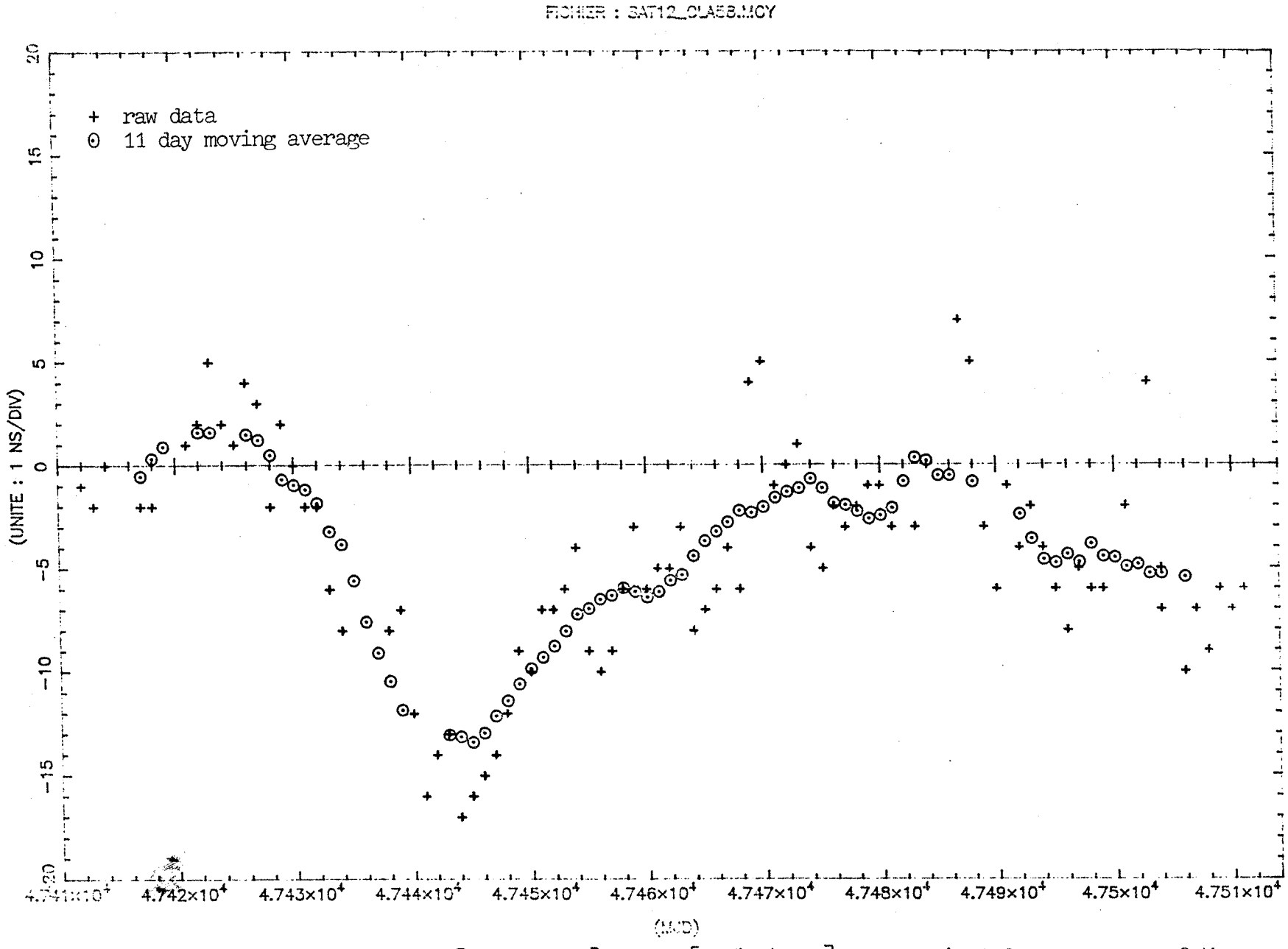


Fig. 4 - Time differences $D(t) = [UTC(OP)-GPS]_{NRT1} - [UTC(OP)-GPS]_{TTR5}$ obtained from one pass of the satellite 12 (pass number 16).

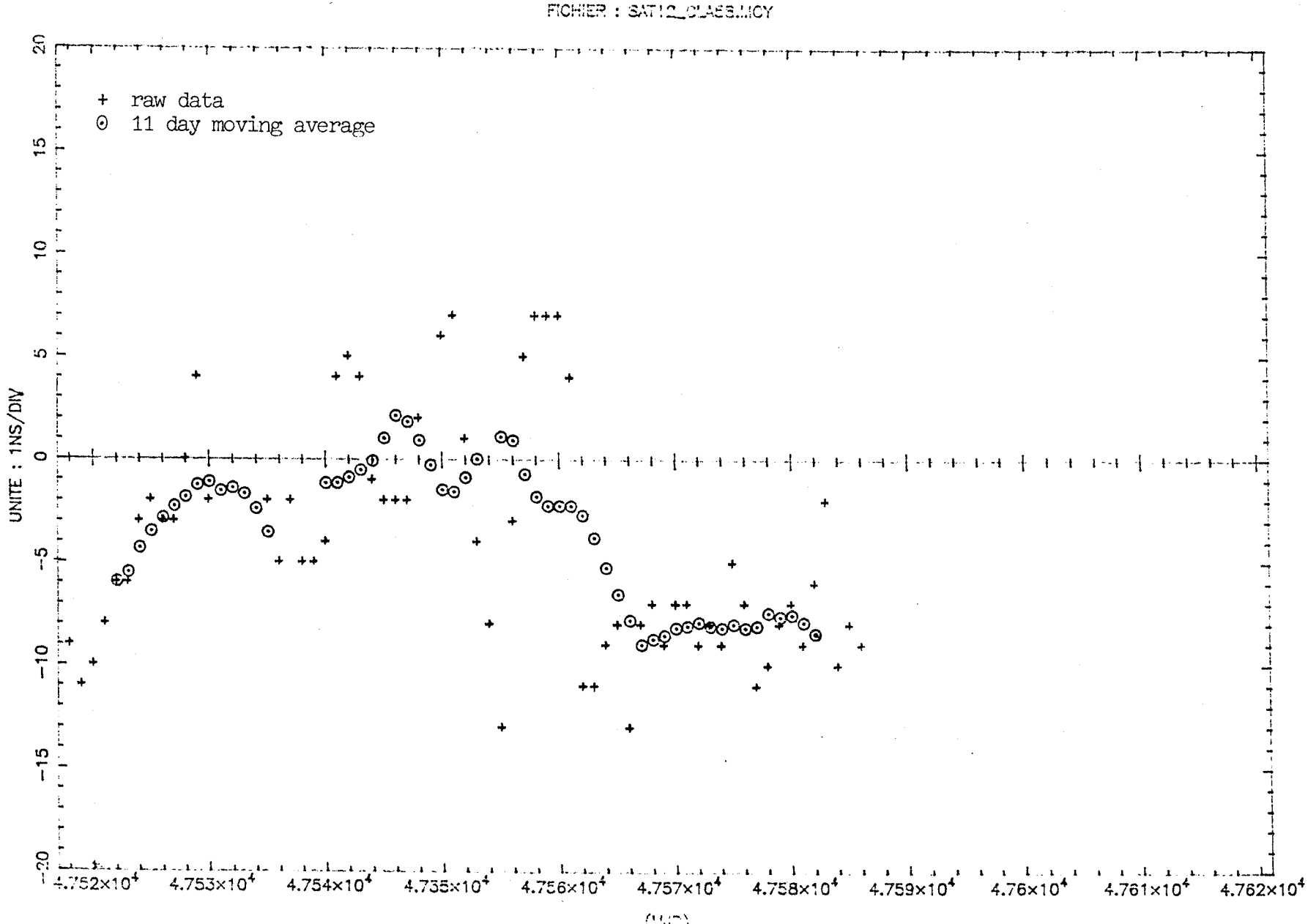


Fig. 5 - Time differences $D(t) = [UTC(OP)-GPS]_{NRT1} - [UTC(OP)-GPS]_{TTR5}$ obtained from one pass of the satellite 12.

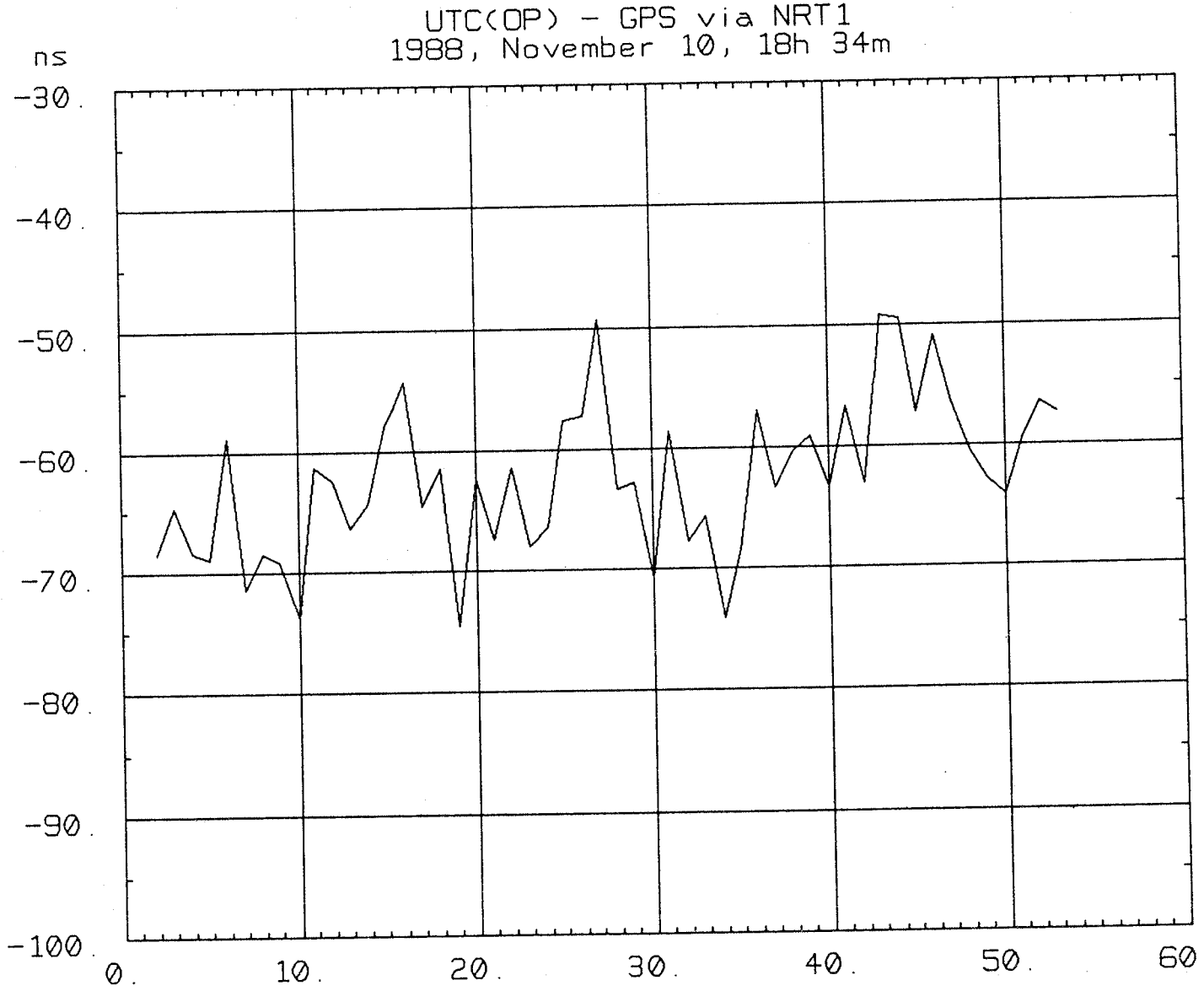


Fig. 6 - Intermediate 15s results UTC(OP) - GPS as obtained by the NRT1 receiver during 780s (satellite 3).

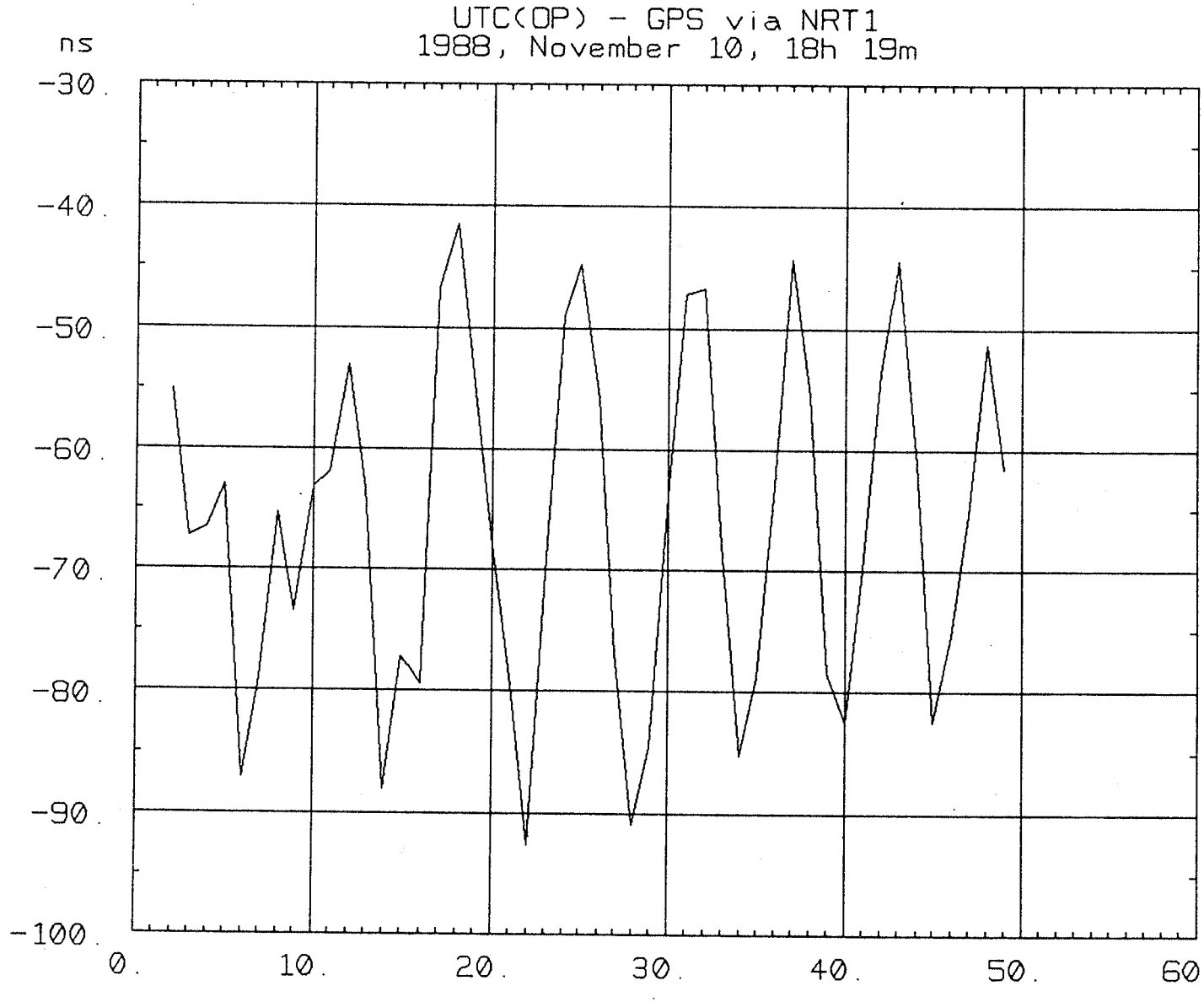
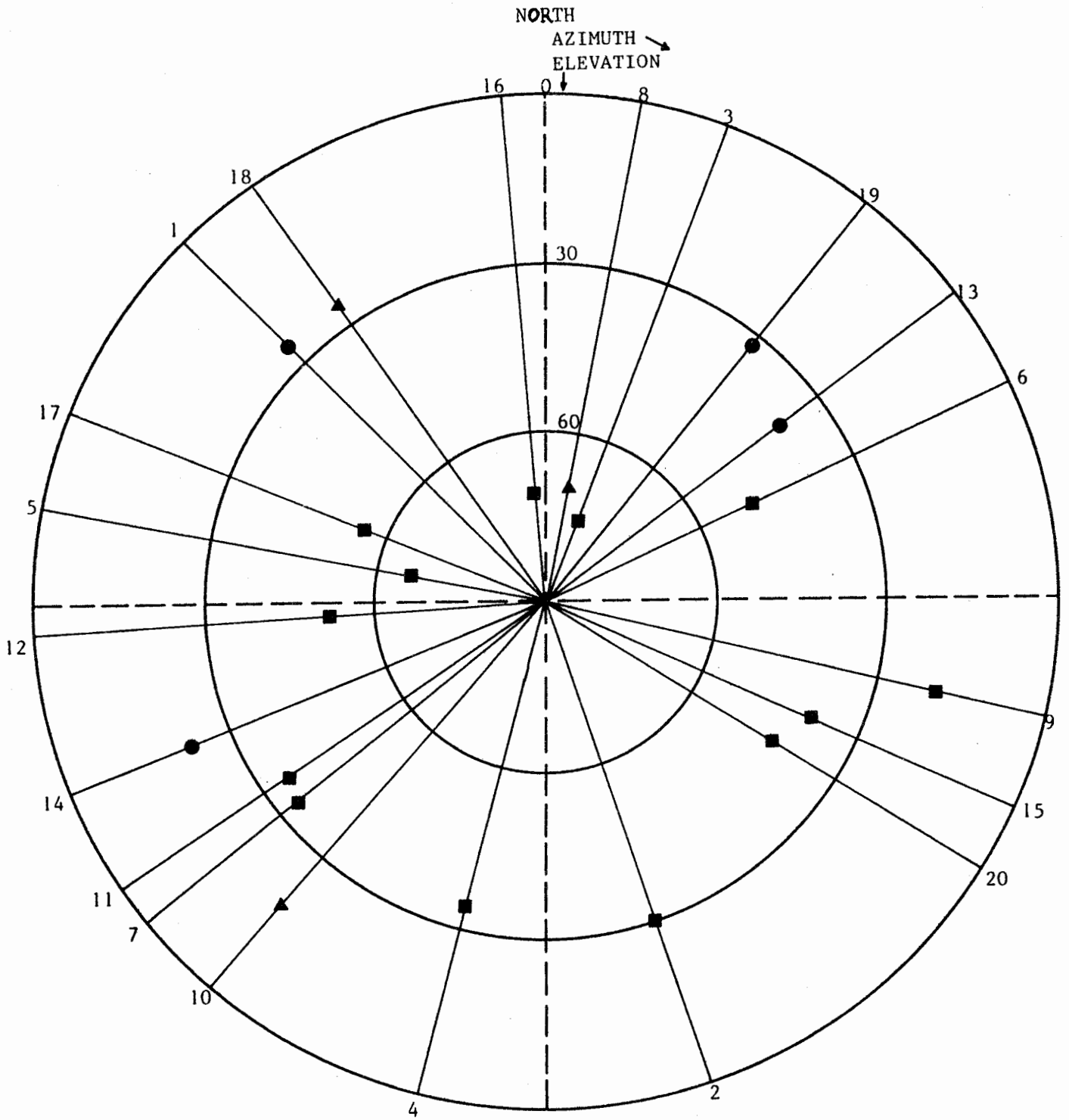


Fig. 7 - Intermediate 15s results UTC(OP) - GPS as obtained by the NRT1 receiver during 780s (satellite 3, taken 15 minutes before the track of figure 6).



● R ≃ 1

■ R ⩾ 2

▲ R changing

Fig. 8 - Space repartition of the 20 passes (end November 1988) with respect to the R values (R = RMS(TTR5)/RMS(NRT1)).
The passes are identified by their numbers of table 3 which are written on the outside circle.

Time References from GLONASS Satellites

I D Kitching and P Daly

Department of Electrical and Electronic Engineering
University of Leeds, Leeds LS2 9JT, United Kingdom

ABSTRACT

The Soviet Union's global navigation satellite system, GLONASS, has eight operational satellites with the full contingent of 24 satellites being available in a possible two years at the present launch rate. Since only a single satellite is required for time transfer to a known location the potential of Glonass as a disseminator of time and frequency can be fully evaluated in this pre-operational phase. Using the Glonass low precision codes, and without making dual frequency corrections, time transfer with Glonass is obtainable to accuracies of the order of 100 nanoseconds or better.

This paper discusses the timing references available from Glonass and how they are applied in a satellite to user time transfer link. Data is presented showing the relationship between Glonass system time and UTC(USNO) over an extended period of time. Finally the linking of Glonass system time to other Soviet atomic time scales is studied.

INTRODUCTION

Glonass provides worldwide time dissemination and time transfer services in the same manner as the NAVSTAR GPS and exhibits the same advantages as Navstar does over other existing timing services [1]. Time transfer is both efficient and economic in the sense that direct clock comparisons can be achieved via Glonass between widely separated sites without the use of portable clocks. Event time tagging can be achieved with the minimum of

effort and users can reacquire Glonass time at any instant due to the continuous nature of time aboard the satellites.

The first release from the Soviet Union of detailed Glonass information occurred at the International Civil Aviation Organisation (ICAO) special committee meeting on Future Air Navigation Systems (FANS) in Montreal in May 1988 [2]. The report contained nominal orbital information as well as detailed descriptions of the Glonass C/A code structure and transmitted data message. The majority of this information had already been made available via other publications [3], [4], [5]. The reader is referred to the references at the end of this paper for generalised Glonass information. Currently eight Glonass satellites are in full operation giving single satellite coverage at Leeds of 21 to 23 hours a day. These satellites are listed in table 1.

SAT ID	COSMOS	GLONASS	CHN	L1/MHz
1988-43A	1946	34	12	1608.7500
1988-43B	1947	35	23	1614.9375
1988-43C	1948	36	24	1615.5000
1988-85A	1970	37	18	1612.1250
1988-85B	1971	38	7	1605.9375
1988-85C	1972	39	10	1607.6250
1989-1A	1987	40	9	1607.0625
1989-1B	1988	41	6	1605.3750

Current Active Glonass Satellites 1-3-89.

Table 1

TIME FROM GLONASS

Time transfer from Glonass is achieved in a straight forward manner, Figure 1. Each satellite transmits signals referenced to it's own on board clock. The Control Segment monitors the satellite clocks and determines their offsets from the common Glonass system time. The clock offsets are then up loaded to the satellites as part of the transmit data message. A user at a known location receives signals from the satellite and by decoding the data

stream modulated on to the transmission is able to obtain the position of the satellite, as well as the satellite's clock offset from the common system time, as a function of time. Hence the signal propagation time can be calculated at any instant. The time at which the signals are transmitted is also contained in the data message and so combining this with the propagation time and correcting for the satellite clock offset allows transfer to Glonass system time. Any other user who has a satellite visible is also able to transfer to the same time scale.

Though this simplistic approach will provide time transfer to the Glonass system time additional errors occur which must be corrected for or an allowance made for them in the error budget.

1) Position errors. Both the transmitted satellite ephemeris and the user's known location can contain errors which appear as biases in the measurements. This is particularly relevant in the case of Glonass as the co-ordinate system reference frame which is used is at present unknown and thus it would prove more profitable to solve for user position as well as time offset to remove position uncertainties.

2) Atmospheric delays. The transit time of the signals are affected by delays in the troposphere and ionosphere. Tropospheric delay can be minimised by selecting satellites at high elevation angles and can also be accurately modelled. The ionospheric delay is usually the largest error in time transfer. Models also exist for this, such as that used by the Navstar GPS, but are generally less accurate. The Navstar model results in a rms reduction in range error of 60 percent [6]. Ionospheric effects can be removed most effectively by making dual frequency measurements.

3) Errors result in the calculation of signal propagation time due to the rotation of the earth, and hence a static user, during the signal propagation time. This is a function of satellite position and user latitude. For the worst case of a user at the equator and a satellite due east or west on the horizon this can result in an error of approximately 128 nanoseconds in the case of Glonass. The earth rotation error can be easily calculated and removed.

4) Receiver noise and biases. Receiver noise with zero mean can be removed by simple averaging, which can be readily applied for a stationary user. Biases such as receiver delay can also be subtracted

from the measurements.

5) Other errors exist such as imprecision in satellite clock correction parameters and affects such as multipath.

SIGNAL TIMING REFERENCES

Timing references are contained in both the satellite code and transmitted data. The Glonass C/A code is a 511 bit maximal length sequence transmitted at a rate of 511 Kbits/s giving a code epoch every millisecond. The Glonass data message is represented as 50 baud data modulated on to the satellite code. The data transitions are co-incident with code epochs. The data is transmitted as 2.5 minute superframes, each superframe is divided into 5, half minute, frames and then each frame is additionally divided into 15, two second, lines. Each frame contains the current time, satellite ephemeris, clock correction parameters and almanacs for five other satellites. Hence five frames are required to obtain all the almanacs. The 1 second epochs occur in the data at the beginning (even second) and middle (odd second) of each line. Figure 2 shows the content of one data frame.

GLONASS DATA MESSAGE

SATELLITE EPHEMERIS

Glonass ephemerides are represented by the satellites ECEF position and velocity vectors as well as acceleration correction components. In general the ephemerides are updated half hourly giving a maximum ephemeris extrapolation period of 15 minutes. To calculate the satellite position the equations of motion of the satellite can be numerically integrated over the prediction period.

$$\ddot{\underline{r}} = -\frac{\mu}{r^3}\underline{r} + \nabla\left[\frac{3}{2}J_2\mu\frac{R_\oplus^2}{r^3}\left(\frac{1}{3} - \frac{z^2}{r^2}\right)\right] \qquad (1)$$

Equation (1) gives the equation of motion of the satellite with a correction for the second zonal harmonic (J_2) of the earth's oblateness which is by far the largest perturbation on the Glonass satellites over the ephemeris validity period. The equation of motion can be expressed in it's constituent components by equations (2) to (4) [7].

$$\ddot{x} = -\frac{\mu}{r^3}x - \frac{3J_2\mu R_\oplus^2 x}{2r^5}\left(1 - \frac{5z^2}{r^2}\right) \tag{2}$$

$$\ddot{y} = -\frac{\mu}{r^3}y - \frac{3J_2\mu R_\oplus^2 y}{2r^5}\left(1 - \frac{5z^2}{r^2}\right) \tag{3}$$

$$\ddot{z} = -\frac{\mu}{r^3}z - \frac{3J_2\mu R_\oplus^2 z}{2r^5}\left(3 - \frac{5z^2}{r^2}\right) \tag{4}$$

These equations can be integrated by any suitable technique (Runge-Kutta 4th order for example) and with a suitable step size the satellite's position can be easily calculated to within 3 metres over half an hour. It is expected that the acceleration terms in the data message will correct for additional perturbations to the satellite which are predominantly luni-solar in origin.

SATELLITE CLOCK OFFSETS

The satellite clock offset from the common Glonass system time is represented by two parameters [2].

1) γ_n - the relative frequency offset between the nth. satellite navigation signal frequency, f_n, and the nominal value, f_{hn}, of the nth. satellite frequency.

$$\gamma_n = \frac{f_n - f_{hn}}{f_{hn}} \tag{5}$$

2) τ_n - the nth. satellite time scale shift relative to the Glonass time scale.

The Glonass system time, t_{sys}, is related to the satellite time, t_{sv}, by,

$$t_{sys} = t_{sv} + \tau_n - \gamma_n(t - t_o) \tag{6}$$

$$\dot{\tau}_n = -\gamma_n \tag{7}$$

where t_o is the time of validity of τ_n and γ_n.

Table 2 shows the range and resolution of the Glonass clock correction parameters.

Glonass	Bits*	Scale	Range	Resolution	Units
τ_n	22	2^{-30}	$\pm 1.953 \times 10^{-3}$	9.313×10^{-10}	s
γ_n	11	2^{-40}	$\pm 9.304 \times 10^{-10}$	9.095×10^{-10}	s/s

* MSB = sign bit.

Glonass clock correction parameters.

Table 2

GLONASS TIME TRANSFER MEASUREMENTS

A series of measurements have been conducted of the difference between UTC(USNO) and Glonass system time. The arrangement of equipment to carry out these measurements is shown in figure 3. A prototype single channel Glonass/Navstar GPS receiver has been constructed [8] which allows time comparisons between Glonass or Navstar system time and a 1 pps reference synchronised to UTC(USNO) available from a commercial Navstar receiver. The Navstar system time / UTC(USNO) comparison is used as a calibration and confidence measurement since the offset between GPS time and UTC(USNO) is known and transmitted as part of the GPS data message. The measurements are conducted as follows.

A time interval counter measures the interval between the UTC(USNO) 1 pps and millisecond epochs decoded from the code generator of the test receiver. Whilst tracking a satellite time interval measurements are thus made of UTC(USNO) against the clock of the satellite currently being tracked but also including the signal propagation time. This measurement is then related to UTC(USNO) against system time by the following equations.

$$\rho = \text{UTC(USNO)} - t_{sv} + t_p \qquad (8)$$

where, ρ = counter reading.

t_{sv} = satellite time.

t_p = signal propagation time (modulus 1 ms)

$$t_{sv} = t_{sys} - \delta t \tag{9}$$

where, t_{sys} = system time.

δt = difference in system time and satellite time.

$$\rho = UTC(USNO) - t_{sys} + \delta t + t_p \tag{10}$$

$$UTC(USNO) - t_{sys} = \rho - t_p - \delta t \tag{11}$$

The resolution on each measurement is 2 ns and the UTC(USNO) 1 pps is accurate to within 100 ns. Measurements are made once per second, averaged over 3 minutes and then the data is stored.

DATE	SATELLITE	READINGS (1/SEC)	AVERAGE OFFSET/ns	STANDARD DEVIATION/ns
12/3/89	NAVSTAR 3	4320	-336	22
12/3/89	NAVSTAR 6	2520	-325	18
12/3/89	NAVSTAR 9	2340	-327	23
12/3/89	NAVSTAR 11	5580	-336	19
12/3/89	NAVSTAR 12	3420	-330	14
12/3/89	NAVSTAR 13	4990	-352	16
12/3/89	GLONASS 34	3960	29696	24
12/3/89	GLONASS 35	4860	29700	24
12/3/89	GLONASS 36	4320	29699	26
12/3/89	GLONASS 38	4320	29705	35
12/3/89	GLONASS 39	3780	29703	35
12/3/89	GLONASS 40	4140	29740	24
12/3/89	GLONASS 41	4680	29713	19

Navstar and Glonass system time offset from UTC(USNO).

Table 3.

Table 3 shows a set of measurements over a typical 24 hour period. The data has been corrected for tropospheric, relativistic and earth rotation effects but not for ionospheric effects. Since the measurements are taken over one day then the data is obtained from two passes of each satellite. An elevation mask of 10 degrees is used. The offset of UTC(USNO) from Navstar system time over this period as transmitted by the Navstar data message is about -344 ns.

Figure 4 shows a plot of UTC(USNO) against Glonass system time over a period of about 8 months together with a plot of UTC(USNO) - UTC(SU) over the same period using data published by Bureau International des Poids et Measures. As can be seen, Glonass system time appears to be a continuous and reliable time scale over this period.

REFERENCES TO OTHER TIME SCALES

An additional clock correction parameter is included in the Glonass data message which relates the Glonass system time scale to the time scale at which ephemeris and satellite clock offsets are calculated. We will call this new time scale Glonass ephemeris time and the offset between Glonass ephemeris time and Glonass system time is denoted by the parameter τ_c. Observations of τ_c over a period in excess of two years shows the offset between the time scales always maintained within approximately $\pm 30\mu s$ of each other. Under normal circumstances application of this parameter in the calculation of satellite location and clock offset is unnecessary since a nominal satellite velocity of 4 km/s and clock frequency offset of the order of 10 ps/s will provide insignificant corrections when changing from Glonass system time to ephemeris time. Figure 5 shows a plot of τ_c over the same time period as figure 4. It can be seen that τ_c provides a somewhat less continuous time scale than Glonass system time. It is unclear at the present time as to whether Glonass ephemeris time is related to or influenced by other external time scales such as UTC(SU).

CONCLUSIONS

It has been demonstrated that time transfer with Glonass to a static user can be achieved to accuracies of the order of 100 ns or better while using the Glonass low precision C/A code phase and without dual frequency measurements. Results are repeatable amongst all the active satellites and a cross calibration with Navstar GPS provides a high confidence level. Glonass system time provides a time scale which has reliability and

performance comparable to that of other international time scales and can be used as an intermediate time scale for clock comparisons between widely separated sites.

It should be pointed out that if the proposed Selective Availability is employed by the DOD on all Navstar GPS satellites, limiting accuracies to the civil user to approximately 100 m (300 ns), then Glonass will supersede Navstar in position and time capabilities to the civil user.

No direct linking has yet been confirmed between Glonass time scales and other external time scales such as UTC(SU) though such a link would of course be of value to the timing community.

REFERENCES

[1] A J Van Dierendonck and W C Melton :
"Applications of Time Transfer Using Navstar GPS", The Institute of Navigation, Washington D.C., U.S., Special Issue on GPS, Volume 2, pp 133-146.

[2] T G Anodina :
Working Paper -
"Global Positioning System GLONASS", Special Committee on Future Air Navigation Systems (FANS/4), International Civil Aviation Organisation (ICAO), Montreal, 2-20 May, 1988.

[3] S A Dale and P Daly :
"Recent observations on the Soviet Union's Glonass Navigation Satellites", IEEE PLANS' 86 (Position Location & Navigation Symposium), Las Vegas, 4-8 November, 1986, pp 20-25.

[4] S A Dale and P Daly :
"Developments in Interpretation of the Glonass Navigation Satellite Data Structure", IEEE NAECON'88 (National Aerospace & Electronics Conference), Dayton, Ohio, 23-27 May, 1988, pp 292-297.

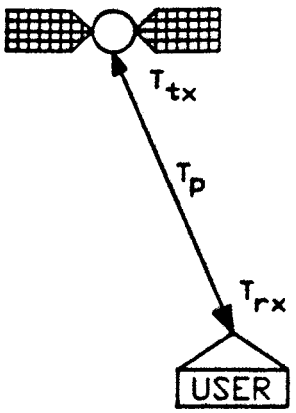
[5] S A Dale, P Daly and I D Kitching:
"Understanding Signals from Glonass Navigation Satellites", International Journal of Satellite Communications (John Wiley), Vol. 7, 1989, pp 11-22.

[6] W A Fees and S G Stephens:
"Evaluation of GPS Ionospheric Time Delay Algorithm for Single-Frequency Users", IEEE PLANS' 86 (Position Location & Navigation Symposium), Las Vegas, 4-8 November, 1986, pp 206-213.

[7] F T Geyling and H R Westerman:
"Introduction to Orbital Mechanics", Bell Telephone Laboratories, Addison-Wesley Publishing Company, 1971.

[8] S A Dale, I D Kitching and P Daly :
"Position-Fixing using the USSR's Glonass C/A Code", IEEE PLANS' 88 (Position Location & Navigation Symposium), Orlando, 29 November - 2 December, 1988, pp 13-20.

TIME TRANSFER LINK



T_{tx} = time of epoch transmission, satellite time

T_{rx} = time of epoch reception, satellite time

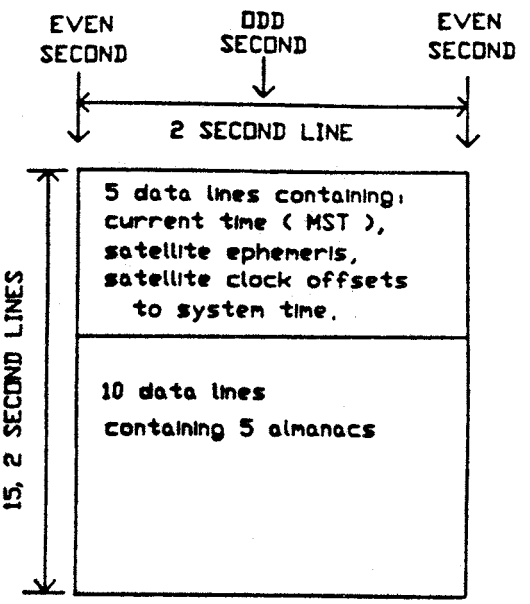
T_p = signal propagation time

$T_{rx} = T_{tx} + T_p$ satellite time

$= T_{tx} + T_p + \Delta T$ system time

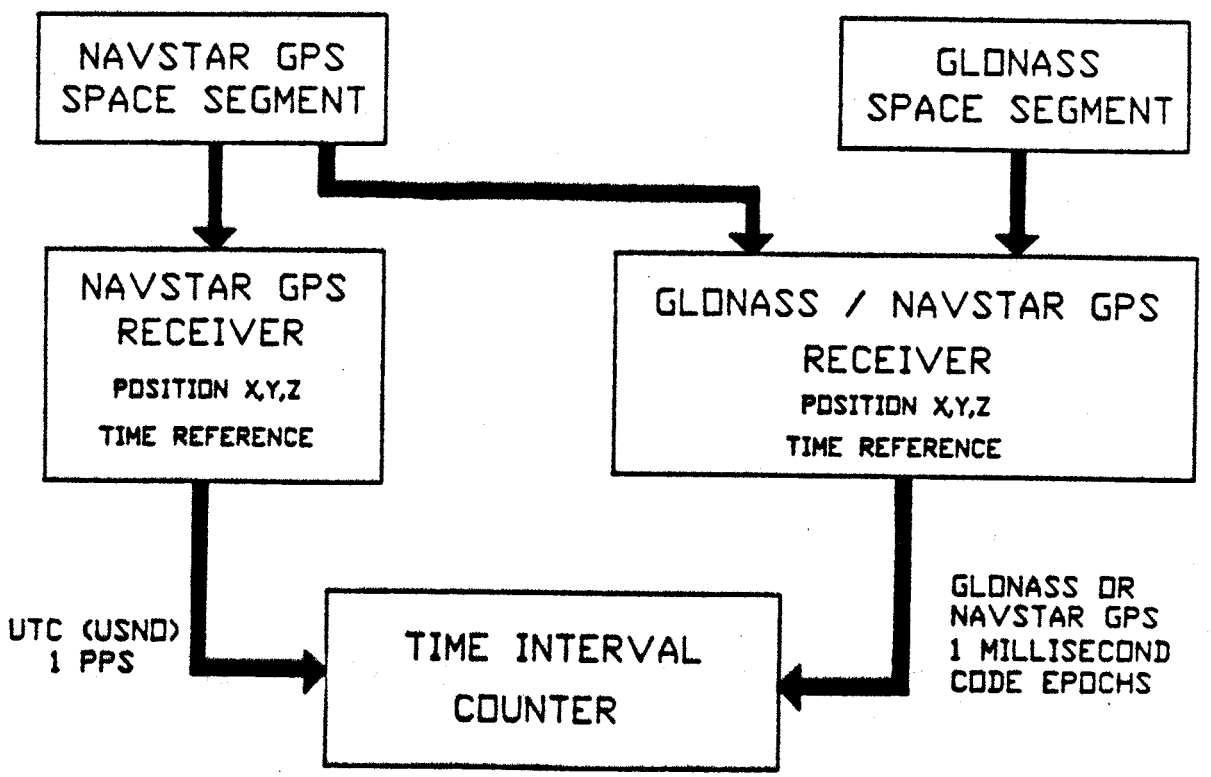
ΔT = system to satellite time offset

FIGURE 1.



DATA FRAME FORMAT

FIGURE 2.



TIME TRANSFER EQUIPMENT ARRANGEMENT
FIGURE 3.

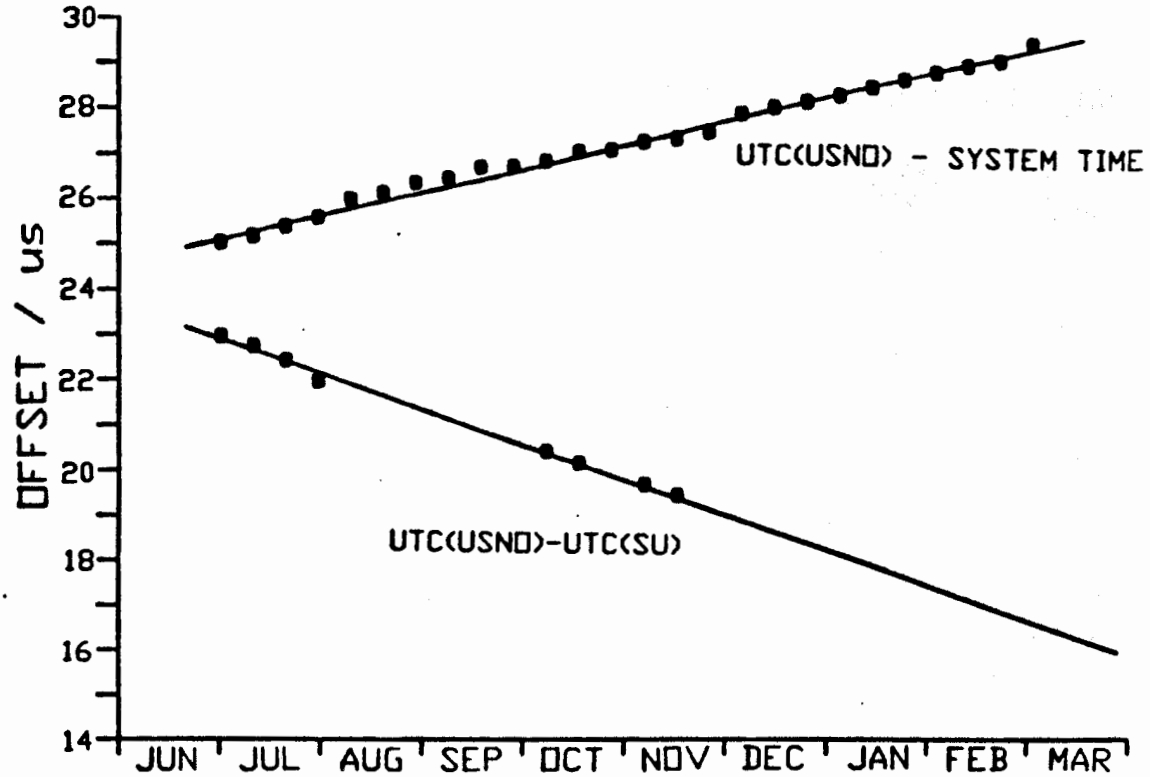


FIGURE 4.

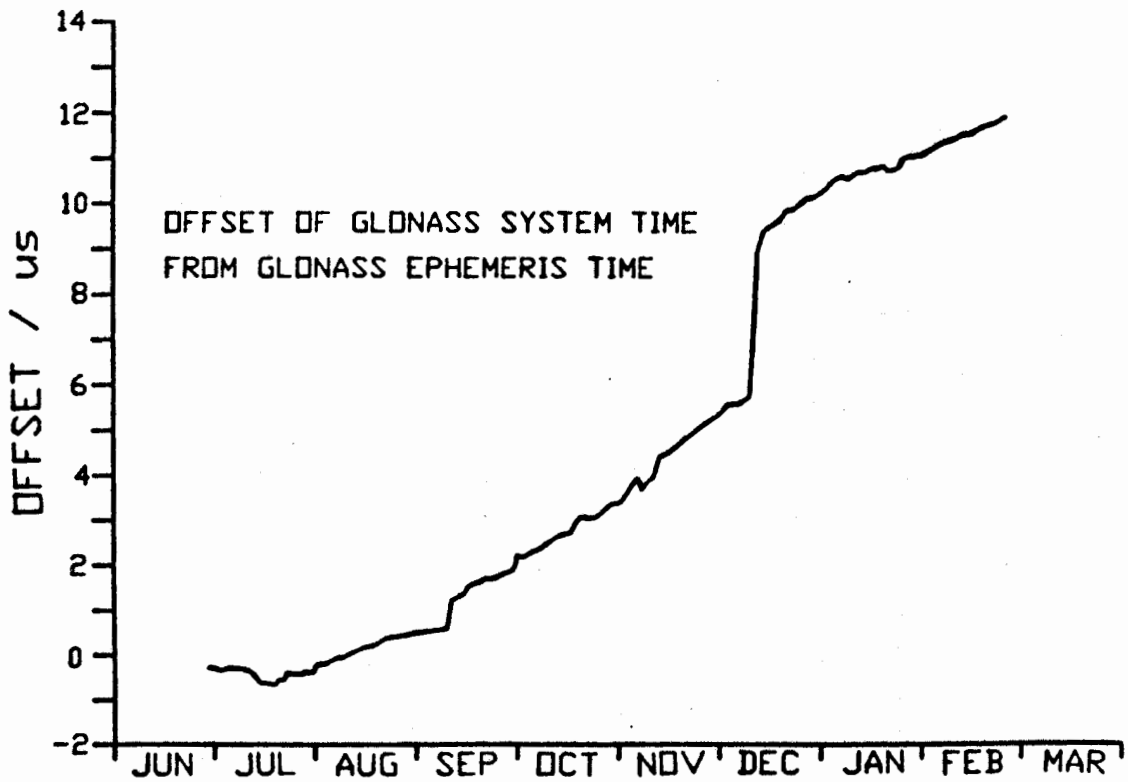


FIGURE 5.

CRYSTAL PARAMETER MEASUREMENT USING A HEWLETT PACKARD HP4195A NETWORK/SPECTRUM ANALYZER

P.E. Morley and R.J.T Marshall, S.T.C. Components, Quartz Crystal Division, Harlow, Essex, U.K.
R.J. Williamson, S.T.C. Technology Ltd., Harlow, Essex, U.K.

INTRODUCTION

The object of this project was to generate a suite of computer programs for use in a production or test environment which could give a substantially complete characterisation of a crystal resonator using a single instrument. This would have been inconceivable a few years ago, but with recent advances in commercially available network analyzer systems it has become a practical proposition.

Several suitable instruments exist, but the one used for this work is a Hewlett Packard HP4195A Network/ Spectrum Analyzer. Apart from the conventional network analysis capability, the main desirable feature of the analyzer is its ability to perform a direct impedance measurement while sweeping either the source frequency or amplitude, and then to display the information in a variety of different formats. This simplifies the manual measurement of crystals and greatly facilitates the development of automatic measurement techniques.

The first technique determines the equivalent circuit parameters of a crystal resonator using a circle-fitting routine performed in the admittance plane. This has been described in detail in a previous paper [1]. Some enhancements have been incorporated which improve both the measurement of very high Q resonators, and also the calibration procedure. The extension of the method to very high frequencies is also discussed.

The other two aspects of resonator characterisation considered here are the measurement of spurious responses, defined in various ways, and the variation of the equivalent circuit values, particularly the Q, with drive level.

ADMITTANCE CHARACTERISTICS OF A CRYSTAL RESONATOR

The generally assumed linear equivalent circuit of a single resonance mode of a bulk wave crystal resonator is shown in fig. 1 [2]. This provides an excellent approximation in most practical cases, particularly at lower frequencies, and treats the resonator as a two-terminal network.

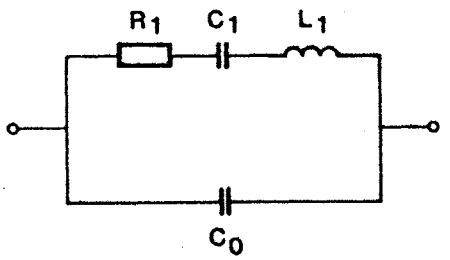


Fig. 1 Simple crystal equivalent circuit

There has been much discussion concerning the deviation of devices from this impedance model, particularly at VHF. This is partly due to the small, but potentially significant, series inductance and resistance in the crystal leads and in the conductive adhesive spots used in mounting the crystal blank [2,3,4]. In the narrow band these series impedances can be modelled adequately by including a parallel conductance in the equivalent circuit as shown in fig. 2. [5].

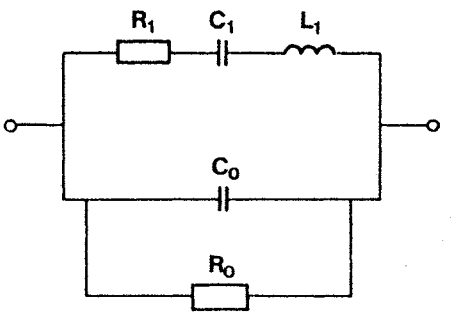


Fig. 2 Crystal equivalent circuit with shunt resistance

Clearly any discussion on the linear equivalent circuit, or even on the admittance, of a crystal assumes linearity in the device. However, the presence of coupled resonance modes can result in major deviations from the equivalent circuit, and their coupling to the wanted mode can be either linear or non-linear. Close-in inharmonic resonances can cause similar problems and the main resonance itself often exhibits non-linearity which can be caused by several different mechanisms. The measurement of the circuit parameters should therefore include an assessment of the quality of fit to the simple linear model.

Another source of measurement ambiguity occurs when the crystal has a third terminal, normally the case connection of a metal holder. This is typically of concern for crystals used in filter applications and for those which are specified with a series load capacitance [6]. In particular the value of static capacitance depends on the way the third terminal is connected. It can be either a) floating or b) connected to one device terminal or c) independently grounded, and this must be clarified before the measurement is made.

MEASUREMENT OF THE EQUIVALENT CIRCUIT PARAMETERS

The crystal measurement itself can be divided into two essentially separate sections; the method of impedance measurement, which includes (a) the type of fixture and (b) the method of data acquisition; and (c) the subsequent manipulation of the impedance data.

a) Measurement Jig

Many different types of jig for impedance measurement have been suggested in the past, each with its inherent advantages and disadvantages [6,7]. For some years now the IEC 444 pi-network has been used by most crystal manufacturers and their customers. A significant practical advantage of this type of network is the use of spring-loaded knife edges for ease of insertion of and connection to the crystal. The original IEC 444 specification went to great lengths to eliminate the reactive parts of the jig by tuning out the inductance of the resistors with capacitors. This, however, is only effective up to about 125 MHz.

The HP4195A instrument offers a possible alternative in its direct impedance measurement facility which uses a 50 ohm reflection bridge as a network. This has potential advantages in resolution, higher available drive level and the ability to calibrate the system with precision coaxial terminations which can be referenced to national standards. However, the absence of a CW mode in the instrument and the excessively slow sweep speed and stabilisation time in the impedance mode make this approach difficult to implement. Also, the spring clip fixture provided with the impedance adaptor is not ideally suited for connection to standard crystal packages.

For these reasons, the initial implementation of the method uses a thick-film type pi-network and the instrument is configured as a conventional network analyzer. The pi-network can be successfully used above 100 MHz if its parameters are accurately known. The admittance of the crystal can then be calculated from the voltage transfer function [1].

Thevenin's theorem shows that any such network can be simplified into an image EMF (E) and a series source impedance together with a load impedance (Fig. 3). The required jig parameters are the total jig impedance ($Z_j = R_j + X_j = Z_1 + Z_2$), the voltage transfer factor (K) and the shunt capacitance (C_j) across the measurement terminals. These can be obtained by measuring three known impedances in the jig at the nominal crystal frequency. In practice the impedances used are a short-circuit, a 50 ohm precision resistor and an open-circuit.

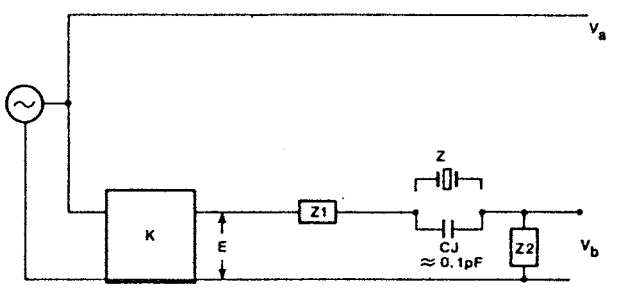


Fig. 3 Equivalent Circuit of Jig

The open circuit measurement consists simply of measuring the empty jig. The other two measurements, however, need some thought. The inductance of both the jig and calibration impedance will depend on the current path through them and this in turn primarily depends on the distance between the contact points. The known impedances should therefore have the same contact separation as the crystals. The short-circuit can be made from a sheet of gold-plated metal with ridges at the required separation Typically this will have an inductance of less than a nanohenry. The 50 ohm impedance can be made from a precision coaxial termination such as SMA or APC 3.5 and a simple modification made to the appropriate socket to provide connection to the knife-edges whilst maintaining the impedance value.

While it is possible to write down the equations relating the impedances to the jig parameters, these are unnecessarily cumbersome, and experience has shown that a simpler procedure is adequate.

Generally $V_b = [Z_2/(Z_j + Z)]E$ (1)

but $E = KV_a$ (2)

so that

$$V_b/V_a = V_{ba} = [Z_j/(Z_j + Z)](Z_2/Z_j)K \quad (3)$$

For a true short (Z=0), the above can be rewritten as

$$V_{ba} = Z_j/(Z_j + Z)\,V_{bas} \quad (4)$$

where the phase is referenced to zero and V_{bas} is the voltage ratio with the short connected. It can be assumed that the effect of the inductive short is negligible in the amplitude domain and the small phase offset which occurs can be calculated once the jig impedance is known.

The jig impedance is calculated by inserting the 50 ohm termination and measuring the voltage ratio V_{ba}

$$Z_j = 50/(V_{bas}/V_{ba} - 1) \quad (5)$$

Finally the jig capacitance is measured by removing the 50 ohm termination. In practice the signal will be very low and the phase measurement will be in error. It is best therefore to assume a pure capacitance and calculate its value from the amplitude.

Once the value for the jig impedance is known, the phase offset from the inductance of the short can be calculated. As this will slightly alter the phase of the 50 ohm measurement and hence the jig impedance it is necessary to iterate around the loop several times.

b) Data acquisition

Using a sophisticated network analyzer, there are clearly many options for obtaining the admittance characteristic with respect to frequency. Some methods use the sweep facilities of the instrument, and then select the appropriate data at a later stage [7]. A disadvantage of this approach is that a slow sweep speed is required to prevent distortion of the observed transfer characteristic for high Q resonators. A different technique which is highly appropriate for crystal resonators is used in commercially available equipment produced by the Transat Corporation. In this dedicated instrument an analogue phase-locked loop is used to step around the resonance by locking to pre-selected phase values. The required frequency information is derived from a counter, and the data are then processed to obtain the crystal component information.

The method used in this work, however, performs admittance measurements at discrete frequencies as predicted by previous data. The unknown admittance can be determined by a rearrangement of equation 4:

$$1/Y = Z = Z_j(V_{bas}/V_{ba} - 1) \quad (6)$$

After any transition in frequency, the stability of the measured impedance Z is determined by taking several successive measurements. A measurement is only considered valid when the stability has attained a pre-determined value.

The initial design of this program was based around the 200 MHz HP3577A Network Analyzer, and this was relatively straightforward. The introduction at STC Quartz Crystal Division of a high frequency fundamental product line prompted the use of the newly introduced

HP4195A Analyzer to extend the upper frequency limit to 500 MHz. Some aspects of the higher frequency instrument were, however, found to cause some difficulty in transferring the data acquisition routines.

Firstly, there is a discrete step of about 1dB in the instrument's oscillator output level at exactly 10 MHz, which is generally undesirable and causes particular problems in the measurement of devices around 10 MHz which exhibit any non-linearity. Similar discontinuities occur at frequencies above 130 MHz, resulting in similar problems.

Unlike the HP3755A, the new analyzer has no 'CW' mode, which allows fast measurements at a single frequency. To overcome this the frequency sweep is set up to give its minimum of three points and a frequency span of 0 Hz. This emulates the 'CW' mode and as the extra data points can be used to determine the stability very little is lost in terms of measurement time. Also the settling time after changing the centre frequency (or any other stimulus parameter) is excessively high when using the internal correction, so all normalizing is performed by the external computer.

c) Data Manipulation

In essence the measurement process consists of finding the crystal resonance, taking some measurements around the resonance, fitting a circle to the admittance data and then fitting a linear function to the rate of change of frequency around the circle.

An admittance equation can be derived from the simple crystal equivalent circuit of fig. 1 which is of the form:

$$(G - 1/2R_1)^2 + (B - 1/\omega C_0)^2 = (1/2R_1)^2 \quad (7)$$

For quartz crystals, the Q is very high and hence ωC_0 is essentially constant over the resonance. The above equation thus describes a circle in the admittance plane. The radius (r) is $1/2R_1$ and the centre coordinates (x_0,y_0) have values of $1/2R_1$ and ωC_0. The effect of the shunt capacitance is to move the circle along the B axis. In like manner, it can be shown that a small amount of series resistance in the narrow band approximation can be transformed into a shunt resistance which simply moves the circle along the G axis (Fig. 4). Thus by performing a circle curve fit on the admittance data of a crystal resonance, it is possible to calculate C_0, R_0 and R_1. Furthermore, by subtracting G_0 and B_0 from the data, only the motional arm (L_1, C_1, R_1) is left and the values can be calculated from the variation of frequency around the circle.

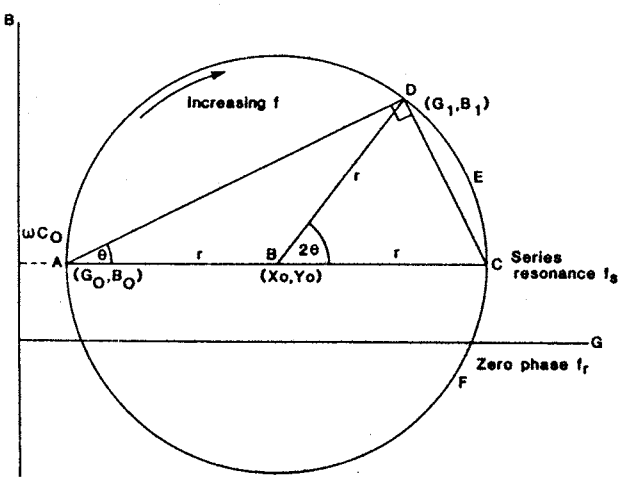


Fig. 4 Admittance circle of crystal

The phase value θ in fig. 4 can be shown to be related to the frequency f by the relationship:

$$\tan(\theta) = 2Q(f_s - f)/f_s \quad (8)$$

This is a linear relationship with a gradient of Q but the value of f_s is unknown. If f_{ref} is chosen close to resonance, then

$$\tan(\theta) = a(f_{ref} - f) + b \quad (9)$$

where $a = 2Q/f_s$ (10)
and $b = 2Q(f_s - f_{ref})/f_s$ (11)

Theoretically, the centre of the circle and its radius give the shunt capacitance and resistance. This assumes that the resonance is linear and that other elements in the equivalent circuit such as series inductance are negligible. Non-linearity causes distortion of the circle, and other factors, such as temperature variation, cause movement of the resonance in frequency. It is safer, therefore, to measure the shunt elements off resonance. Typically this is measured at 10% below the expected frequency, but to improve the measurement accuracy of C_0 for low frequency crystals, the shunt capacitance is measured at 30 MHz on crystals whose resonance is below 27 MHz. In either case several measurements are made close to this frequency to ensure the absence of spurious resonances.

The effects of non-linearity will be reduced if amplitude variations are limited. This is achieved by taking measurements symmetrically around the series resonance. This makes sense since this is the frequency at which the crystal is likely to be used. The method adopted is to use the off-resonance measurement together with two others close to the resonance to calculate the crystal parameters approximately. These can then be used to predict the frequencies which will place the subsequent measurements at regular intervals around the admittance circle. One of the advantages of this method is its ability to deal with high Q crystals which have a tendency to "ring" by checking for impedance stability after each frequency step. It also enables averaging of the data after stabilisation.

EXTENSION TO HIGHER FREQUENCIES

Recent developments in crystal resonator technology have resulted in a progressive increase in the achievable upper frequency limit [8,9]. As the device frequency increases, so the impedance of the static capacitance C_0 decreases. Any series resistance can then become a significant factor in the impedance model. Also the impedance of any inductive series component increases in value with an increase in frequency, and this can also become significant.

Although the simple model of fig. 2 has been used to obtain parameter measurements on devices above 200 MHz (table 1), their validity should be regarded as somewhat suspect. Moreover, as has been observed at the recent IEC TC49 WG6 measurement workshop [10], different fitting techniques to the same impedance model can result in significant differences in measured parameter values. It is clearly necessary to adopt internationally an equivalent circuit which has been optimised for high frequency devices.

The next phase of the program design is now under way to improve the measurement above 200 MHz. The implementation will utilize the direct impedance measurement facility of the HP4195A, although it is considered that the spring clip fixture requires some modification to minimise the crystal lead length, whilst maintaining a reasonable impedance match to the measurement port. The method will use a frequency sweep to obtain data, as there is only a marginal constraint on the sweep speed due to the resonator Q at these frequencies. An initial wide-band measurement will be performed to characterize the C_0 and the series

components, and then a second narrow-band sweep will determine the motional parameter values.

STC Components Ltd
Quartz Crystal Division

Date: Tue 7 Mar 1989 Operator: pem
Time: 12:20 Spec no:

Crystal code	fs (kHz)	drive (uW)	R1 (ohms)	C1 (fF)	Q	C0 (pF)
33/fund	100000.738	95.80	21.89	2.65	27431	1.02
33/fund	100000.746	95.80	21.90	2.65	27440	1.02
33/fund	100000.743	95.80	21.89	2.65	27413	1.02
33/fund	100000.744	95.80	21.90	2.65	27412	1.02
31/fund	100001.665	96.14	27.02	2.74	21457	1.06
31/fund	100001.671	96.14	27.02	2.74	21468	1.06
31/fund	100001.673	96.14	27.01	2.75	21463	1.06
31/fund	100001.674	96.14	27.01	2.74	21467	1.06
25/3rd	299965.393	68.85	102.48	.19	27533	1.27
25/3rd	299965.395	68.87	102.42	.19	27525	1.27
1/fund	139264.671	100.36	36.96	2.19	14119	.88
1/fund	139264.671	100.36	36.96	2.19	14118	.88
3/fund	139262.832	103.60	28.44	2.49	16136	.95
3/fund	139262.828	103.61	28.41	2.49	16138	.95
12/fund	124666.226	96.25	26.62	3.77	12722	1.20
12/fund	124666.225	96.27	26.47	3.76	12825	1.20
7/fund	124981.684	94.18	18.50	3.49	19732	1.15
7/fund	124981.688	94.26	18.61	3.49	19596	1.15

Table 1. typical parameter values for high frequency fundamental resonators

MEASUREMENT OF SPURIOUS RESPONSES

The fundamental mode and its overtones each have an associated family of unwanted responses. These typically occupy a narrow frequency band immediately above the major resonance and are usually of higher impedance. The most appropriate method for measuring these resonances depends on the intended resonator application.

For crystal oscillators, the series resistance of the spurious response is of greatest concern, particularly in the severe case where mode hopping can occur, ie where the oscillator loop gain is greater than unity for two or more modes.

In filter applications, the main concern is the frequency of the spurious modes. The influence of spurious impedance level on the filter stop-band is complex, but its effect is often estimated by inserting the resonator into a fixture of known series impedance and testing the difference in response level in dB between the wanted and spurious modes. The fixture typically contains a circuit for C_0 cancellation, usually in the form of a balanced hybrid coil [11].

The measurement program is capable of characterising the spurious modes in several ways, in an attempt to cover these variations in specification methods. The four main options are as follows:

a. Level in dB relative to main response.
b. Level in dB relative to short circuit.
c. Resistance value.
d. Resistance ratio relative to main mode.

The user is always prompted for the crystal frequency and the frequency range over which the search for resonances will be performed. When necessary, the user is also prompted for fixture impedance, and is given the opportunity of tuning the jig for cancellation of shunt capacitance.

The technique for searching for significant spurious modes simply utilises the facilities of the instrument to display the amplitude of the transmission response in the frequency domain, and to search for maxima. This process is not entirely trivial, mainly because of the finite number of data points used by the analyzer for determination of the frequency response. For resonances with high Q factors, it is clearly possible for the maximum transmission points to lie between sample frequencies, and to result in significant amplitude errors. For this reason, a multiple-sweep system is employed in which the frequencies of the most severe spurious modes are estimated using the widest sweep range, and then the resonances are more closely examined with later measurements using a narrower span. This process is repeated until there is no ambiguity in measurement.

MEASUREMENT OF DRIVE LEVEL DEPENDENCE

The presence of non-linearity in crystal resonators can be manifested in several forms, some of which have been known for many years [2]. The deviation from linearity, which is almost always undesirable, appears to be divided into two main categories: bulk inelastic effects, which typically occur at high levels of drive, and surface phenomena, which normally occur at low current levels. Bulk non-linearity results in intermodulation distortion in crystal filters, in non-linear mode coupling resulting in anomalous resonance behaviour over temperature, and in the well-known distortion of the amplitude/ frequency curve at high crystal current [12,13]. Surface non-linearity can also result in forms of intermodulation [14], and in high levels of close-in phase noise in crystal oscillators, but its main effect is high resonator starting resistance, sometimes called 'sleeping sickness' or second level of drive (SLD) [15]. The primary intention of this measurement method was to detect the presence of a large drive level dependence of the latter type in resonators for application in crystal oscillators.

Several methods have been used for the measurement of sensitivity of crystal frequency and/or resistance to drive current. Clearly, with computer-controlled impedance analyzers it is possible to perform a full measurement of frequency and resistance at previously selected discrete drive levels [16]. While this is a highly accurate method, it is slow, and limited in the number of oscillator levels which can be used. Another widely used technique is the ramp method [15], in which each crystal is placed in an oscillator circuit where the initial crystal current is extremely low. The current is then ramped to a maximum level, and then reduced back to the initial level. The graphical display of detected oscillator output level against the ramp level provides a clear indication of the presence of resistance anomalies, but is rather difficult to interpret, and is also difficult to relate to national standards. A third method is also described [15] which utilises a matched pair of steppable attenuators situated before and after the crystal network in an otherwise conventional vector-voltmeter based measurement system. This offers the advantage of frequency and resistance measurement, and has revealed examples of highly peculiar crystal behaviour.

The relatively recent introduction of highly versatile network/spectrum/impedance analyzers, in particular the Hewlett Packard HP4195A, has greatly facilitated the development of techniques for this type of measurement. The instrument is capable of ramping either the frequency, the d.c. bias level or the R.F. signal level whilst performing an impedance measurement. The measurement network is a 50 ohm reflection bridge, and the impedance can be displayed in all the appropriate formats (with the exception of a Smith Chart, a curious omission by HP).

To perform the measurement on this instrument, it is first set to impedance measurement mode, and is calibrated with the standard impedances, which are provided with the system, over a wide frequency range. For each crystal the measurement sequence is then as follows:

a. In the rest state before a crystal is inserted, the amplitude is set to the minimum possible value. The

display is set to |Z|/θ, and the stimulus is set to ramp in the frequency domain. The frequency range is centred on the nominal crystal frequency, with a default span of 4000 ppm, a value which can be changed if necessary by the user.

b. After insertion of the resonator, the marker is set to the minimum impedance value. The frequency span is then reduced by a factor of 50 and the process is repeated, thus obtaining a reliable estimate of the resonant frequency.

c. The display format is changed to R/X, since the resistance component of the impedance around a resonance is fairly insensitive to the proximity to the resonant frequency, particularly for lower frequency devices (fig. 5). The instrument is then instructed to ramp the oscillator level, typically from -56 dBm to +4 dBm into 50 ohms, at the frequency as determined by the first sweep measurement. Although the display is not truly continuous, the number of data points per sweep can be set to as high as 101. In practice this must be performed in three passes, due to a limitation in the instrument of 26 dB in power level per sweep, giving a total of 303 discrete power levels.

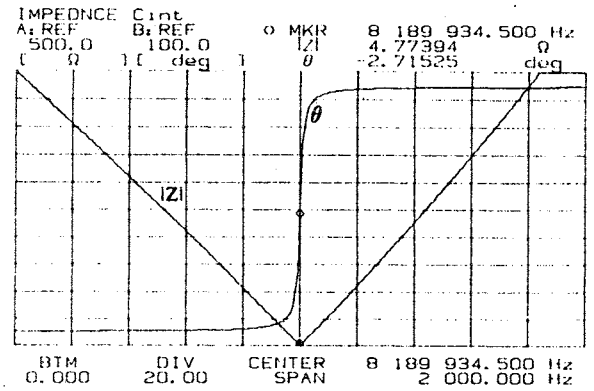


Fig. 5a Plot of resonator impedance in |Z|/θ format

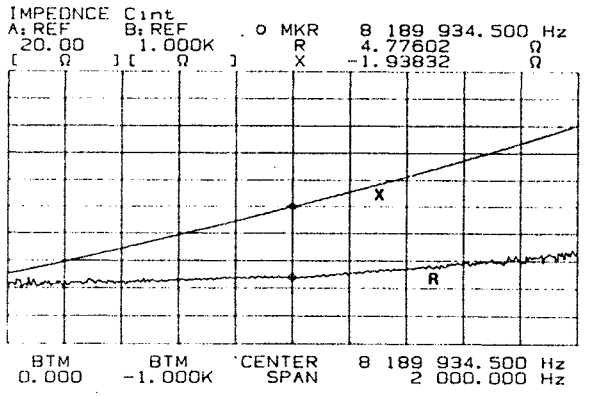


Fig. 5b Plot of resonator impedance in R/X format

d. The instrument display is then instructed to show the resistance value against drive level, and gives an excellent illustration of deviations from linearity. The maximum and minimum resistances are available from this display, and can be tested against previously defined limits, for example, those set out in the recent proposal for IEC 444. Typical examples of plots for good and bad devices are shown in figs. 6a and 6b respectively.

As the source frequency is in no way locked to the crystal resonance, the main inherent error in this method is due to possible large changes in crystal resonant frequency with drive level, which may result in the applied signal frequency significantly deviating from the minimum impedance point. This will result in irregular deviations in crystal current with applied signal level. However, this process is not intended for precision parameter measurement but as a method for rejection of units which exhibit poor linearity; it is ideal for this purpose, and can be implemented on 'off the shelf' equipment.

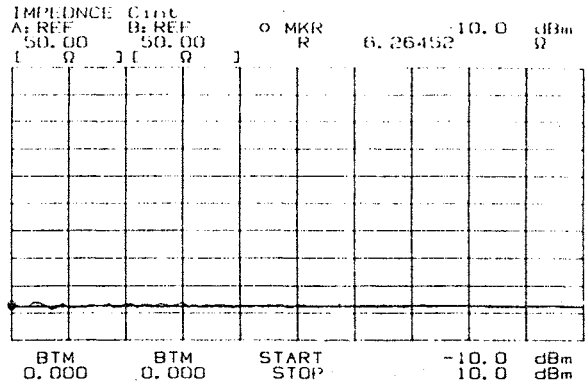


Fig. 6a Plot of real part of impedance of good unit vs drive

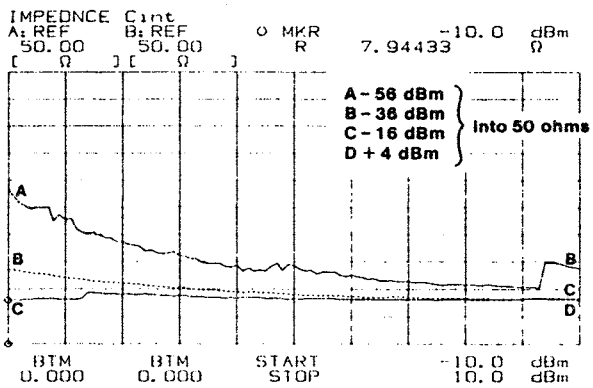


Fig. 6b Plot of real part of impedance of bad unit vs drive

SUMMARY

A suite of computer programs has been developed which utilises a single instrument to perform most of the measurements required for crystal resonators. This includes a precision technique for characterization of the equivalent circuit parameters (which can also be extended for measurements over temperature), a spurious response test and a means of selecting crystals for drive level dependence. The modules are designed for use in a production environment, so speed and simplicity of operation are of the highest importance.

References

1 R.J. Williamson, 'An Improved Method for Measuring Quartz Crystal Parameters', IEEE Trans UFFC-34, No. 6, Nov. 1987.

2 E. Haffner, 'The Piezoelectric Crystal Unit-Definitions and Methods of Measurement', Proc. IEEE, Vol. 57, No. 2, Feb. 1969.

3 F. Neuscheler, 'Schwingquartz-Daten mit Netzwerk-Analysatoren gemessen', Elektronik, 19/18.9.1987.

4 A. J. Flok, 'Equivalent Circuits of One Class of Three-Terminal Resonant Elements', Bulletin de l'Academie Polonaise des Sciences, Vol. XXIX, No. 5-6, 1981.

5 A. J. Fiok and M. R. Wernik, 'A New Approach to Equivalent Circuits of High-Q Resonant One-ports', Bulletin de l'Academie des Sciences, Vol. XXVIII, No. 7-8, 1980.

6 W.H. Horton, 'Measurement Systems for Quartz Crystal Units', 5th Quartz Crystal Conf., Kansas, Sept 13, 1983.

7 R.C. Peach and S.E. Morris, 'A System for Precision Parameter Measurements on Quartz Crystals and Bipoles', Proc. 39th AFCS, 1985, pp 527-534.

8 F.M. Stern, J.Dowsett, D. Carter and R.J. Williamson, 'The Fabrication of High Frequency Fundamental Crystals by Plasma Etching', to be published in the 43rd Ann. Freq. Cont Symp. May 1989.

9 J.R. Hunt and R.C. Smythe, 'Chemically Milled VHF and UHF AT-Cut Resonators', Proc. 39th Ann. Freq. Cont. Symp., May 1985.

10 IEC TC49 Working group 6 draft, 'Assessment of the Quartz Crystal Measurement Data Obtained at the IEC TC49 WG6 Measurement Workshop held in Bled, Yugoslavia on 17th-21st October 1988'.

11 W. H. Horton and R. C. Smythe, 'The Hybrid-Coil Bridge Method of Measuring Unwanted Modes of Vibration in Quartz Crystals', Proc 17th Ann. Freq. Cont. Symp., May 1963.

12 R. C. Smythe and P. E. Morley, 'Experimental Evaluation of the Effective Non-linear Elastic Constant for Trapped Energy and Contoured Resonators', Proc, 39th Ann. Freq. Cont. Symp., May 1985.

13 H. F. Tiersten, 'Analysis of Nonlinear Resonance in Rotated Y-Cut Quartz Thickness-Shear Resonator', Proc, 29th Ann. Freq. Cont. Symp., 1975.

14 W. H. Horton and R. C. Smythe, 'Experimental Investigations of Intermodulation in Monolithic Crystal Filters', Proc. 27th Ann. Freq. Cont. Symp., May 1973.

15 J. S. Yerna, 'Resistance-Measurements of Quartz Crystals at very low Drive Levels', 38th Ann. Freq. Cont. Symp., 1984.

16 R. C. Smythe, 'An Automated Resonator Measurement System using a Reflection Coefficient Bridge', 35th Ann. Freq. Cont. Symp., May 1981.

NINE-YEAR AGING BEHAVIOR OF THE CERAMIC FLATPACK RESONATOR

Dale E. Beetley

General Electric

ABSTRACT

GE has developed a multichannel, high precision aging* measurement facility capable of high volume testing of resonators. Features of the facility considered unique for production aging systems tests include: 1) Loran-C/disciplined time-frequency (DTF) oscillator frequency standard, 2) direct current power bus design, 3) measurement and switching techniques, and 4) high volume automatic precision resonator aging. Computer-controlled data acquisition is used for unattended operation.

Facility requirements included frequency measurement with sufficient precision to allow 20-year extrapolation of resonator frequency shift using 30 data points. The frequency reference is traceable to the National Bureau of Standards (NBS).

Long-term extrapolation required selection of a model which would most accurately reflect the major processes involved in aging. In order to verify the accuracy of model extrapolation, a group of resonators has been maintained in test for more than nine years.

INTRODUCTION

The major subsystems of the GE Quartz Crystal Automatic Aging Measurement Facility include the frequency standard, measurement scanning console, computer system, and power distribution system. This paper reviews the design considerations applied to achieve a frequency drift measurement capability of parts in 10^{-11} per day in a production oriented test facility. Equipment capability and the long- and near-term aging affects and their correlation are discussed. Data model accuracy is considered, based on long-term data taken over a period approaching half the extrapolated life of the resonator.

BACKGROUND AND DISCUSSION

The initial impetus for this project was the need to obtain information relating to the long-term (20 years) frequency behavior of a ceramic, flatpack resonator [1]. The resonator, shown in Figure 1, consists of a ceramic "frame" on which a quartz crystal blank is mounted using molybdenum ribbon mounting "clips." Ceramic covers are installed using a thermocompression bond which provides a hermetically sealed package.

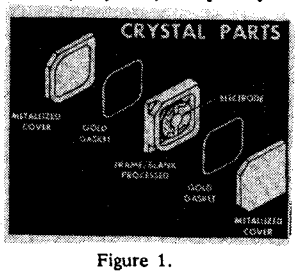


Figure 1.
Ceramic Flatpack Resonator

To minimize contamination problems and thereby promote good aging characteristics, high vacuum, high temperature processing is used throughout the resonator fabrication cycle. Since previous work has indicated that stress relaxation in the quartz blank is one process involved in aging [2,3], stress in the mounted blank was measured (see Figure 2).

*Aging refers to the frequency drift per day from an initial value.

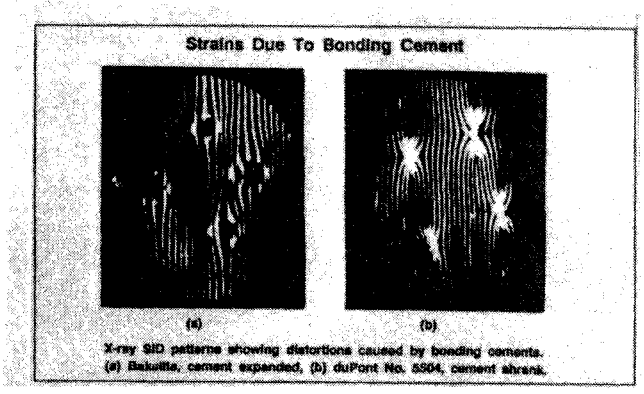


Figure 2. Stress in Mounted Quartz Crystal Resonator Blank

It can be seen that stress in the blank is heavily dependent on the mounting cement. However, other factors will also introduce stress in the blank; among these is the nonplanarity of the mounting clips. Precise clip forming and cutting tools have been developed to minimize this effect. Efforts to provide additional improvements in this area are ongoing.

Fabrication of the automatic aging measurement facility presented a number of technical challenges which required resolution for proper system operation [4]. Facility requirements included frequency measurement with sufficient precision to allow 20-year extrapolation of resonator frequency shift using 30 data points. It was also necessary that the facility operate unattended for extended periods of time. Long-term extrapolation then required selection of a model which would most accurately reflect the major processes involved in aging.

Facility Equipment And Operations

To provide the required stability, Loran-C transmission is used to supply the frequency measurement reference source. This also provides National Bureau of Standards (NBS) frequency traceability [5]. The Loran receiver, shown in Figure 3, produces a "raw" frequency reference phase-locked to a Loran-C signal. The signal is further conditioned using a disciplined time-frequency (DTF) standard oscillator [6]. The resulting phase-locked output of the DTF has excellent long-term and short-term frequency stabilities which are typically 5 x 10^{-12} per day [5] and an Allen variance (1 second) of 2 x 10^{-12} respectively [6].

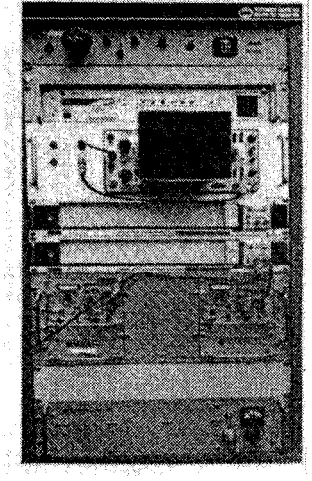


Figure 3.
Loran-C Transmission Receiver

The aging test console consists of 13 test drawers (Figure 4), each containing positions for 20 ovenized oscillators (Figure 5). Channel selection for 260 channels is provided by commercial scanner units (Figures 4 and 6).

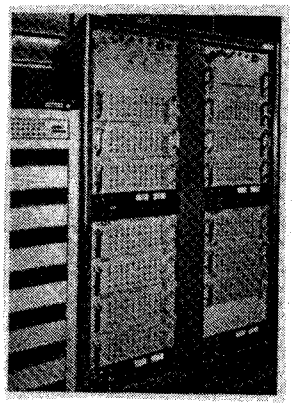


Figure 4.
Aging Measurement Console
With 13 Test Drawers

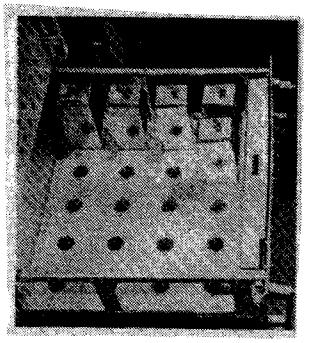


Figure 5.
Console Drawer Positions
for Ovenized Oscillators

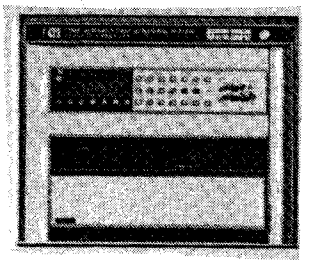


Figure 6.
Console Scanner Unit
for Channel Selection

Individual oven temperature is monitored using a thermistor. Thermistor output is measured at the socket using a four-terminal technique, essentially eliminating errors due to lead resistance (see Figure 7).

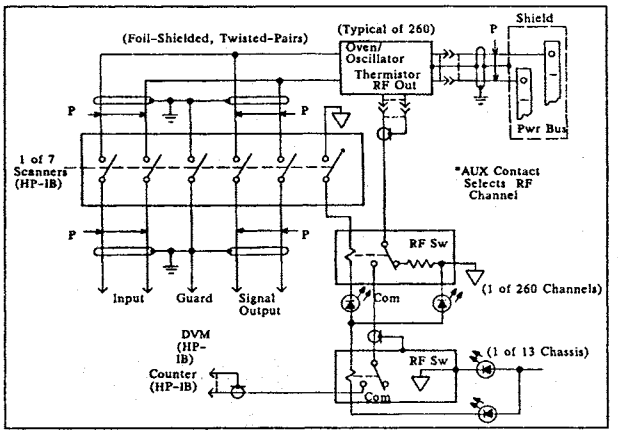


Figure 7.
Oscillator Aging Measurement Network

The ohmmeter range must be chosen carefully. An incorrect range can cause excessive current which would result in thermistor self heating, thereby causing large errors in the oven temperature signal. The large temperature offsets (several tenths of a degree) would mask the true reading, which is stable to ± 50 millidegrees or less. When measuring the thermistor (5000 ohms), a test current of 5 microamperes was found to be adequate to overcome errors due to ohmmeter output current.

Direct current (dc) power is obtained from high stability, high current supplies that power a maximum of 80 oscillators each. In order to minimize oscillator coupling, a special dc bus was fabricated for each supply. The bus consists of short, heavy, tin-plated copper bars. Voltage and sense leads for the associated power supply are connected between the two bars with the attachment point at the bar center point (see Figure 8). In order to minimize voltage variation resulting from oscillator loading, individual power leads were used for each oscillator.

Figure 8.
Direct Current Power Bus

The bus is installed in a shielded enclosure, with radio frequency (RF) interference decoupled, at the bus, between groups of oscillator power source leads associated with each individual chassis.

Interconnecting power leads, with unbroken shields, were constructed using foil-shielded, twisted-pair cable (see Figures 9 and 10). Power is also RF decoupled at each oscillator socket. In addition, the oscillator output signal is switched through RF matrix switches which provide a constant 50-ohm load to the oscillators and exhibit a minimum RF isolation of 80 decibels (dBm) between channels. The signal is then routed to a precision frequency counter with a resolution of 0.0001 Hz. The frequency reference for the counter is the signal generated by the Loran-C system.

Figure 9.
Interconnecting Power Leads
With Foil-Shielded, Twisted-Pair Cable

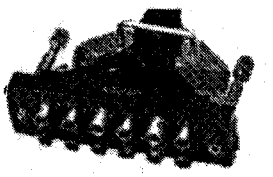


Figure 10.
Power Cable Interconnector

Through spectrum analysis of an unshielded bus with 22 operating oscillators, it was learned that the coupled noise and signal from other oscillators was down 40 dBm from reference. The same bus, when shielded, had a noise level 50 dBm down from reference, an additional 10 dBm noise margin (see Figures 11 and 12).

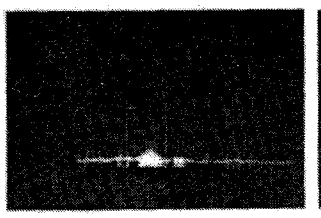


Figure 11.
Coupled Noise and Signal from
Unshielded Bus

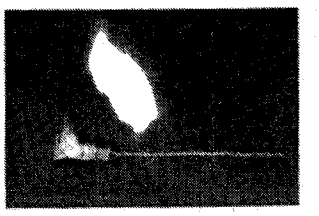


Figure 12.
Coupled Noise and Signal from
Shielded Bus

The channel selection provided by the scanner units is capable of the required four-terminal switching. Good RF grounding and shielding techniques are used throughout each chassis [7] (see Figures 10 and 13).

Signal leads other than the oscillator output signal were fabricated using foil-shielded, twisted-pair conductors in order to minimize coupled noise problems. Cables for the oscillator output signal were fabricated using double shielded coax. Extreme care in this area is essential. RF coupling through a poorly designed dc bus, such as one that does not incorporate "single-point feed, single-point ground," may cause fractional frequency pulling, a problem which will manifest itself as similar aging characteristics for all units. Correction of this problem could require the complete dismantling of a test station [8].

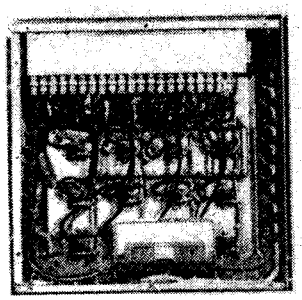


Figure 13.
RF Grounding and
Shielding Method

Figure 14.
Test Facility Computer System

The computer system for the aging test facility, shown in Figure 14, consists of a desktop calculator, two floppy disks, a real time clock, and a digital plotter. The system interface is the Institute of Electrical and Electronics Engineers (IEEE) 488 instrument control bus. Once started, all measurements are made daily, automatically, for a specified length of time, typically 30 to 45 days. Daily measurements of 260 oscillators usually can be completed in 45 minutes.

As a result of the need for uninterrupted operation in the event of power failure, an uninterruptable power supply (UPS) is used. The UPS is a charger/battery/inverter unit capable of supplying the required power for approximately 60 minutes, which is sufficient time for emergency generators to come on-line and take up the load. The UPS has prevented the interruption of several critical tests during weather-related power outages.

Aging Model

Resonator aging is currently modeled using the following equation, where $\Delta f/f$ is frequency change in parts per billion (ppb) and T is time in days with $T > 1$ [9]:

$$\Delta f/f = A + B \ln(T) \qquad (1)$$

A least squares fit to the aging data collected for a particular resonator is used to generate values for the A and B constants. This allows the model to predict resonator frequency or frequency rate of change over some arbitrary period of operation. Aging rate of change is determined by differentiation of equation 1 with respect to T:

$$d(\Delta f/f)/dt = d[A + B \ln(T)]/dt \qquad (2)$$

Therefore, the rate of change per day = B/T, where B is slope in parts per billion and T is days.

An objection raised with the current model is the inability to project frequency back to day "0." As a result, an alternate model was proposed as follows:

$$\Delta f/f = C + D \ln(ET + 1) \text{ [10]} \qquad (3)$$

where C, D, and E are constants and T, as before, is time in days. The proposed model assumes that stresses in plastically deformed electrodes are balanced by perfectly elastic stresses in the quartz, and that stress in the quartz is directly related to frequency [2]. While equation 3 is valid for $T = 0$, evaluation of the constants is significantly more difficult than for the current model, requiring either an iterative or a differential process. Also, for large values of T (approximately 15 days or greater), the constant (1) in equation 3 becomes negligible. In this case, equation 3 can be rewritten as:

$$\Delta f/f = C + D \ln(E) + D \ln(T) \qquad (4)$$

By using the equation

$$F = C + D \ln(E) = \text{constant} \qquad (5)$$

equation 4 becomes

$$\Delta f/f = F + D \ln(T) \qquad (6)$$

This has the same form as the model presently being used. It appears, then, that the primary advantage of equation 3 is that it permits an accurate model to be generated with fewer data points. However, the reduced number of data points does not reduce the labor involved, only the elapsed time. Any time savings will be offset by the additional time required for data reduction.

RESULTS AND CONCLUSIONS

The aging measurement test facility has been in use for over ten years. During this time, several thousand operations of the test bed oscillators have been performed, successfully, for their specified aging cycle. In order to verify the accuracy of model extrapolation, a group of 17 resonators has been maintained in test for more than nine years. The average aging rate was calculated and plotted periodically. Aging rates at the 500-day point were predicted for most of the units [9] and measured to be a few parts in 10^{-9}/day [11] (see Figure 15).

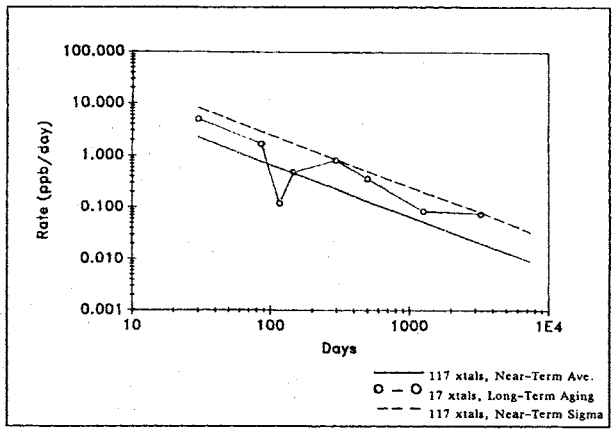


Figure 15. Resonator Frequency Change (20-year aging rate)

Extremely good source voltage stability must be maintained over the test period to prevent power change influence on the natural crystal aging, an effect demonstrated in Figure 15. A dc bus powering 17 oscillators, which had already undergone more than 60 days aging, was loaded with an additional 40 oscillators. The system measurement resolution was such that, following the increased bus loading, a decrease was noted in the average aging rate for the original 17 oscillators. The decrease, less than 1 ppb/day, later recovered when the bus was unloaded. The effect, however, was not sensed by the highly aged resonators when 200 additional oscillators were loaded onto other power buses in the test system. The magnitude of this effect is such that it is negligible during the normal aging cycle.

Several very stable 5-MHz resonators, with aging rates of parts in 10^{-11}/day, were measured on the system directly. A typical aging plot is shown in Figure 16. Note the relationship of the best fit of the "normalized" aging versus the actual data points. Further confirmation of model validity also can be seen with the relatively high value for the determination

coefficient, 0.993 in this example. In addition, note the trace of oven temperature regulation over the test period. This trace helps to resolve anomalies in frequency behavior of a resonator or in the resolution of test bed oven/oscillator malfunctions.

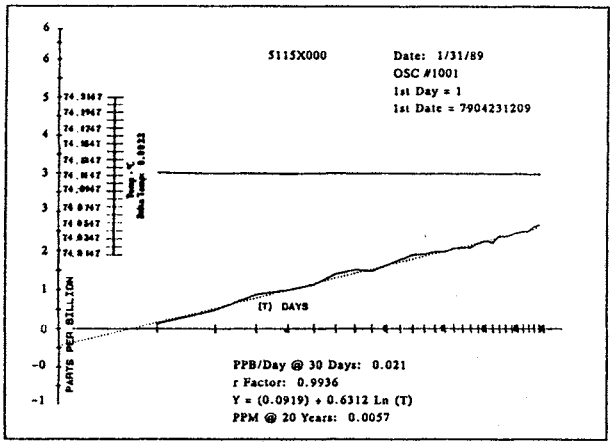


Figure 16. Resonator Aging, Typical Plot

To provide an average population aging prediction, a statistically significant control group of 117 resonators, which underwent aging during the same initial 30-day period as the long-term sample, was plotted. The average rate and 3 sigma limit at 30 days, along with the extrapolated rates for 20 years were plotted (see Figure 15). It can be seen that the average rate for the long-term test samples is approximately centered in the area bordered by the comparison sample data. The measured average rate (plotted values) for the test samples compares well with the projected values, with one exception. The measurement system is sufficiently sensitivity to detect a shift in rate of approximately 0.8 ppb. This occurred when the remaining channels on the power bus supplying the long-term samples were loaded. As can be seen, the aging rate recovered when the additional resonators were removed.

The close correlation between measured and projected values for the test samples appears to support the aging model used. The extrapolated rate for the test samples is approximately 0.05 ppb at 3288 days, while the measured value is 0.075 ppb. The data indicate that measurement capability for frequency change approaches the noise floor of the system. Previous Allen variance measurements of the oscillator/oven stability yielded a noise floor of parts in 10^{-10}. System measurement resolution is parts in 10^{-11}.

ACKNOWLEDGMENTS

Sincerest thanks are extended to T. M. Snowden and C. J. Biver for their helpful comments and to H. W. Hazzard and J. K. Pringle for their invaluable efforts in measuring and collecting the aging data.

REFERENCES

1. R. D. Peters, "Ceramic Flatpack Enclosures for Precision Quartz Crystal Units," Proceedings, AFCS 30, June 1976, p 22-231.
2. M. R. Miljkovic, G. Lj. Trefunovic, and V. J. Brajovic, "Aging Prediction of Quartz Crystal Units," Proceedings, AFCS 42, June 1988, pp 404-11.
3. J. C. Brice, "Crystals for Quartz Resonators," Reviews of Modern Physics, Vol. 57, No. 1, January 1985, p 105.
4. Dale E. Beetley, Baynard R. Blitch, and Thomas M. Snowden, "The Quartz Resonator Automatic Aging Measurement Facility," Proceedings, AFCS 35, May 1981, pp 263-70.
5. U.S. Coast Guard Headquarters, Loran-C System Performance Inquiry, April 1981.
6. D. Babitch, J. Ho, and M. Bloch, Frequency Electronics Inc., "The Disciplined Time/Frequency Standard: A New Multifunction Crystal Oscillator," Proceedings, AFCS 26, 1972.
7. R. Morrison, "Grounding and Shielding Techniques in Instrumentation," John Wiley and Sons, Inc., 1964.
8. R. Greene, Private Communication, Greenray Industries, August 1978.
9. D. R. Koehler, Test Requirements, Sandia National Laboratories, Albuquerque, NM, December 1978.
10. J. H. Armstrong, P. R. Blomster, and J. L. Hokanson, "Aging Characteristics of Quartz Crystal Resonators," Proceedings, AFCS 20, April 1966, pp 192-9.
11. R. L. Filler, L. J. Keres, T. M. Snowden, and J. R. Vig, "Ceramic Flatpack Enclosed AT and SC-Cut Resonators," Proceedings, IEEE Ultrasonics Symposium, 1980.

AUTOMATED MEASUREMENT SYSTEM FOR FINAL FREQUENCY PLATING AND CHARACTERIZATION OF HIGH PRECISION SC-CUT CRYSTALS

Pierre Stoermer

Ball Corporation, Efratom Division
Irvine, California U.S.A.

INTRODUCTION

Final frequency operation of a precision oscillator is dependent on the accuracy to which the frequency of the resonator can be set. This becomes extremely important when the crystal resonator has a small tuning range and when there are limited provisions for tuning reactance adjustments in the oscillator.

An automated measurement and processing system capable of characterizing, tuning, and sealing two hundred precision SC resonators during a single vacuum pump-down is described. Tuning of 10 MHz resonators to within 0.1 ppm at the operating temperature has been achieved. Correlation between resonator frequency operation in the tuning system and in the final oscillator is better than 0.5 ppm. Agreement between temperature slope measurements in the processing system and in an external temperature bath agree to within 10%.

Electrical and thermal performance characteristics of the resonators are stored in a data base, statistically analyzed and used for production process monitoring and modification.

PROCESS SYSTEM

The tuning/measurement process station is illustrated in Figure 1. The system consists of a stainless steel, cryogenically pumped, high vacuum chamber containing thermal evaporative deposition equipment, manipulation devices, measurement feedthroughs, and cold-weld sealing equipment. Within the vacuum chamber, a one meter diameter copper turntable holds two hundred resonator assemblies and their enclosures. The copper plate is heated by quartz lamps situated around the perimeter and is proportionally temperature controlled through a thermocouple embedded in the copper plate. Four rings of machined pockets are located at different diameters in the copper plate. Resonator blanks and headers are placed within pockets in the outer two rings. These pockets serve as the mask for gold deposition onto the blanks during frequency tuning. Two resistance heated gold sources are located beneath the copper plate, one for each of the two resonator diameter positions. Resonator covers are placed in the two inner rows. An electrical feedthrough/manipulator provides for connection between the resonator package pins and the outside measurement system. Figure 2 shows one of the resonators over one of the gold sources.

Electrical monitoring of the crystals during processing is done using an impedance analyzer and an instrumentation controller.

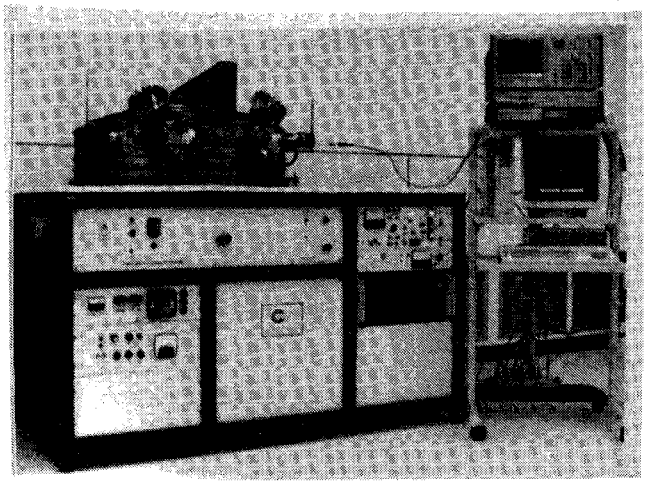


FIGURE 1: TUNING/MEASUREMENT PROCESS SYSTEM

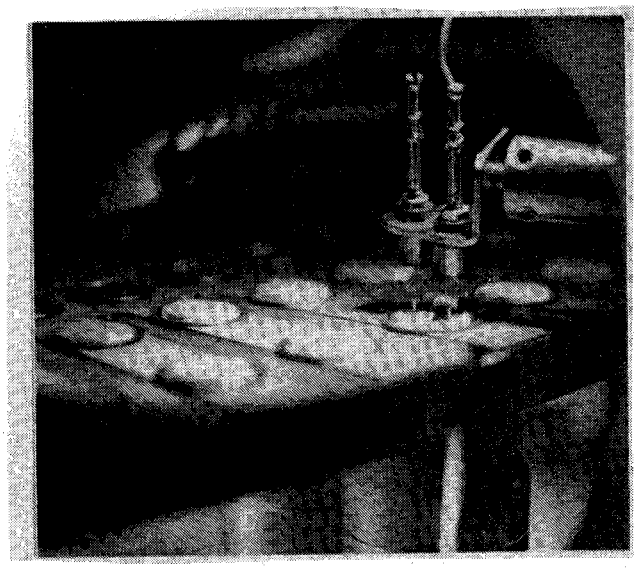


FIGURE 2: CONNECTION OF RESONATOR FOR TUNING

In production, resonator blanks, headers, and covers receive an extended vacuum baking at 275°C for approximately 30 hours. This also provides a stabilization period for the quartz blank. Cleanliness of the system is monitored using a residual gas analyzer. When the desired level of remaining contaminants and vacuum level is reached, the heaters are turned off and the plate is allowed to stabilize at the operating temperature. At this point, background pressure in the vacuum system is less than 1.3×10^{-6} Pascals.

Each resonator is rotated in turn to the tuning position. Electrical connection is made to the resonator during the gold deposition process. The instantaneous frequency is monitored by the impedance analyzer. Each resonator is tuned to approximately 50 Hz high of its final frequency during the first pass. On the second pass, the gold deposition rate is lowered and the resonators are tuned to their final frequency. The reduced deposition rate minimizes gold source heating effects of the crystal blank.

Once the entire batch of resonators have been tuned, the 40 kilogram copper plate is allowed to cool naturally by radiative heat loss to the surrounding vacuum system enclosure. The thermal time constant of this system is on the order of four hours. During cool down, several measurements of the C-mode and B-mode frequencies of each resonator are taken. These results are compared to previously measured data on the two modes at their operating temperature and the frequency-temperature slope calculated for each resonator. Resonators meeting both the electrical and thermal specifications are cold-weld sealed in the vacuum system prior to backfilling of the vacuum chamber.

ASUREMENT APPROACH

Electrical characterization of the crystal resonator has been accomplished using accepted one-port techniques. All measurements are taken with an impedance analyzer which is driven by a computer/controller. The equipment block diagram is shown in Figure 3.

Overall accuracy of the impedance analyzer relies on establishing a calibration plane for measurement of the resonator blanks. Calibration standards are maintained on the copper plate. These consist of an open, short, and 50Ω load. Calibration is done only at the normal operating temperature of the blanks. At this point the error in the calibration standards is less than 1% in resistance. The open and short are not affected by temperature. This allows compensation for parasitic reactances in the cables and measurement connectors so that only the electrical characteristics of the resonator are measured.

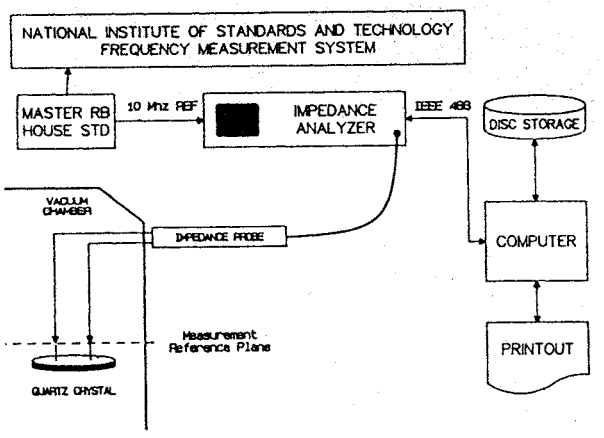


FIGURE 3: MEASUREMENT EQUIPMENT BLOCK DIAGRAM

A software routine has been written to determine the series resonant frequency, f_s, the motional reactances, C_m and L_m, and the resistance, R_s, of the crystal. The shunt capacitance, C_o, is measured at approximately 500 kHz above and below series resonance, averaged and used for subsequent calculation of the resonator's operating frequency during tuning. The tuned or operating frequency is calculated by:

$$f_o = (1 + C_m/2(C_o + C_l)) f_s$$

where C_l is the operating load capacitance in the oscillator. Since there is a significant dependency of the C-mode series resonant frequency on the resonator drive level, all measurements are taken with a resonator drive current of 0.6 ma which corresponds with:

1) minimal frequency deviation with drive current

2) the actual drive current in the oscillator.

An SC-cut resonator exhibits a mode commonly referred to as the "B-mode" which is approximately 9% higher in frequency than the desired "C-mode". Figure 4 illustrates the impedance response of the B and C-modes for a 10 MHz SC-cut resonator. Typically the B-mode has higher Q and lower resistance than the C-mode. Although this mode can present problems in the oscillator if not properly trapped, its existence allows a very convenient way of measuring relative temperature changes in the resonator. B-mode frequency measurement provides the most accurate data on the resonator's true operating temperature since it is a measure of the temperature at the center of the quartz blank and responds immediately to temperature variations. The B-mode exhibits a frequency-temperature slope of approximately -330 Hz/°C near the operating temperature of the resonator. Figure 5 illustrates the typical frequency-temperature characteristics of both the B and C modes over the desired operating temperature range.

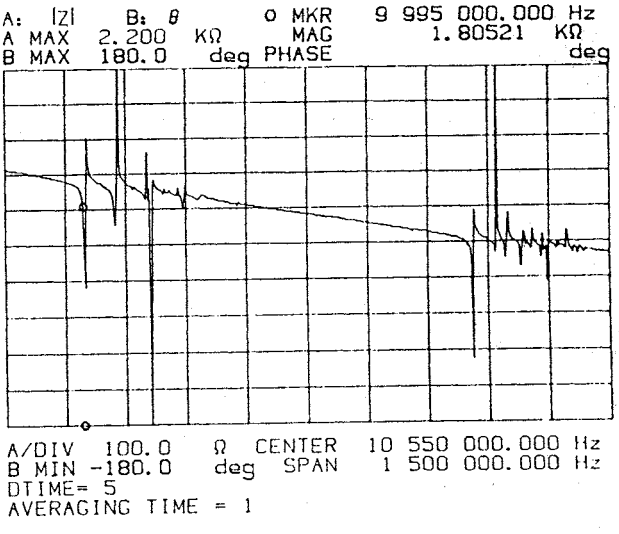


FIGURE 4: IMPEDANCE RESPONSE OF B AND C-MODES

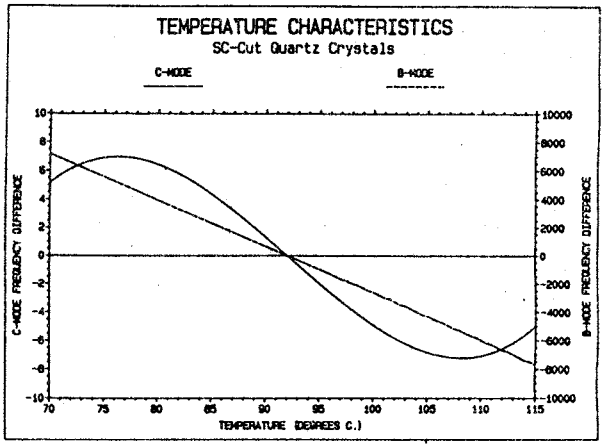


FIGURE 5: B AND C-MODE FREQUENCY-TEMPERATURE CHARACTERISTICS

Determination of the SC resonator's (C-mode) frequency-temperature behavior requires nearly simultaneous measurement of both modes at two different operating temperatures. This is accomplished by finding series resonance for both modes with the impedance analyzer in approximately five (5) seconds, then changing the resonator temperature and repeating the measurements. The difference in the B-mode frequencies divided by the B-mode slope yields the relative temperature change seen by the resonator. Frequency-temperature slope of the C-mode is determined by dividing the difference in the measured C-mode frequencies by the calculated temperature change derived from the B-mode data.

ASUREMENT SYSTEM

All measurements on the resonators are taken with an HP 4194A impedance analyzer with an HP model 310 controller. An in-house 10 MHz rubidium oscillator is used as a house frequency standard for the impedance analyzer. This standard is monitored continuously on equipment provided by the United States National Institute of Standards and Technology. It is monitored using a common-mode view of a Loran-C station. Average daily deviation from UTC is held within 5 parts in 10^{-12}.

Series resonance of both the B and C-modes is determined by measuring the complex impedance over a narrow range about series resonance. For frequency deviations sufficiently small, both the real and imaginary parts of the impedance can be represented by essentially straight lines. Given the requirement to make accurate measurements quickly, 19 data points are measured over a frequency sweep of less than 10 Hz for the C-mode. With sufficient dwell time at each measurement point, the total measurement cycle is completed within 2 seconds. A second order, least-mean-square fit is done to both the real and imaginary data.

Series resonant frequency is determined where the imaginary part of the equivalent series impedance is equal to 0. Series resistance is determined from the real part of the impedance at the series resonant frequency. The slope of the equivalent series impedance with respect to frequency is calculated at series resonance and used to calculate the resonator Q. From this and the shunt capacitance and motional resistance data measured earlier, the motional inductance and capacitance are calculated.

Electrical parameters and temperature data for each resonator during a particular process run is stored on the computer hard disk for subsequent statistical analysis as part of a continuing statistical process control program [1]. These results are used for correcting process drift in the areas of X-ray orientation, contouring, and base plating of the resonator blanks.

RESONATOR MEASUREMENTS

Figure 6 illustrates the typical measurement output from the system for the C-mode resonance on a 10 MHz resonator. This data is taken before tuning occurs to verify the quality of the resonator. The standard deviations on the equivalent series resistance, calculated motional parameters and Q are typically two to three orders of magnitude lower than the average measured values. During the tuning process the standard deviations of the measured values increase to one to two orders of magnitude lower than the average values due primarily to the rapid change in resonant frequency with time and heating effects during gold deposition. The standard deviations return to more respectable levels when the gold deposition rate is lowered during the final tune to operate frequency.

CRYSTAL IDENTIFICATION: A20

1 AVERAGE(S)

C MODE - BASE FREQUENCY = 10 MHz

26 Jan 1989
12:41:16

MEASURED SHUNT CAPACITANCE IS: 4.63E-12 FARADS

Eqv R	Delta Fo	Cm	Lm	Q
42.779	+1.860730E+03	+3.043607E-16	+8.319363E-01	+1.222130E+06
42.759	+1.860727E+03	+3.042016E-16	+8.323715E-01	+1.223358E+06
42.727	+1.860720E+03	+3.044169E-16	+8.317828E-01	+1.223404E+06
42.785	+1.860720E+03	+3.043653E-16	+8.319237E-01	+1.221947E+06
42.723	+1.860711E+03	+3.041366E-16	+8.325493E-01	+1.224629E+06
42.645	+1.860700E+03	+3.041770E-16	+8.324388E-01	+1.226714E+06
42.627	+1.860701E+03	+3.041226E-16	+8.325877E-01	+1.227456E+06
42.751	+1.860698E+03	+3.041228E-16	+8.325872E-01	+1.223905E+06
42.672	+1.860685E+03	+3.039754E-16	+8.329909E-01	+1.226745E+06
42.687	+1.860685E+03	+3.040303E-16	+8.328404E-01	+1.226096E+06

C MODE MEAN VALUES:

42.716	+1.860708E+03	+3.041909E-16	+8.324009E-01	+1.224639E+06

C MODE STANDARD DEVIATIONS:

.055	+1.646535E-02	+1.472446E-19	+4.028828E-04	+2.001350E+03

FIGURE 6: MEASUREMENT SYSTEM OUTPUT

After final frequency tuning the frequency-temperature slope of each resonator is measured. Figure 7 shows the deviation in slopes measured within the process system as compared to measurements made on the sealed resonators in a temperature controlled bath. Temperature accuracy and repeatability within the bath was better than 0.02°C. Measurements of the series resonant frequency of both the B and C-modes were manually made with a different set-up which utilized an impedance analyzer tied to the house reference. Thirty four measurements were made on twelve different resonators at temperatures ranging between 90 and 110°C. The standard deviation of the data is 7.5 parts in 10^9. The production requirement for temperature slope of the resonator is ± 4.5 parts in 10^8; we know with 95 % confidence that any resonator with a measured slope of ± 3 parts in 10^8 will meet the final oscillator requirements.

Correlation between the final oscillator operating frequency and the tuned resonator frequency is approximately 0.5 ppm even though resonators are tuned to 0.1 ppm. The major contribution to this discrepancy is believed to be with changes in the actual load capacitance presented by the oscillator circuit.

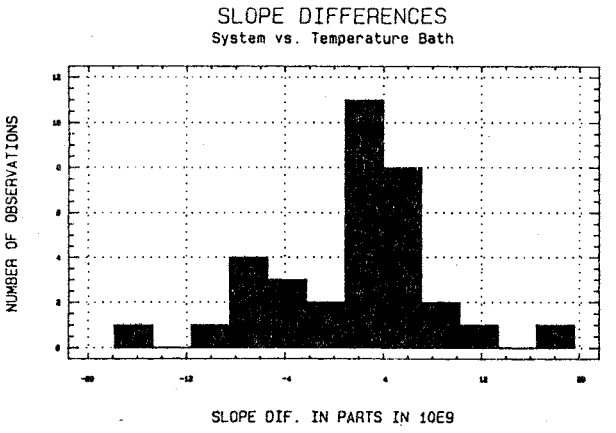


FIGURE 7: C-MODE SLOPE DIFFERENCES

OCESS CONTROL

Measurements of the motional parameters, series resonant frequency before tuning and temperature slope give a good picture of the health of various processes used for resonator production. Differences in the motional capacitance from batch to batch indicates variation in the contour of the blank.

Initial measurement of the series resonant frequency at operating temperature before tuning is a direct measure of the control achieved in the chemical polishing operation and the calibration of the base plating system.

Temperature slope data and its variation with time is used to correct the X-ray orientation system. Collimating crystals which provide the Bragg reflections to properly orient the SC-cut are periodically adjusted to keep the frequency-temperature slope characteristics within specification.

Once resonators are assembled into oscillators, additional feedback on actual load capacitance presented by the oscillator network can be derived. This information is used to correct the tuned frequency of the resonator in the system so that final operating frequency in the oscillator meets specification.

CONCLUSION

A process/measurement system has been described which provides for high volume fabrication and test of precision SC-cut quartz resonators. A virtue of the system is that all frequency adjustments and electrical characterization of the resonators are accomplished within an environment simulating final use. Resonators are hermetically sealed within the system after characterization and are ready for installation into an oscillator when removed. Resonator data is analyzed and used to maintain the health of the fabrication process.

REFERENCES

1. J.A. Kusters and C.A. Adams, "Applications of Total Process Control Techniques in the Production of High Precision Quartz Resonators," Proceedings of the 39th Annual Symposium on Frequency Control, May 1985, p. 475 ff.

DUAL-MODE TEMPERATURE CONTROLLED QUARTZ RESONATOR

Serge GALLIOU - Marc MOUREY

Laboratoire de Chronométrie, Electronique et Piézoélectricité (ENSMM)
Route de Gray - La Bouloie - 25030 BESANCON CEDEX - FRANCE

ABSTRACT

Miniaturization is nowadays a leading feature of oven controlled crystal oscillators. In this paper, the authors propose a solution which is well adapted to that demand. It is founded upon the use of a SC cut resonator vibrating in two simultaneous modes ; the 10 MHz frequency is the C mode's, the B mode's being used as a temperature sensor. The heaters are set inside the resonator enclosure.

A description of the designed resonator structure is followed by a recall of the dual-mode oscillator principle, which was developped in order to take out the two frequencies. Then, the temperature control loop using those frequencies is presented. Finally, a few measurements summary the designed system performances.

INTRODUCTION

Compact oven controlled crystal oscillators appeared on the market (1, 2, 3). They provide lower power consumption and faster warm up than the previous generation. we target those features when using the SC cut resonating in two simultaneous modes. The crystal is cut for the third-overtone C-mode resonance at 10 MHz ; it may resonate simultaneously on its third-overtone B-mode at 10.9 MHz, which is also a mode of thickness shear, orthogonal to the C-mode. The B-mode is strongly sensitive to temperature. Therefore, with a heat source closed to the crystal, the quartz temperature may be directly controlled by the B-mode frequency used as a temperature sensor.

The goal of our investigations consists on analysing the performances we may expect from such a system. It is obvious the ultimate resolution is that of the temperature sensor. Within the temperature range, the B-mode frequency changes according to the relation $\Delta f_B = K\Delta T$ with $-300 \leq K(Hz/°C) \leq -260$. In other words, a B-mode frequency control with a resolution of 1 Hz means a temperature resolution of $4\ 10^{-3}$ °C which must lead to a C-mode frequency stability better than 10^{-9} at the turnover temperature. The design goal for warm-up time to within an accuracy of $1\ 10^{-9}$ is less than 15 minutes, with a maximum heater power consumption of 500 mW;

For that first step, experiments are done with standard electronic components, while keeping in mind that the electronic might be integrated in a next step. In the same way, the resonator structure is designed from standard's matched for our experiments.

This is discribed in the first part of the paper. The second and third parts concern the electronic design :

- the dual-mode oscillator which supplies the C-mode frequency f_C and the mixing of both frequencies, $f_{BC} = f_B - f_C$, sensitive to temperature

- the controller loop which processes f_{BC}.

1 - RESONATOR STRUCTURE

The tested resonators are standard 10 MHz-SC cut crystals in standard enclosures. Heaters have been added inside those enclosures, and the crystal holders insulated with regard to the latter in order to reduce the power consumption.

The first resonator design we have tested is a BVA assembly (see fig. 1). The heaters are resistive paste printed on the "Condensors", on the opposite surfaces of the electrodes'. The two condensors and the resonator are sandwiched together by two insulating pieces in P.T.F.E.

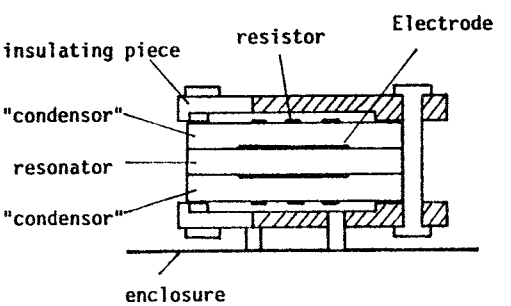


Figure 1 : Schematic cross-section-"BVA-type"

A second design has been tested (see fig. 2) ; the resonator with deposited electrodes is heated with I.R. radiation (4). The crystal is enclosed in two reflecting half-shells. To minimize thermal losses this assembly is insulated with regard to the enclosure by a piece of lead glass.

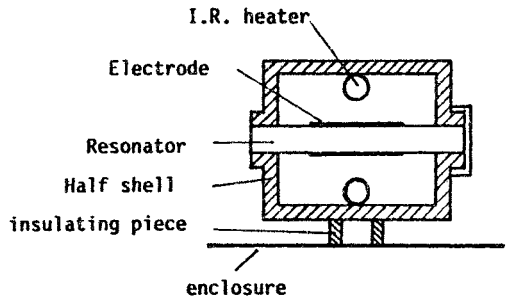


Figure 2 : "V-type" structure.

Before using the B-mode frequency as a temperature sensor, both designs have been evaluated with a thermistor sticked on the heaters (5, 6).

2 - DUAL-MODE OSCILLATOR

Fig. 3 shows a block diagram of the dual mode oscillator. Previously used with a 5 MHz-SC cut (7, 8), it has been translated at 10 MHz. Nevertheless, that transposition was not direct because of the B-mode dominance in 10 MHz-SC cuts : the C-mode frequency stability is closely related to the ratio of the 10 MHz-signal level to the 10.9 MHz-signal level (8). The bigger this ratio, the better the frequency stability is. An automatic gain control is then necessary ; it holds the crystal drive level constant.

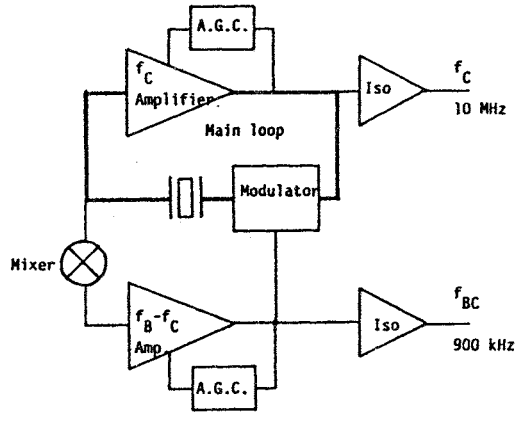


Figure 3 : The dual-mode oscillator (O.B.M.T.F.)

We can recall the operating principle of the dual-mode oscillator called O.B.M.T.F. ("Oscillateur Bi-Mode à Translation de Fréquence") : the crystal resonates on both modes C and B, provided the input modulating signal of the modulator inserted in the main loop is at $f_B - f_C$, where f_B and f_C are the B-mode frequency and the C-mode frequency respectively. One way to implement that is to self-mix the output signal of the resonator. Here the signal processed in the second loop is at a low frequency $f_B - f_C$, far from the frequency f_C of the main loop ; then filtering and level controls are made easier.

3 - TEMPERATURE CONTROL LOOP

The oscillator provides a f_{BC} signal. This is the frequency processes in the temperature control loop, instead of the B-mode frequency f_B. Indeed, the C-mode frequency sensitivity versus temperature is negligeable compared with that of the B-mode frequency (see fig. 4). Therefore, the relation $df_{BC}/dT = df_B/dT$ is almost true for static temperature changes.

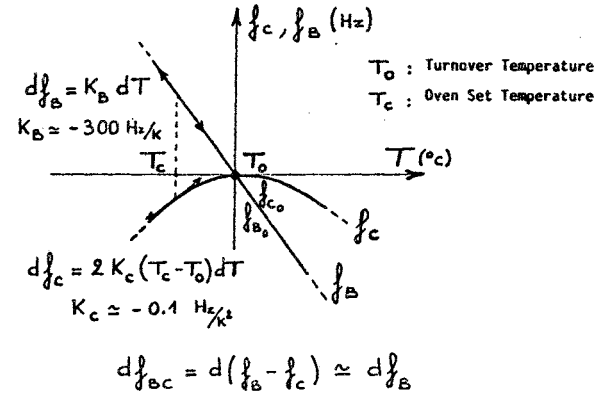


Figure 4 : B and C mode frequency sensitivity near the turnover point.

The temperature control consists on accurately monitoring the frequency f_{BC} by applying power to the heaters. The frequency f_{BC} has to be counted and compared to its reference value at the turnover temperature.

Many designs may be applied : for that first step, we have chosen a simple design, mixing analog and digital electronic, which provides a resolution of 1 Hz for the frequency f_{BC}, by counting it in a measuring time of 1 second. The latter is the result of f_C-divisions (see fig. 5) ; then, its stability is $\Delta f_C/f_C$.

Let N the counting result of f_{BC} in a n-bit binary counter ; then $f_{BC} = x \cdot 2^n + N$ ($0 \leq N \leq 2^n-1$), if x is the number of counter rotations. The operating temperature range determines the frequency range and the counting result must be such as $0 \leq N \leq 2^n-1$ within the latter. From - 20°C to 80°C i.e. for ΔT = 100°C, the frequency change is about $\Delta f_{BC} \simeq 30000$ Hz which leads to $2^n \geq 30000$ or $n \geq 15$. Components impose an even number, so, here n = 16.

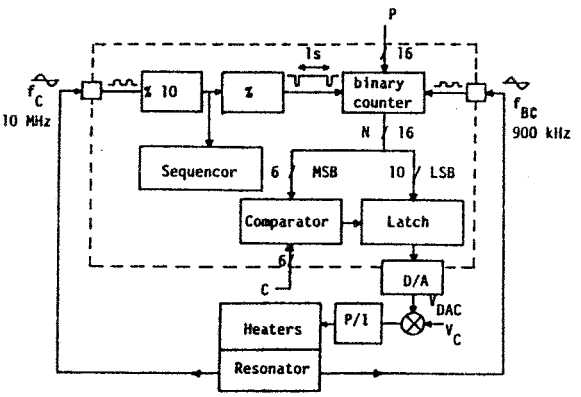


Figure 5 : Block diagram of the temperature-control electronic.

In addition, counter is preset at P before beginning to count, so that the number x of counter rotations is the same for all the frequencies within the range, whatever the resonator may be. Like that, only one number N is associated to one frequency value.

Only the ten least significant bits (L.S.B.) of N are converted into voltage to be compared with a reference voltage. They cover a frequency change of 2^{10} Hz, i.e. approximately 3°C. The six most significant bits (M.S.B.) are capable of indicating the temperature by step of 3°C (refer to fig. 5 and fig. 6). When comparing them with a proper value C, the temperature control loop may be locked in its working area.

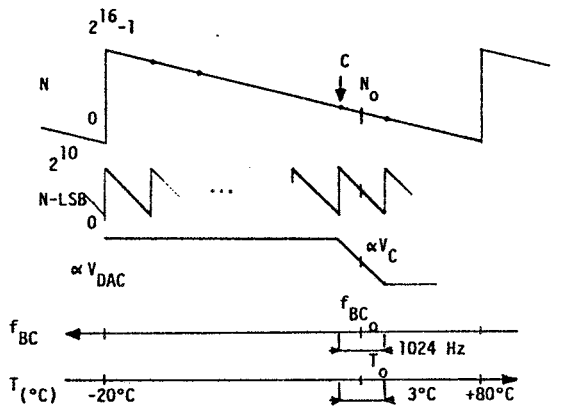


Figure 6 : Tuning principle of the temperature control loop.

At the end of each 1 second-measuring time the ten LSB of N are transfered in latches and held up to the end of the following measuring time. This transfer occurs while the MSB-comparator output allows it. It is not the case during the warm-up time where the state of the MSB-comparator output means the system is far from the turnover point : then the latch outputs force the heating.

All the digital electronic is monitored by a sequencor which uses 4 µs between two measuring times to preset the components.

The analog part of the electronic allows to introduce proportional and integral controls. Dead times may also be compensated in this section. A satisfactory analysis of the feedback control system can be accomplished by the use of the figure 7.

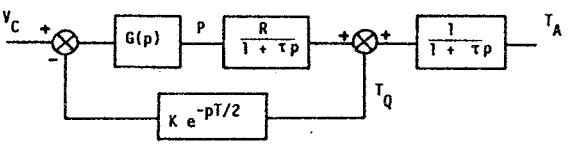


Figure 7 : Temperature control system.

The symbols represent the following :

T_Q : crystal temperature
T_A : Ambient temperature
P : Heater power
V_C : Reference voltage
R_C : thermal resistance between the crystal and the ambient temperature
τ : Resonator assembly time constant
T : Frequency measuring time
G(p) : Transfer function of P/I actions and lag-lead systems.

4 - MEASUREMENTS - ANALYSIS

The first trials had been done with the BVA-type assembly. We achieved satisfying performances with regard to the warm-up : the system required about 12 minutes to stabilize within 1×10^{-9} of its final frequency, as shown in fig. 8, with a maximum heater power consumption of 445 mW. However, the tests on the temperature control loop efficiency revealed a lack of thermal gain.

Figure 9 shows the change in the temperature sensor (the frequency f_{BC}) which results of a change in the enclosure temperature. It follows that the thermal gain, about 250 is unacceptable. When increasing the loop gain, it is necessary to compensate the time constant (about 150 seconds ; it is called τ in fig. 7) between the heaters and the sensor, to avoid oscillations. Our attempts with simple compensating networks aborted. Investigations concerning the BVA-type assembly were stopped here while waiting to find a solution to that problem of the time constant reduction. Measurements of its short term frequency stability and heater power consumption are shown in fig. 10 and 11.

The other design, we call V-type, does not have the disadvantage of the former : the radiative heating decreases the above-mentioned time constant. The temperature control loop may be adjusted such as the

frequency f_{BC} is locked in the system resolution, that is 1 Hz (or 4 10^{-3} °C), even if the enclosure temperature is a step greater than 10°C. Fig. 12 shows the warm-up behavior : the oven controlled oscillator needs 12 mn to reach the operating frequency within 1 x 10^{-9} ; during this warm-up, the heater power is about 420 mW. The steady-state heater power consumption in moving air is shown fig. 11. At least, fig. 10 shows the short-term frequency stability. All these measurements had been done with the electronic at the room temperature ; only the resonator enclosure was submitted to temperature changes.

These results may be considered as satisfying. They prove the temperature control loop efficiency. However, we does not find the expected stability of the frequency f_C when the enclosure temperature changes, as shown in the example fig. 13. We attribute this behavior to thermal gradients due to a non-uniformed heating of the bulk and, more important, to stresses induced by the holders. It must be pointed out that the controller loop locks the frequency f_{BC} in a constant value. Therefore, frequency shifts depending on mechanical stresses for example will also cause heating power effects.

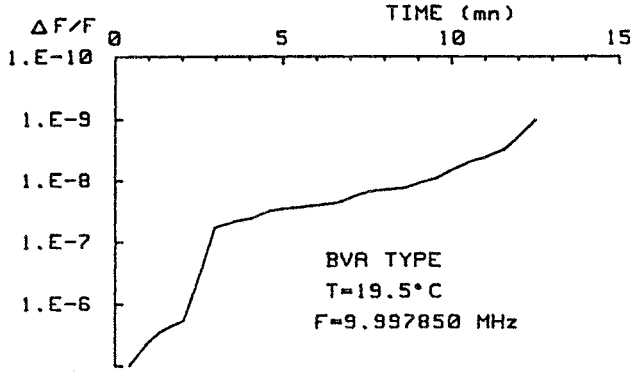


Figure 8 : Warm-up - BVA-type resonator.

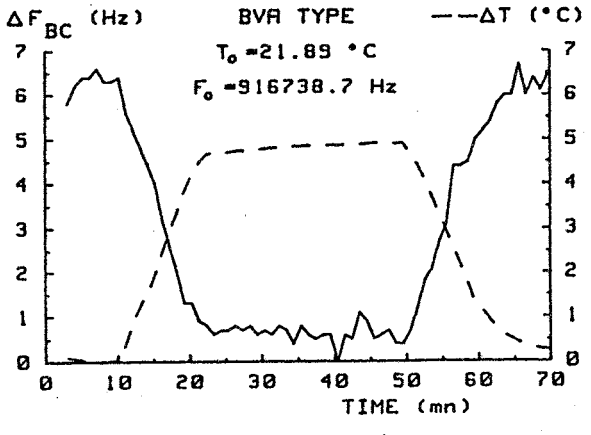


Figure 9 : Effect of an ambient temperature change on the temperature control loop (with a BVA-type resonator).

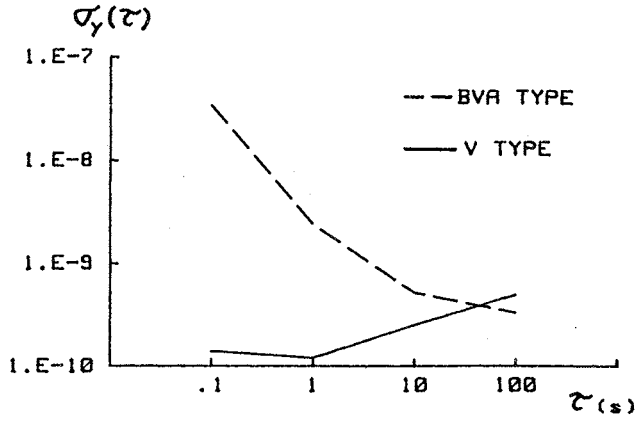


Figure 10 : Short term frequency stability.

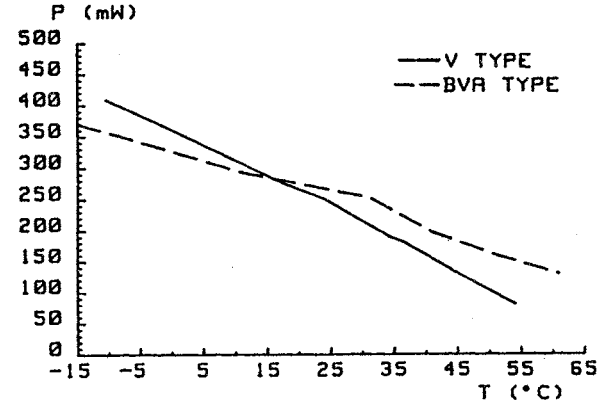


Figure 11 : Heater power consumption (see Fig. 7 : thermal resistance :
R = 230°C/W for the BVA-type
R = 195°C/W for the V-type).

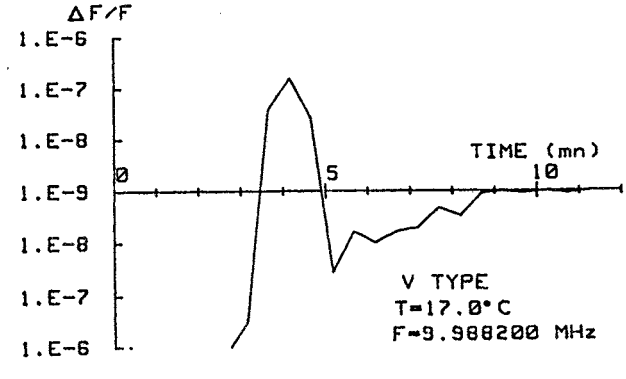


Figure 12 : Warm-up - V-type resonator.

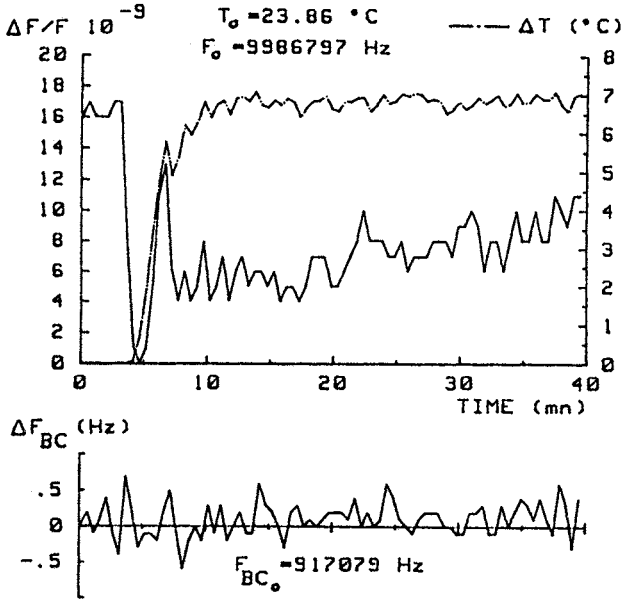


Figure 13 : **Effect of an ambient temperature step on the frequency f_C (10 MHz). (The temperature control loop is working).**

CONCLUSION

This work is one application of the dual-mode principle. We have to improve the electronic design (the temperature control loop and the dual-mode oscillator) before its miniaturization. This first step of investigations has particularly shown the importance of the mechanical aspect. That is the point we are working on now.

This work is supported by D.R.E.T. (D.G.A.).

REFERENCES

(1) H.W. JACKSON, "Update on the tactical miniature crystal oscillator program", 36th Annual Frequency Control Symposium, pp 492-498, 1982.

(2) G. MARIANNEAU, D. HAUDEN and J.J. GAGNEPAIN, "Integrated oven controlled quartz crystal oscillator", 40th Annual Frequency Control Symposium, pp 350-354, 1986.

(3) K. SCHWALLER, K. SCHULTHESS and W. ZINGG, "Miniature Oscillators", 2d European Frequency and Time Forum, pp 817-822, 1988.

(4) J.P. VALENTIN, "Résonateur à thermostat infrarouge intégré", brevet d'invention n° 83-07307 du 03/05/1983.

(5) S. GALLIOU, "Chauffage rapide avec résonateurs BVA "tactiques"", p 31-43, rapport de synthèse contrat DRET n° 85/180.

(6) C. NOEL, "Conception et réalisation d'un thermostat pour résonateur à quartz", D.E.A. Acousto-Opto-Electronique et Vibrations n° 1132; université de Franche-Comté, octobre 1988.

(7) M. MOUREY, J.L. VATERKOWSKI, "New design of a dual mode quartz crystal oscillator", Electronics Letters, Vol. 21, n° 5, p 185, 1985.

(8) S. GALLIOU, M. MOUREY, "Compensation statique et dynamique en température d'un oscillateur à quartz bi-mode", 1st European Frequency and Time Forum, p 267-272, 1987.

THE STW RESONATOR-A NEW CONTROL ELEMENT FOR HIGH STABILITY OSCILLATORS

B.A. Auld

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

Surface wave resonators for oscillators operating in the UHF and low microwave frequency ranges have, in the past, always operated in the Rayleigh surface wave mode (SAW). Another type of surface wave, the surface transverse wave (STW), has been recognized since 1980 as offering a number of potential advantages for resonator application. But it was not until 1987 that this promise was realised with the publication by T.L. Bagwell and R.C. Bray of experimental results for a set of industrially-produced STW resonators. As predicted, these resonators were found to exhibit significantly higher power-handling capabilities than SAW resonators with comparable parameters, an important feature with regard to long term aging. In addition, no evidence was found for the existence of higher order transverse modes in the resonator gratings - a significant design constraint in SAW resonators. These devices also had values for Q and phase noise level that are comparable to those of the best SAW resonators to date.

By contrast with Rayleigh surface waves, STW are polarized in the horizontal plane and exist only on corrugated surfaces. Consequently, the character of an STW depends not only on the properties of the substrate material, but also on the dimensions of the corrugations. This feature introduces additional design parameters which can be manipulated to vary, for example, the penetration depth of the surface wave and the variation of propagation velocity with temperature. Optimization of an STW resonator design therefore involves many more possibilities for trade-off than is possible in the case of SAW resonators.

This paper reviews the basic concept and character of STW and the elements of resonator design. Trade-offs used in the design of temperature-compensated structures will be explained and illustrated. A survey of experiments will be given; and outstanding problems, as well as new directions for research will be indicated.

Introduction

Surface elastic wave resonators are now well-established as control elements for high stability oscillators operating in the UHF and low microwave frequency range, where a bulk wave fundamental thickness mode resonator is too thin to be fabricated as a self-supporting structure. The original conception of a surface wave resonator and its practical realizations were based on the use of Surface Acoustic Waves (Rayleigh Waves). Most surface wave resonators now in service are of this type. The top line of Fig. 1 shows such a resonator in section view. In this device the smooth center region between the two periodic arrays of etched grooves is the standing wave resonance region. The two periodic grooved gratings are designed to provide "mirrors" containing a Fabry-Perot surface wave resonance. This is achieved by spacing the grooves at one-half an acoustic wavelength at the operating frequency of the resonator. At this spacing there is a constructive combination of the small reflections from the individual grooves (Bragg scattering), producing total reflection and a stopband in the propagation characteristics. In the resonator region proper there may be a number of half-wavelengths in the spacing of the mirrors, leading to the possibility of multiple resonances. But just one of these is selected by designing the mirrors to have a narrow reflection stopband synchronous with the selected resonance, thereby achieving single-mode operation. Electrical coupling to the resonator (in either one-port or two-port operation) is realized by means of one or two interdigital surface wave transducers placed either in the resonator region or just inside the mirror regions. Design of these SAW resonators requires a phenomenological or analytical model for the propagation and reflection characteristics of a finite length grooved grating. Reference 1 reviews the general properties of SAW resonators and details a commonly used approach to the modeling problem.

In the late 1970's another type of planar elastic wave device appeared on the scene - the surface skimming bulk wave (SSBW) delay line[2]. Propagation in these delay lines takes place in a horizontally polarized bulk shear wave traveling almost parallel to the surface of the substrate. The SSBW, or SBAW (Shallow bulk acoustic wave) delay lines use interdigital transducers for excitation and detection. For certain orientations of a rotated Y-cut quartz crystal, these horizontally polarized shear waves propagate at an unusually high velocity ($\approx$ 5000 m/s). This type of device is therefore very suitable for operation at microwave frequencies, because the wavelength (and therefore the spacing of the interdigital fingers) is larger than for SAW at the same frequency. Consequently, fabrication of microwave frequency devices is greatly simplified. Temperature compensated rotated cuts, analogous to those used for quartz bulk thickness shear resonators, also exist for SSBW propagation. Figure 2 compares bulk wave resonance and SSBW propagation in quartz for two rotated cuts. Because of the change in propagation direction SSBW propagation in AT-cut quartz is analogous to thickness resonance in BT-cut quartz, and a similar relationship exists between BT SSBW propagation and AT thickness resonance. It is the AT cut that has the high SSBW propagation velocity mentioned above (Fig. 3), but the BT cut has a flatter temperature compensation curve.

Because of the high propagation velocity of SSBW on AT-cut quartz, this type of propagation appears to be an attractive candidate for use in surface wave resonators at microwave frequencies[3]. The large wavelength of the bulk shear waves means that the mirror gratings have larger period and can therefore be fabricated at higher frequencies. Unfortunately, SSBW propagation is not a surface wave normal mode (as, for example, SAW), but is a narrow spectrum of radiated plane waves excited by the interdigital transducer. Consequently, an SSBW resonating between mirrors has significant radiation losses, and therefore a relatively low Q. This difficulty can be overcome by etching a grooved grating on the surface of the substrate. As will be seen in the following analysis, the presence of the grating creates a horizontally polarized surface wave mode (a surface transverse mode,STW). As in the case of the SAW grating at the top of Fig. 1, this grating structure creates a stopband near the first Bragg scattering frequency, where the grating period is equal to one-half an acoustic wavelength. At frequencies below the lower edge of the stopband propagation is possible, but at a lower phase velocity than that of a SSBW. This passband behavior is illustrated by the curve labelled STW resonator in Fig. 4. Note that all points on this curve have a phase velocity (the ratio of ω to β) that is lower than the SSBW, or bulk shear wave, velocity. The surface wave nature of the

solution in this frequency range can be explained physically by analogy with optical refraction. When a light ray is incident on a medium with a higher velocity, a critical angle of incidence is reached when the refracted ray propagates parallel to the interface. For larger angles of incidence, and slower effective velocities along the interface, the light amplitude decays exponentially as a function of penetration into the second medium. Analysis shows that the surface wave nature of the STW solution extends into the part of the stopband below the SSBW line in Fig. 4.

The SAW resonator structure of Fig. 1 is modified for STW operation by etching a propagation grating in the central region between the mirrors. This resonator grating may have the same period Λ_m as the mirrors, but with different depth (Fig. 1). Reducing the groove depth decreases the individual reflections and narrows the stopband. This permits, at a given frequency, operation in the stopband of the mirrors and in the pass band of the resonator, An alternate structure, shown in the bottom line of Fig. 1, has a resonator grating with the same depth as the mirror grating, but with a smaller period. As shown in Fig. 4 this also permits operation in the stopband of the mirrors and the passband of the resonators. Since there are no depth variations in this case, fabrication requires only a single masking step. SAW resonators are also constructed using metal strips rather than grooved gratings. The interdigital transducers can then be incorporated in the mirrors in a single fabrication step, as shown in Fig. 5. To adapt for STW operation a smaller period grating is added in the resonator region, and the resonator again operates as shown in Fig. 4.

Current-State-of-the-art

Reference 4 reviews the development of transversely polarized grating waves, outlines the analysis of STW and demonstrates experimentally the existence of the STW stopband. In a recent paper[5] Bagwell and Bray at Hewlett-Packard reported the first industrial STW resonator with performance comparable to a high quality SAW resonator. Transducer analysis for this device was carried out by Flory and Baer[6].

Figure 6 illustrates the general characteristics of one Hewlett-Packard resonator operating at 500 MHz, with an unloaded Q of 11,000 and an insertion loss of 7 dB. In the experimental frequency response curve of Fig. 6, it is to be noted that the transverse modes observed in SAW resonators do not appear. The physical reason for this is not yet understood. Figures 7 and 8 show the phase noise and aging characteristics of another STW resonator at 632 MHz. By comparison with SAW, STW have a substantially greater penetration depth. This feature led in the past to predictions that such resonators would be less sensitve to surface effects and could be operated at higher power levels. The results shown for phase noise, burn-in, and aging appear to be consistent with these predictions. But more research will be required to understand the phenomena in detail. Finally, Fig. 9 illustrates measured temperature-frequency characerisitics of a 500 MHz STW.

Temperature Compensation

In a SAW resonator the temperature-frequency behavior is entirely controlled by the temperature dependence of the elastic constants, the density of the substrate, and the thermal expansion of the resonator length. The existence of a resonator grating in a STW resonator adds another parameter, the temperature variation of the grating dimensions. Figure 10 illustrates the role played by this factor in determining the turnover point of thetemperature-frequency behavior for a grooved STW resonator[7]. The starting point is the temperature variation of the SSBW velocity $V_s(T)$, which is determined by the properties of the substrate. Other parameters enter into the temperature variation of the stopband width $\delta\omega$, including the height and period of the grating. This latter variation shifts the STW dispersion curve relative to the SSBW dispersion curve, so that a particular point on the STW curve remains fixed, even though $V_s(T)$ varies with temperature. The result is that the turnover point on the temperature curve is shifted. Figure 11 is a theoretical prediction of the amount of this shift for a resonator on 37°-cut quartz[8].

For metal strip resonators the temperature behavior is more complicated because of the differential thermal expansion of the strips and the substrate. There is, as yet, no adequate model for this effect.

STW Propagation Characteristics

Figures 12 and 13 show in section view the mirror and resonator regions of grooved and metal strip STW resonators. The method of analysis follows the procedure described in Reference 1 for SAW resonators, except that in this case an analytical field model is used for propagation under a grating. The field model is first developed for a grating of infinite length, giving the dispersion relation for both the passband and the stopband. From this result the reflection coefficient of a semi-infinite grating can be obtained by the Kogelnik procedure, used in Reference 1 and reviewed briefly in the next section. The resonant frequencies are then obtained by a standard Fabry-Perot calculation.

Analysis of STW propagation is performed using the Floquet procedure. For an infinite grating the particle velocity field in the substrate is expressed as an infinite series of space harmonics.

$$v_x = \Sigma a_n e^{-\alpha_n y} e^{-i\beta_n z} v_0 e^{i\omega t} \tag{1}$$

$$\beta_n = \beta + \frac{2\pi n}{\Lambda}$$

where the coordinates are defined as in the central region of Fig. 12. Each space harmonic propagates along the grating, with an exponential decay into the substrate. These harmonics, with amplitudes a_n, are assumed to satisfy the acoustic wave equation for x-polarized propagation on rotated y-cut quartz.

$$c_{66}\frac{\partial^2 v_x}{\partial y^2} + 2c_{56}\frac{\partial^2 v_x}{\partial y \partial z} + c_{55}\frac{\partial^2 v_x}{\partial z^2} = \rho\frac{\partial^2 v_x}{\partial t^2} \tag{2}$$

This allows each harmonic decay factor α_n to be expressed in terms of the corresponding propagation factor β_n

$$\alpha_n = -\frac{ic_{56}\beta_n}{c_{56}} \pm \left\{\frac{c_{eff}}{c_{66}}\beta_n^2 - \frac{\rho\omega^2}{c_{66}}\right\}^{\frac{1}{2}} \tag{3}$$

with

$$c_{eff} = \frac{c_{55}c_{66} - (c_{56})^2}{c_{66}}$$

The acoustic fields in the substrate, given by Eq. (1), and the corresponding stresses must then be matched to the acoustic fields in the grating teeth. For shallow gratings a perturbation procedure is used[9]. This leads to two equations containing summations over the space harmonic amplitudes a_n, one equation for the region under a tooth and another for the adjacent groove[8]. By multiplying these equations by the complex conjugate of the z-dependent part of the wave function of an arbitrary space harmonic, one obtains an infinite set of linear algebraic equations for the space harmonic amplitudes. These can be solved numerically by truncating the set and solving for increasing numbers of harmonics, until the desired accuracy is achieved.

For the shallow gratings typical of surface wave resonators it is usually sufficient to apply the coupled mode approximation, where only the n = 0, -1 harmonics are retained[1]. In this case the set of linear equations reduces to two, which can be solved analytically for the dispersion

relation between ω and β and for the relative amplitudes of the 0 and -1 space harmonics. The resulting dispersion relation is most conveniently expressed in terms of the relative frequency and propagation factor variables defined in Fig. 14. For a grooved grating the dispersion relation is given by the following formula, with M equal to 1.

V_s = SSBW velocity in the substrate
V_m = shear velocity in the metal (isotropic)
ρ = density in the substrate
ρ' = density in the metal

$$\left\{\frac{\delta\omega}{V_s}\right\}^2 - \left\{\delta\beta\right\}^2 = \left\{MK\right\}^2 \tag{4}$$

$$M = \left\{\frac{1}{2}\frac{\rho'}{\rho}\left|1 + (\frac{V_m}{V_s})^2 + \frac{\pi}{2}(1 - (\frac{V_m}{V_s})^2\right|\right\}^2$$

$$K = \frac{2\pi}{\Lambda}\frac{\rho V_s^2}{C_{66}}(\frac{h}{\Lambda})^2$$

For the metal strip grating M has a value determined by the relative properties of the metal and the substrate. An alternative approach to the analysis of metal strip gratings is given in Reference 9.

It will be noticed that the right-hand side of Eq. (4), the so-called mode coupling constant, is proportional to the square of the grating depth h. By contrast, for a SAW grating it is linearly proportional to h. A general comparison of the two kinds of gratings is given in Reference 4. As will be seen, one consequence of this difference is that the mirror grating stopbands are narrower for STW applications that for SAW, and this makes STW resonator design more delicate.

Reflection Coefficient of a Finite Grating

In Reference 1 the analysis proceeds from a pair of differential equations for the 0 and -1 space harmonics. An alternative approach is to start from positive- and negative-traveling solutions to the truncated Floquet equations. By combining these solutions and suitably matching combinations of the forward and backward space harmonics to the incident, reflected and transmitted waves at the terminals of the grating, one obtains an analogous expression for the reflection coefficient. A comparison between STW and SAW gratings is given in Reference 4. For a semi-infinite STW grating the final result for the input reflection coefficient of the mirror is

$$\Gamma_m = \frac{\overline{K}}{S + jD} = |\Gamma_m| e^{j\Phi_m}$$

$$\overline{K} = (MK_m)^{1/2} \tag{5}$$

$$|\Gamma_m| = 1$$

$$D \approx Re(-\delta\omega/V_s + \delta\beta)^{1/2}$$

$$S \approx Im(-\delta\omega/V_s + \delta\beta)^{1/2}$$

in terms of the variables defined in Fig. 14.

Resonant Frequency Calculation

Figure 15 is a schematic of the standing wave resonance in the central region of an STW resonator, where the reflection coefficients are defined by Eq. (5). The analysis assumes a standing wave in the resonator region,

$$u = (A e^{-j\beta z} + B e^{+j\beta z}) e^{j\omega t} \tag{6}$$

$$\beta(\delta\omega) = \pi/\Lambda_r + \sqrt{(\delta\omega/V_s)^2 - (MK_r)^2}$$

with boundary conditions

$$\frac{B}{A} e^{j\beta_r L_r/2} = |\Gamma_m| e^{j\Phi_m}, Z = +L_r/2 \tag{7}$$

$$\frac{A}{B} e^{j2\beta_r L_r/2} = |\Gamma_m| e^{j\Phi_m}, Z = -L_r/2$$

at the mirrors. Solution of Eqs. (6) and (7) gives the frequency equation

$$\beta(\delta\omega) L_r = \Phi_m(\delta\omega) + \pi N \tag{8}$$

$$(N = 1,2,3,\ldots.)$$

where N defines the order of the resonance (approximately the number of half-wavelengths in the standing wave pattern).

For design calulations it is more convenient to cast Eq. (8) into the form

$$\beta - \frac{\Phi_m}{L_r} = \frac{\pi N}{L_r} \tag{9}$$

It then becomes easier to graphically visualize the position of the resonant frequencies. This is an important feature for STW resonator design because, as was noted above, the narrowness of the STW stopbands requires a delicate treatment of the design. This point is illustrated by Fig. 16, where the lower half of the mirror stopband (where a surface wave exists) has a relative width less than one-third of a percent. The design problem is to choose a resonator length L_r that is an integral multiple plus half of the grating period, and at the same time satisfy Eq. (9) inside the lower half of the mirror stopband.

Figure 17 illustrates a graphic procedure, on an enlarged view of the mirror stopband region in Fig. 16. The solid curve marked STW defines β as a function of δω, the frequency displacement from the crossover of the straight dashed lines. At the right of the STW line is a curved dashed line, with the offset marked. Here, Φ_m is the phase angle of the reflection as a function of δω. To find the resonant frequencies, the right side of Eq. (9) is marked off along the horizontal axis and the frequencies are read from the vertical axis. For an arbitrary choice of parameters solutions are not always found within the required frequency range (that is, the lower half of the mirror stopband). Carrying out a design therefore requires an empirical adjustment of parameters, using a computer.

Summary

A brief review has been given of the properties of STW surface wave resonators, giving an overview of basic concepts and the method of analysis. These resonators were predicted to have certain technological advantages, such as ease of fabrication, improved aging properties, and reduced phase noise. Results reported in the recent literature have confirmed these expectations, but the physics of this improved performance is not yet fully understood. A program of STW resonator fabrication is now under way at CNRS LPMO, with the goal of researching in depth the noise and aging characteristics of this new type of resonator.

References

1. P.S. Cross and R.V. Schmidt, *Coupled Surface-Acoustic-Wave Resonators*, BSTJ 56, 1447-1482 (1977).

2. M. Lewis, *Surface Skimming Bulk Waves, SSBW*, Proc 1977 IEEE Ultrasonics Symposium, 744-752.

3. K.H. Yen, K.F. Lau and R.S. Kagiwada, *Recent Advances in Shallow Bulk Acoustic Wave Devices*, Proc 1979 IEEE Ultrasonics Symposium, 776-784.

4. A. Renard, J. Henaff and B.A. Auld, *SH Surface Wave Propagation on Corrugated Surfaces of Rotated Y-Cut Quartz and Berlinite Crystals*, Proc 1981 IEEE Ultrasonics Symposium, 123-128.

5. T.L. Bagwell and R.C. Bray, *Novel Surface Transverse Wave Resonators With Low Loss and High Q*, Proc 1987 IEEE Ultrasonics Symposium, 319-324.

6. C.A. Flory and R. L. Baer, *Surface Transverse Wave Mode Analyses and Coupling to Interdigtal Transducers*, Proc 1987 IEEE Ultrasonics Symposium, 313-318.

7. B.A. Auld and D. F. Thompson, *Temperature Compensation of Surface Transverse Waves for Stable Oscillator Applications*, Proc 1987 IEEE Ultrasonics Symposium, 305-312.

8. D. F. Thompson, *Temperature Compensation of Microwave Acoustic Resonators*, PhD Thesis, Department of Electrical Engineering, Stanford University, June 1986.

9. S. Datta and B.J. Hunsinger, *First-order Reflection Coefficient of Surface Acoustic Waves from Thin-Strip Overlays*, J.Appl. Phys. 50, 5561-5655 (1979).

10. E. Danicki, *Propagation of Transverse Surface Acoustic Waves in Rotated Y-Cut Quartz Substrates Under Heavy Metal Periodic Electrodes*, IEEE Trans. SU-30, 304-312 (1983).

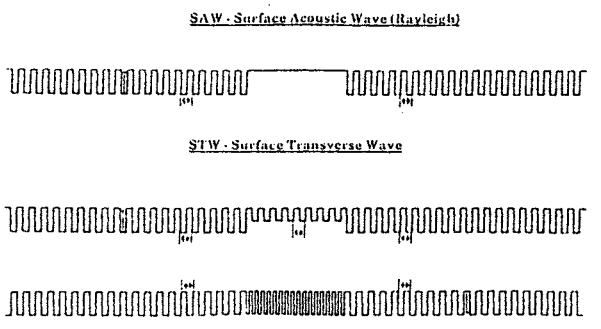


Figure 1 Grooved SAW and STW Resonators

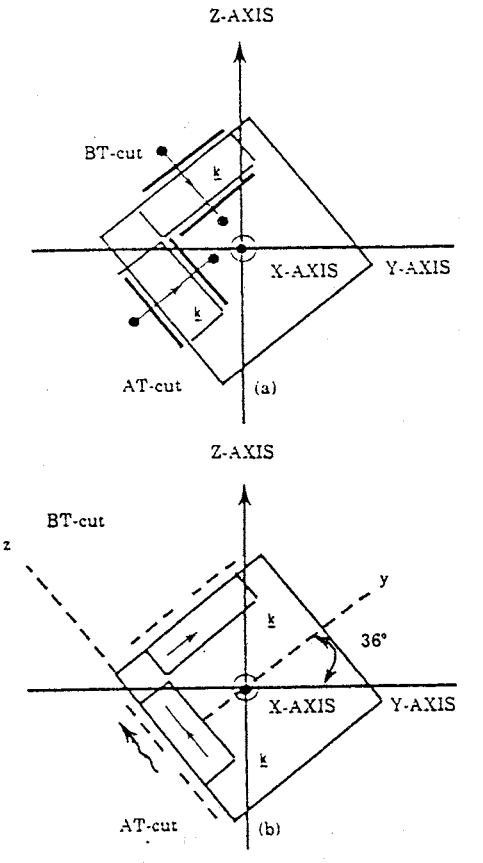


Figure 2 (a) Bulk Propagation on AT and BT Quartz (b) SSBW on AT and BT Quartz (After Reference 2)

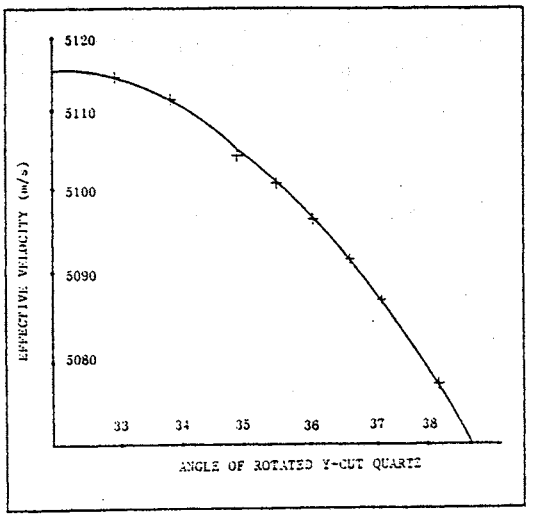


Figure 3 SSBW Velocity versus Crystal Cut (After Reference 2)

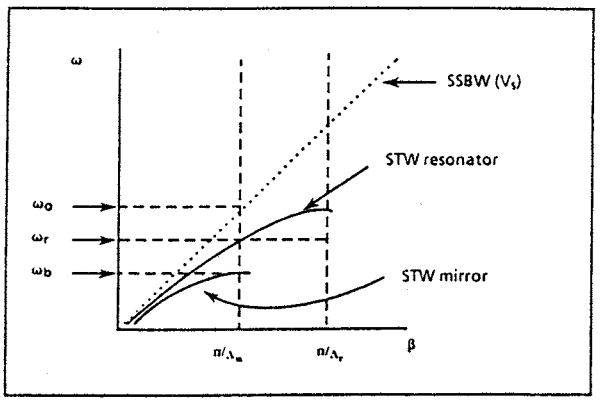


Figure 4 Typical Dispersion Curves for the Resonator and Mirror Gratings of an STW Resonator

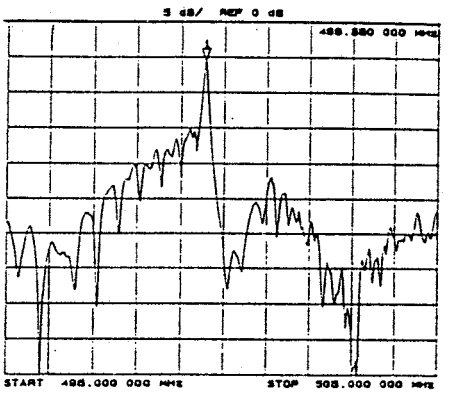


Figure 6 Resonance Frequecny Response of a 500 MHz STW Resonator (After Reference 5)

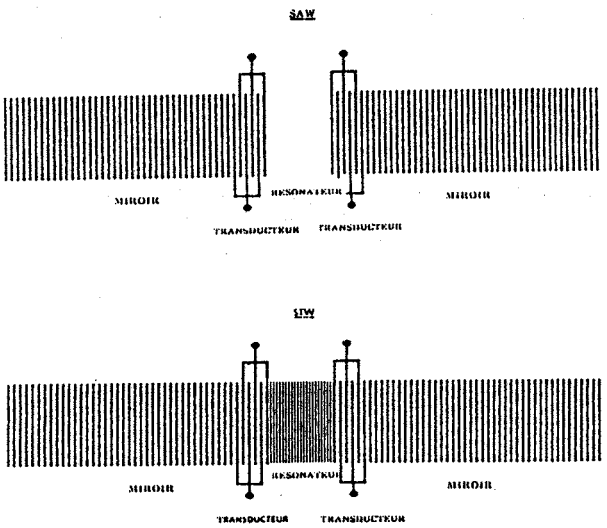


Figure 5 Metal Strip SAW and STW Resonators

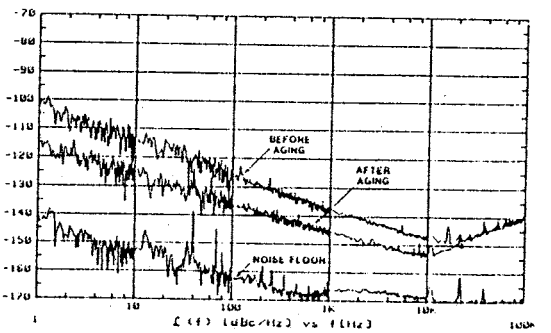


Figure 7 Reduction of Phase Noise by Burn-in (After Reference 5)

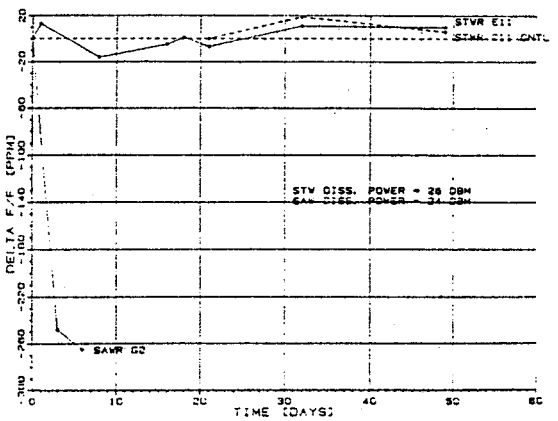


Figure 8 Comparison of SAW and STW Aging Characteristics (After Reference 5)

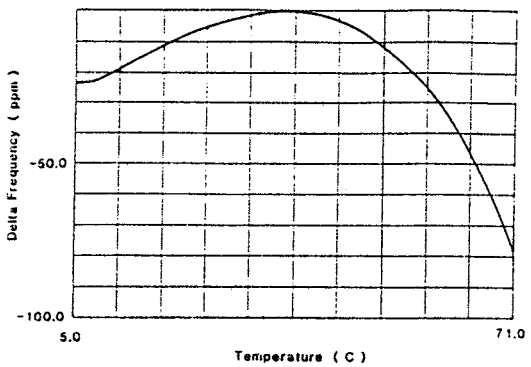


Figure 9 Typical STW Temperature/Frequency Characteristic (After Reference 5)

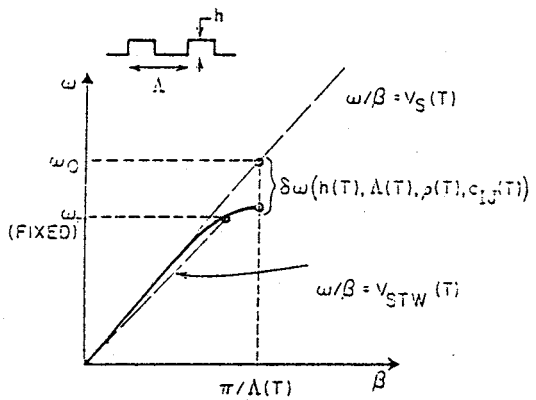


Figure 10 Temperature Compensation Mechanisms in a Grooved STW Resonator

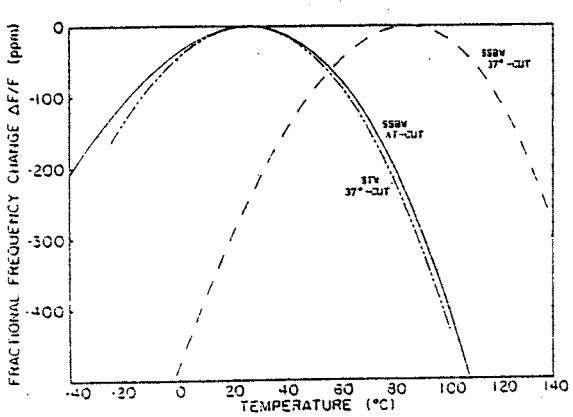


Figure 11 Illustration of the Shift in Turnover Temperature Due to an STW Grooved Grating

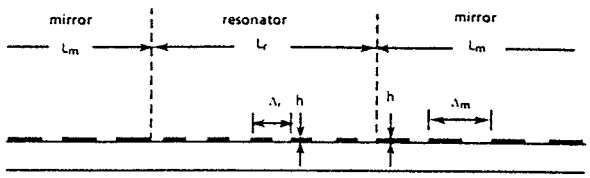


Figure 12 Section View of a Grooved STW Resonator

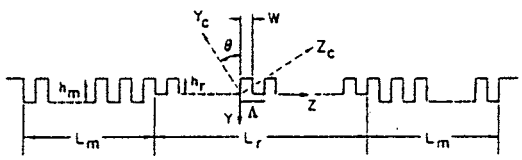


Figure 13 Section View of a Metal Strip STW Resonator

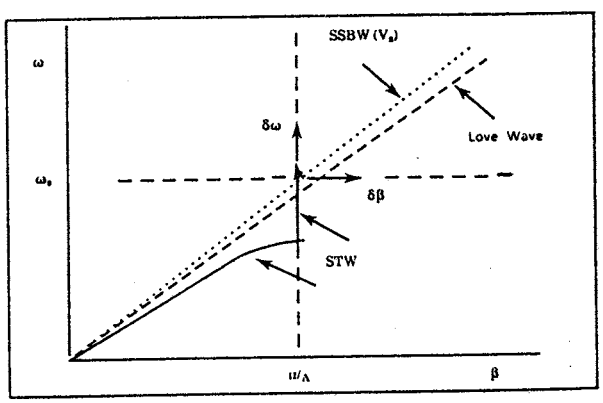


Figure 14 Definition of the Relative Frequency and Propagation Factor Variables δω and δβ

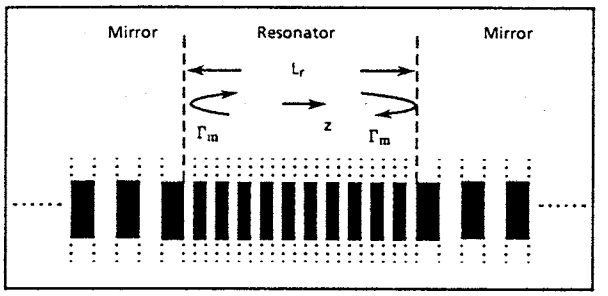


Figure 15 Fabry-Perot Resonances in an STW Resonator

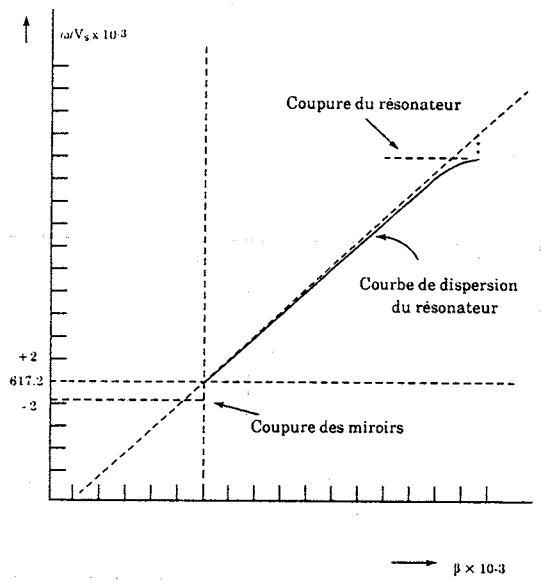


Figure 16 Typical Disposition of the Resonator Grating and Mirror Grating Stopbands

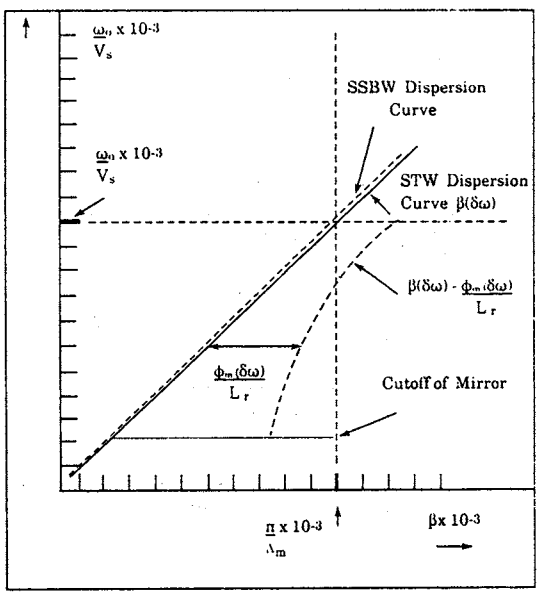


Figure 17 Graphical Construction for Finding STW Resonant Frequencies

EXPERIMENTAL STUDY OF PLANAR STRESSES EFFECTS ON SURFACE ACOUSTIC WAVE DEVICES

E. Bigler, S. Ballandras, D.Hauden

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

The influence of mechanical stresses on SAW devices may be described by stress-sensitivity coefficients which are the tensorial equivalents of the first order static temperature coefficients. Crystal anisotropy is used to find quartz crystal cuts exhibiting zero-stress coefficient for planar isotropic stresses. Among singly rotated quartz cuts ($\phi = 0$), two SAW quartz cuts exist satisfy both conditions of temperature -and planar isotropic stress compensation. These cuts have been experimentally tested with SAW delay lines built on circular quartz plates. Comparison of both temperature and stress effects is made with respect to the classical (ST,X) cut measured in the same experimental conditions. Potential applications involve new devices with low sensitivity to thermoelastic effects.

Introduction

High stability SAW oscillator applications require a careful design of the device to minimize the frequency shifts due to the influence of external physical parameters such as temperature, forces, accelerations and vibrations. Static temperature effects may be avoided using crystal anisotropy by the choice of a temperature-compensated cut like the classical (ST,X) cut of quartz. Dynamic temperature effects occur in warm-up conditions or when external temperature cycling is made with a characteristic time shorter than the thermal time constant of the device. In that case, the substrate will be submitted to thermal gradients, inducing thermoelastic stresses. Additional stresses may be transmitted into the plate from external forces applied to the mount holders, or due to acceleration effects.

A first method to minimize stress effects is to use a classical temperature-compensated cut (like the ST,X) cut and to design suitable plate shape and fixation conditions, yielding reduced stress sensitivity [1,2]. Another method which is presented here consists in using crystal anisotropy to find new cuts exhibiting simultaneously zero-temperature sensitivity and zero-stress sensitivity to certain kinds of stresses [3,4]. The sensitivity of a SAW device built on a thin plate of quartz crystal may be defined [5] according to the relation

$$\frac{\Delta V}{V_o} = \sum_{ij=1}^{3} {}^{s}a_{ij}\,\overline{T_{ij}} \qquad (1)$$

In eq. (1), $\Delta V/V_0$ represents the relative velocity shifts of the SAW induced by a quasi-static mechanical bias T_{ij}.

The case of planar isotropic stress compensation will be considered here. This configuration appears for example in the case of a circular plate submitted to a symmetrical distribution of mechanical or thermoelastic stresses. Within these assumptions, it has been shown [3,4] that several SAW quartz cuts are both temperature -and stress-compensated, defining for SAWs the equivalents of the SC-cuts of bulk waves [6].

This work is supported by Rome Air Development Center, Hanscom Field, Air Force Base, Bedford, MA 01731 (USA) under AFOSR grant # 87-0241

Determination of temperature -and stress- compensated quartz cuts for planar isotropic stresses

Stress and temperature coefficients may be calculated using the same formalism developed by Tiersten [7] for wave propagation in a prestrained medium. In the case of an homogeneous distribution of stresses and strains at constant temperature, the relative velocity shifts of the SAW may be expressed using the perturbation method by [7,8,9]

$$\frac{\Delta V}{V_o} = {}^{s}a_{pq}\,\overline{T_{pq}} \qquad (2)$$

where T_{pq} are the components of the quasi-static stress tensor and sapq the stress coefficients defined as

$${}^{s}a_{pq} = \left\{\delta_{ik}\delta_{sp}\delta_{rq} + C_{iskruv}\,s_{uvpq} + C_{\ell skr}\,s_{i\ell pq} + C_{is\ell r}\,s_{k\ell pq}\right\} U_{iskr} \qquad (3)$$

where δ_{ik} is the Kronecker delta,
C_{iskruv} are the 3rd order elastic constants,
s_{uvpq} are the 2nd order compliance constants,
$C_{\ell skr}$ are the 2nd order stiffness constants
and U_{iskr} depends only on the propagation terms of the SAW on the unperturbed crystal.

In the case of a thin plate of quartz crystal submitted to in-plane, isotropic stresses, eq. (2) takes the simpler form

$$\frac{\Delta V}{V_o} = ({}^{s}a_{11} + {}^{s}a_{33})\,T \qquad (4)$$

where the indices (1,2,3) refer to a set of axes (a_1,a_2,a_3) taken with a_1 parallel to the SAW, a_2 normal to the plate and a_3 parallel to the surface. T is the only non zero stress component: $T = T_{11} = T_{33}$; $(T_{13} = 0)$.

The stress coefficient ${}^{s}a_{11} + {}^{s}a_{33}$ for planar isotropic stresses has been computed for quartz, in the non piezoelectric approximation, as a function of the cut angles (Φ,θ) and propagation direction Ψ corresponding to a $(YXw\ell t)\Phi\theta\Psi$ plate, [10] SAW propagation in the ℓ-direction.

Taking the angle Ψ as a parameter, the loci of stress-compensated cuts are defined in the (θ,Φ) plane by the condition

$$({}^{s}a_{11} + {}^{s}a_{33})(\Phi, 0, \Psi) = 0 \qquad (5)$$

These loci are represented in solid lines on Fig. 1 for propagation directions $\Psi = 30°$ (Fig. 1a), $\Psi = 45°$ (Fig. 1b). On the same graphs are represented in dotted lines the loci of zero first-order temperature coefficients calculated using Tiersten's formalism of temperature derivatives of fundamental elastic constants [11,12] in the non piezoelectric approximation.

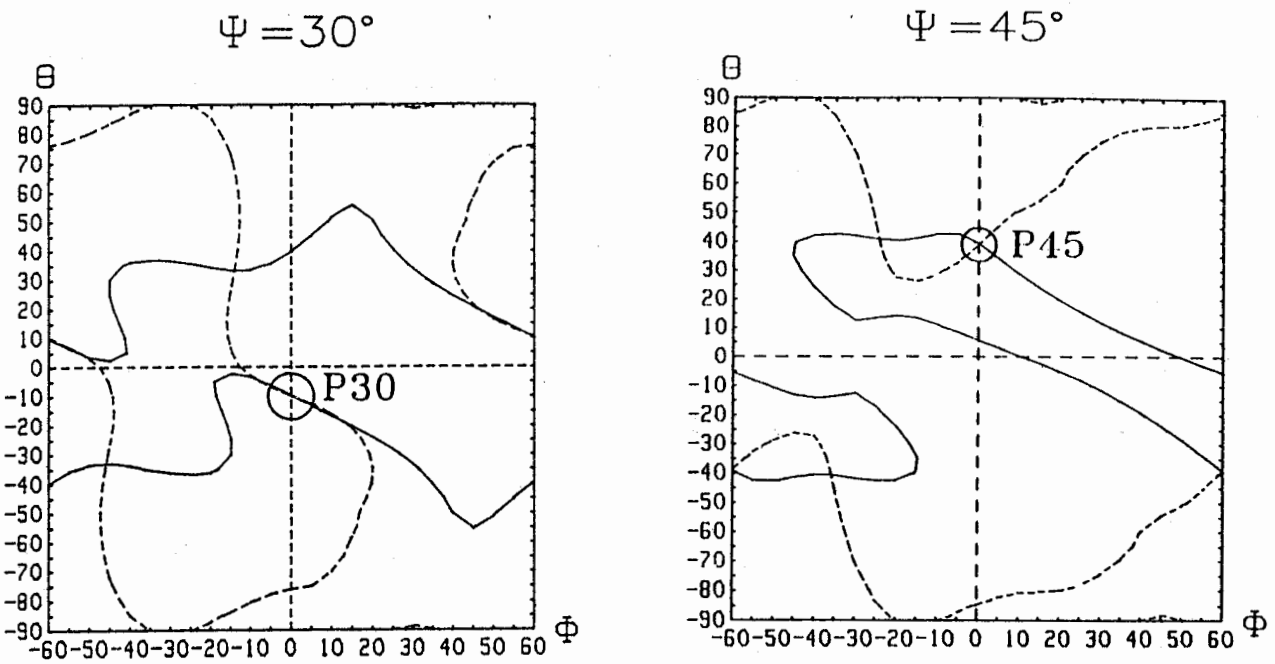


Fig. 1 :
- Solid lines : locus of quartz cuts compensated for planar isotropic stresses effects
- Dotted lines : locus of zero first order temperature coefficient
Cut angles θ, Φ and propagation direction Ψ correspond to a (YYwℓt)Φθ plate, SAW propagating in the ℓ-direction

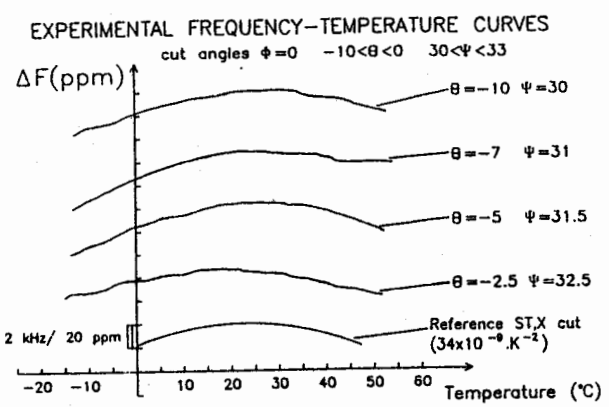


Fig. 2 : Experimental static temperature characteristics for the cuts $\phi = 0, -10° < \theta < -2.50, 30° < \psi < 33°$
Comparaison with the classical ST,X cut

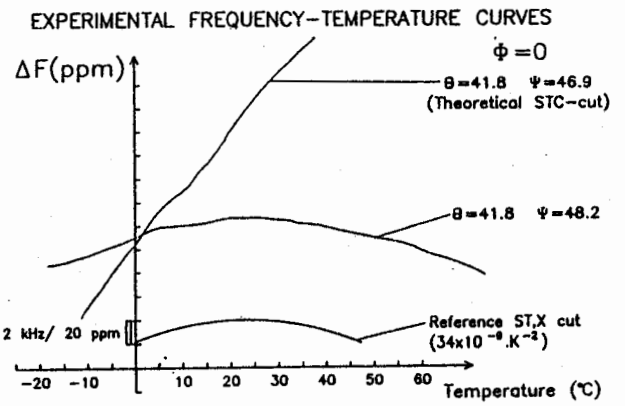


Fig. 3 : Experimental static temperature characteristics
a - for the cut ($\phi = 0, \theta = 41.8°, \psi = 46.9°$) corresponding to the theoretical values of B.K. Sinha STC cut [13]
b - for the corrected cut ($\theta = 0, \theta = 41.8°, \psi = 48.2°$) close to the STC-cut

The intersections between the two loci define quartz cuts satisfying both conditions of planar isotropic stress and temperature compensation.

The cuts defined on Fig. 1 have been tested experimentally; they are the only solutions among singly rotated plates ($\Phi = 0$) for simultaneous temperature and planar stress compensation. Singly rotated plates have been preferred in this experimental study since the precision of angular machining is better than for doubly rotated plates.

It must be pointed out that the second cut (point P45 on Fig. 1b) is close to B.K. Sinha's theoretical STC cut [13] ($\Phi = 0, \theta = 41.8°, \Psi = 46.9°$).

Experimental results

Experiments are performed on quartz delay lines built on circular plates (diameter : 22 mm, thickness : 1 mm or 2 mm).

The device is used in an oscillator operating at 100 MHz.

Static temperature effects

Temperature compensation was checked by plotting the frequency vs temperature curve between -20°C and +50°C. The propagation angle Ψ was adjusted to keep the inversion point of the frequency/temperature curve between 20-25°C (Fig. 2 and 3).

A discrepancy of about 1.3° on the angle Ψ was observed between B.K. Sinha's theoretical STC cut (Fig. 3a) and the experimental temperature-compensated cut (Fig. 3b).

As a reference, the theoretical curve for the classical (ST,X) cut is shown. It may be seen that the thermal behavior of the new cuts is very similar to the classical (ST,X) cut.

In-plane stress effects

A delay line satisfying the condition of temperature compensation is tested with a mechanical system represented on Fig. 4.

The circular plate is submitted to a single or double diametral compression obtained by two pairs of knife-edges acting at 90° to each other.

The plate may be rotated around its axis, and the force-frequency effect is measured for each azimuthal angle β between the propagation direction of the SAW and the direction of the force (Fig. 5).

The experimental procedure is as follows :

The knife-edge are removed, and the plate is held by a clamp until a new angular position is set. The plate is then submitted to a diametral pre-compression of about 1N. The clamp is removed, so that the plate is submitted only to the compressive forces. An additional force between 1N and 12N is set, and the frequency shifts are measured with respect to the pre-strained state. The azimuthal angle is measured without contact by optical means.

Theoretical model of the force-frequency effect in an anisotropic disk

The force-frequency effect in a circular disk may be derived from the stress sensitivity coefficients (5) and from the classical theory of elasticity [14].In the case of a single diametral compression, the relative frequency shifts in the isotropic model are given by

$$\frac{\Delta F}{F_o} = -\frac{2P}{\pi d t}\left[({}^s a_{11} + {}^s a_{33}) + 2({}^s a_{11} + {}^s a_{33})\cos 2\beta - 2\,{}^s a_{13}\sin 2\beta\right] \quad (6)$$

where P is the applied force along a diameter
d is the diameter and t the thickness of the disk
${}^s a_{11}$, ${}^s a_{33}$, ${}^s a_{13}$ are the in-plane coefficients expressed in the axes (a_1, a_3) parallel to the SAW
β is the azimuthal angle between the SAW and the force (Fig. 5).

The simple geometry of the circular disk allows an exact evaluation of the quasi-static mechanical bias at the center, taking into account quartz anisotropy.

According to Janiaud's calculation [15], the stress distribution is still a simple radial distribution and its expression at the center of the plate is :

$$T_{11}^o(\beta + \Psi) = \frac{\rho\delta}{\pi d t}\left(\frac{4A}{\delta + \varepsilon + \zeta} - \frac{A + B}{\delta - \zeta + \gamma}\right) \quad (7)$$

$$T_{33}^o(\beta + \Psi) = -\frac{\rho\delta}{\pi d t}\left(\frac{A + B}{\delta - \zeta + \gamma}\right) \quad (8)$$

$$T_{33}^o = 0 \quad (9)$$

where β is the azimuthal angle between the SAW and the force (Fig. 5)
Ψ is the propagation angle of the SAW
(ε, ζ, γ, δ are the following combinations of compliance constants expressed in the axes (a_1^o, a_3^o)

$$\varepsilon = 4(S_{11} - S_{33}) \quad (10)$$

$$\zeta = S_{11} + S_{33} - 2S_{13} - S_{55} \quad (11)$$

$$\gamma = 4(S_{15} + S_{35}) \quad (12)$$

$$\delta = 2S_{13} + S_{55} + 3S_{33} + 3S_{11} \quad (13)$$

Coefficients A and B are determined numerically by the static equilibrium conditions for the plate. These coefficients depend on crystal anisotropy and are computed for each set of cut angles (Φ,θ) and propagation direction Ψ.

Final expression of the relative frequency shifts for a single diametral compression of magnitude P is

$$\frac{\Delta F}{F_o} = \left(\frac{T_{11}^o(\beta + \Psi) + T_{33}^o(\beta + \Psi)}{2}\right)({}^s a_{11} + {}^s a_{33}) + ({}^s a_{11} + {}^s a_{33})\left(\frac{\overline{T_{11}^o(\beta + \Psi)} - T_{33}^o(\beta + \Psi)}{2}\right)\cos 2\beta - {}^s a_{13}\left(\frac{T_{11}^o(\beta + \Psi) - T_{33}^o(\beta + \Psi)}{2}\right)\sin 2\beta \quad (14)$$

In that case, the "mean value" corresponding to a 4-point symmetrical compression of magnitude P/2 is no longer independent of the angles β and Ψ and is given by

$$\frac{\Delta F}{F_o} = \left(\frac{|T_{11}^o(\beta + \Psi) + T_{33}^o(\beta + \Psi)|}{2}\right)({}^s a_{11} + {}^s a_{33}) \quad (15)$$

Experimental results for single and symmetrical diametral compression are given on Figs. 6 and 7.

Fig. 6 shows a series of results obtained with the cuts ($\Phi = 0, -10° < \theta < -2.5°, 30 < \Psi < 32.5°$). Stars represent the experimental values of single 2-point diametral compression of magnitude P and circles represent the experimental value for a 4-point symmetrical compression of magnitude P/2. As a reference, broken lines show the theoretical values given by the isotropic stress calculation.

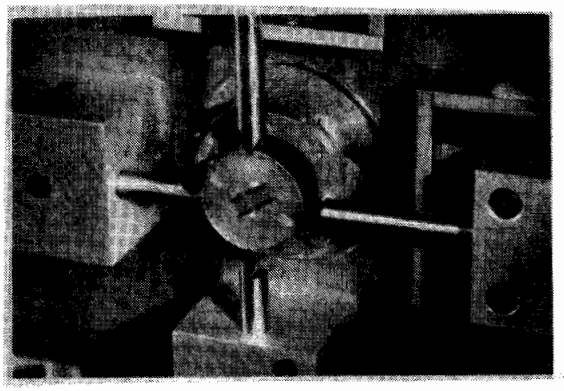


Fig. 4 : Experimental setup for single and symmetrical diametral compression of a circular quartz delay line

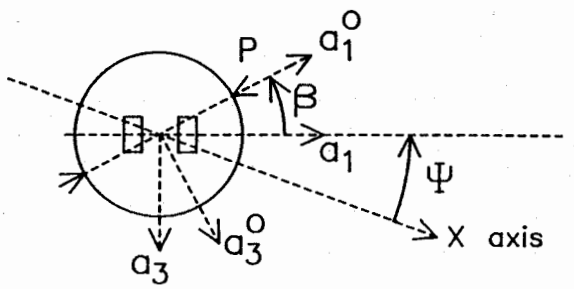


Fig. 5 :
Definition of the azimuthal angle β between the direction of SAW and the direction of diametral compression

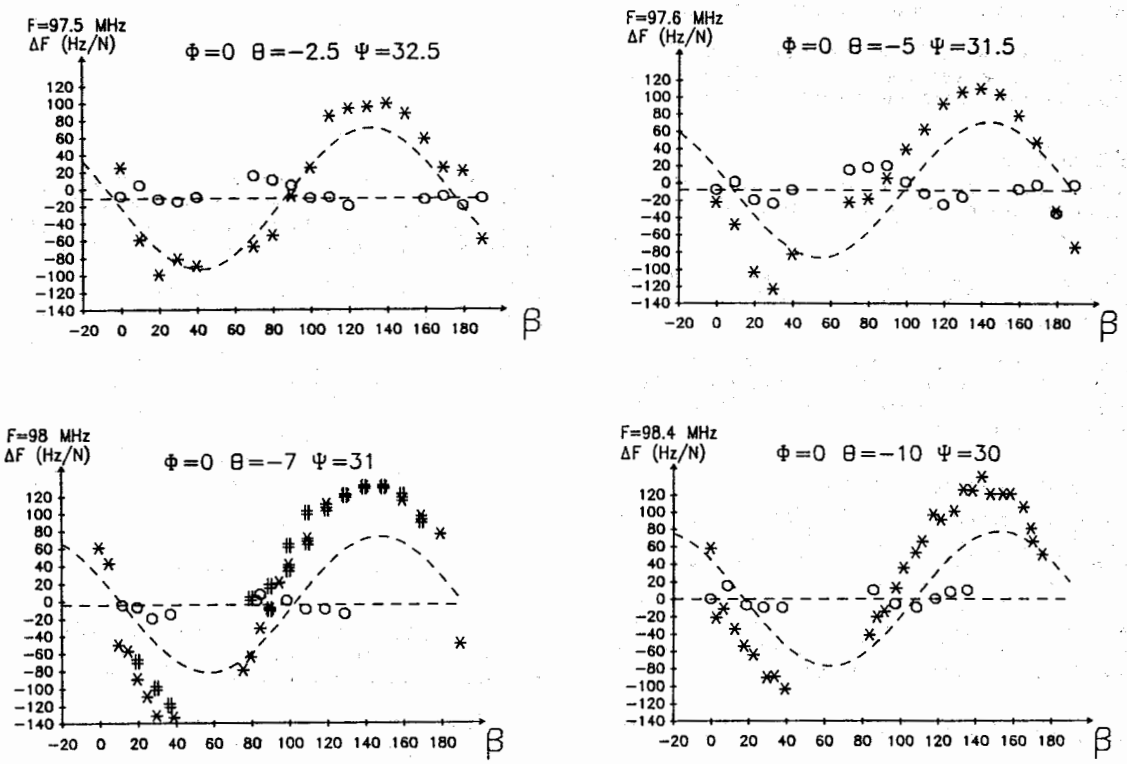


Fig. 6 : Experimental results of force-frequency effect on the cuts ($\phi = 0, -10° < \theta < -2.5°, 30° < \Psi < 32.5°$)
* : experimental values of 2-point compression
o : experimental values of 4-point compression
broken lines show the theoretical values (isotropic stress calculation)

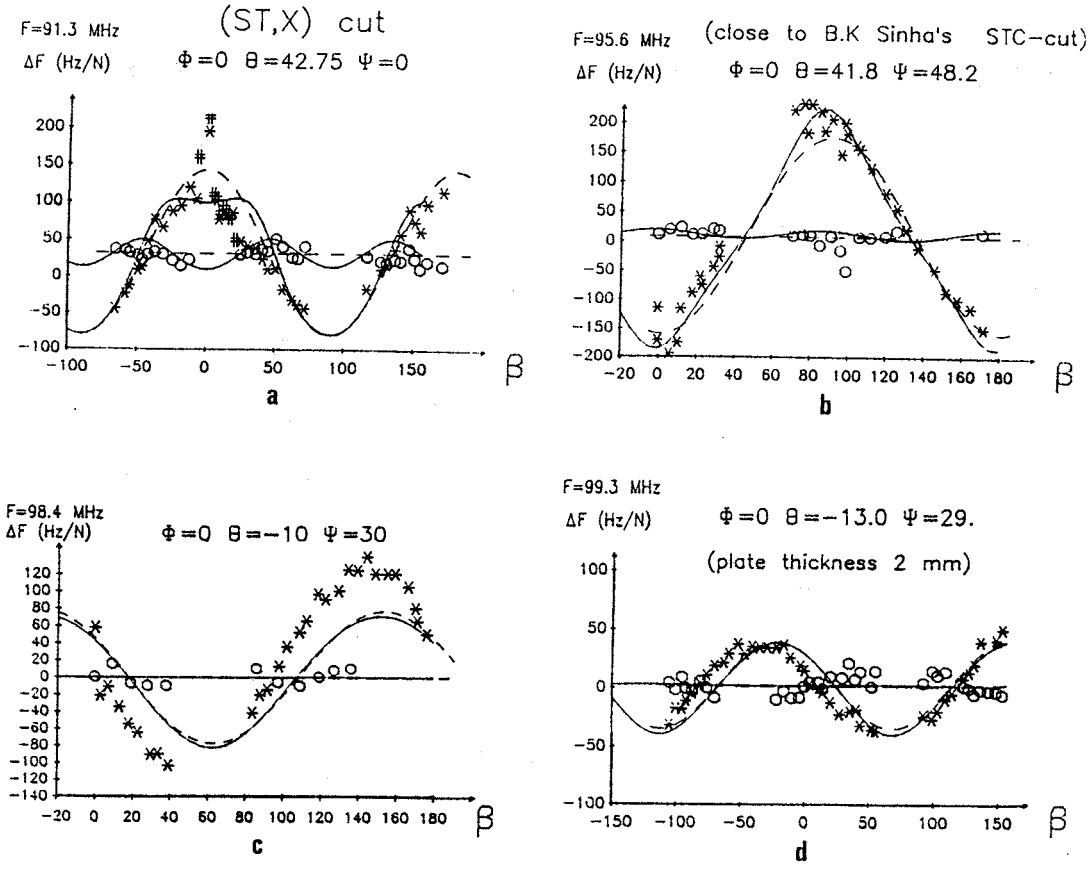


Fig. 7 : Experimental results of force-frequency effect on the cuts ($\phi = 0$) and
a) $\theta = 42.75$ $\Psi = 0$ (ST, X cut) b) $\theta = 41.8$ $\Psi = 48.2$ (close to STC-cut)
c) $\theta = -10$ $\Psi = 30$ d) $\theta = -13$ $\Psi = 29$
Results a) b) and c) are obtained with plate thickness = 1 mm, results d) with plate thickness = 2 mm
* : experimental values of 2-point compression
o : experimental values of 4-point compression
Solid lines show theoretical values with anisotropic stress calculation, broken lines = isotropic stress calculation

Table I
Comparison of planar isotropic stress-compensated cuts with the classical (ST,X) cut

Singly rotated cuts $\Phi = 0$	$\theta = -10°$ $\Psi = 30°$	$\theta = 41.8°$ $\Psi = 48.2°$	ST, $\theta = 42.75°$ X $\Psi = 0°$	
coupling factor $\Delta V/V$	4.9×10^{-4}	4.5×10^{-4}	5.8×10^{-4}	Theoretical values
beam steering angle	- 5.7°	$\simeq 6°$	0	
2nd order temperature coefficient	$-30 \times 10^{-9}\ K^{-2}$	$-30 \times 10^{-9}\ K^{-2}$	$-34 \times 10^{-9}\ K^{-2}$	Experimental values
sensitivity to planar isotropic stresses	$< 3 \times 10^{-12}\ m^2.N^{-1}$	$< 5 \times 10^{-12}\ m^2.N^{-1}$	$\simeq 12 \times 10^{-12}\ m^2.N^{-1}$	

The experimental values of the 2-point compression follow a sinusoïdal shape, although a certain discrepancy in amplitude and phase exist between theory and experiment. To investigate the possible origin of the discrepancy, the following attemps were made :

- using the anisotropic stress calculation
- increasing plate thickness to check the influence of parasitic bending forces.

Comparison between isotropic and anisotropic stress calculation is shown on Fig. 7 where solid lines represent theoretical values of the anisotropic stress calculation (broken lines = isotropic calculation).

The difference between the two models depends on the crystal cut. In the case of the ST-cut (Fig. 7a) the anisotropic model predicts a well-marked oscillation in the case of a 4-point compression. The average value of the oscillation corresponds to a constant value given by the isotropic stress calculation.

The difference between the two models is also significative in the case of Fig. 7b ($\Phi = 0, \theta = 41.8, \Psi = 48.2$). In the case of the cuts ($\Phi = 0, \theta \simeq -10°, \Psi \simeq 30°$) the two models predict almost the same result, and do not explain the discrepancy in amplitude and phase between theory and experiment on Fig. 7c. Fig. 7d shows the investigation on a thicker plate (2 mm thick) with a better agreement between theory and experiment ; this could be explained by a significant reduction of parasitic bending forces with respect to the previous experiments (1 mm thick).

Table 1 summarizes the experimental results obtained.

Conclusion

Experimental tests were performed on two series of cuts
$(-13° < \theta < -2{,}5°, 29° < \Psi < 33°)$
$(\theta = 41.8°, \Psi = 48{,}2°)$

Temperature-compensated cuts were found in close vicinity of the theoretical predictions. Stress compensation was observed on these cuts in the case of a 4-point symmetrical compression. The upper limit of sensitivity to planar isotropic stresses was estimated experimentally four times lower for the cut ($\theta = -10°, \Psi = 30°$) than for the classical (ST,X) cut. For the cut $\theta = 41.8°, \Psi = 48.2°$), stress sensitivity has been experimentally found to be reduced by a factor 2 at least with respect to the ST cut. Different cuts in the vicinity of ($\theta = 40°, \Psi = 45°$) are under test to check for a possible better stress-compensation.

Next step will be an investigation of dynamical temperature effects in SAW devices built with the new cuts presented here. Such devices should exhibit a reduced sensitivity to thermoelastic stresses in the case of a symmetrical distribution of temperature gradients.

Acknowledgement

We are very grateful to R. Coquerel and C. Bonjour for SAW-delay line fabrication.

References

[1] D.V. Schick and H.F. Tiersten, "An analysis of the in-plane acceleration sensitivity of ST-cut quartz surface wave resonators with the substrate extending beyond the supports", Proc. 42nd Ann. Freq. Cont. Symp., IEEE cat. n° 88 CH 2588-2, pp. 230-238.

[2] J.A. Greer and T.E. Parker, "Improved vibration sensitivity of the all quartz package surface acoustic wave resonator", Proc. 42nd Ann. Freq. Cont. Symp., IEEE cat. n° 88 CH 2588-2, pp. 239-251.

[3] E. Bigler, G. Théobald, D. Hauden, "Dispositifs à ondes élastiques de surface insensibles aux contraintes planes", Proc. 1st European Time & Frequency Forum, Besançon, 1987, pp. 311-314.

[4] E. Bigler, R. Coquerel, D. Hauden, "Experimental temperature and stress sensitivities of surface acoustic wave quartz cuts", Proc. 42nd Ann. Freq. Cont. Symp., IEEE cat. n° 88 CH 2588-2, pp. 224-229.

[5] E. Bigler, R. Coquerel, D. Hauden, "Stress sensitivity mapping for surface acoustic waves on quartz", IEEE Trans. Ultr. Ferro. Freq. Contr., vol. UFFC-36, n° 1, pp. 57-62 (jan. 1989).

[6] E.P. Eernisse, "Calculations on the stress-compensated (SC-cut) quartz resonator", Proc. 30th Ann. Freq. Cont. Symp., 1976, pp. 8-11.

[7] H.F. Tiersten and J.C. Baumhauer, "Nonlinear electroelastic equations for small fields superposed on a bias", J. Acoust. Soc. Am., vol. 54, n° 4, pp. 1017-1074 (oct. 1973).

[8] H.F. Tiersten and B.K. Sinha, "A perturbation analysis of the attenuation and dispersion of surface waves", J. Appl. Phys., vol. 49, n° 1, pp. 87-95 (jan. 1978).

[9] D. Hauden, M. Planat and J.J. Gagnepain, "Nonlinear properties of surface acoustic waves : applications to oscillators and sensors", IEEE Trans. Sonics Ultrason., vol. SU-18, n° 5, pp. 342-348 (sept. 1981).

[10] IEEE Standard on Piezoelectricity 176-1949, Proc. IRE, vol. 37 (1949), pp. 1378-1395.

[11] H.F. Tiersten and B.K Sinha, "Temperature derivatives of the fundamental elastic constants of quartz", Proc. 32nd Ann. Freq. Cont. Symp., 1978, pp. 150-154.

[12] B.K. Sinha and H.F. Tiersten, "On the temperature dependence of the velocity of surface waves in quartz", J. Appl. Phys., vol. 51, n° 1, pp. 659-4665 (jan. 1980).

[13] B.K. Sinha, "A stress and temperature compensated orientation and propagation direction for surface acoustic wave devices", IEEE Trans. Ultrason. Ferro. Freq. Cont., vol. UFFC-34,n° 1, pp. 64-74 (jan. 1987).

[14] S.P. Timoshenko and J.W Goodier, in "Theory of Elasticity", p. 122, 3rd edition (International Student Edition), McGrawHill, ed. (1970).

[15] D. Janiaud, "Modélisation de l'influence d'une accélération sur la fréquence des résonateurs à quartz", Thèse, Université de Besançon, 1978, pp. 63-67.

ALL QUARTZ PACKAGED S.A.W. RESONATOR *

S. Calisti**, Ph. Defranould**, L. Penavaire**, D. Hauden***, and A. D'Almeida***

** TH SINTRA ASM DTAS, 399 route des crêtes BP 38, 06561 VALBONNE CEDEX (FRANCE)

*** LPMO CNRS, 32 avenue de l'observatoire, 25000 BESANCON CEDEX (FRANCE)

ABSTRACT

In order to obtain a new kind of device for military applications, UHF frequency source or sensor, we have developed an "all quartz" package for S.A.W. resonators, which allows considerable latitude in the choice of mounting and should reduce an oscillator's sensitivity to acceleration. This package is achieved by the use of a low temperature (400 to 500°C) firing glass to seal a quartz cover above the substrate, both cover and substrate coming from the same ST cut material. This concept was previously introduced by PARKER et al. [1].
We will first describe how we have built the final device, starting from a 320 MHz center frequency etched groove resonator which includes improved metallisation layer and ending with a device sealed under clean high vacuum. A second part will deal with acoustical measurement and will show quality factors up to 20000 obtained on the first devices made. Other results on temperature, short and long term stability will also be presented, including a mesurement of the accelerometric sensitivity of the device.

1.- INTRODUCTION

Today a new generation of aeronautical materials have become available with improved flying performances obtained by the fly-by-wires system which allows high G factors on improved mechanical structures.
These performances called for new RADAR materials which matched the flying capacity, this led to electronically scanned antennas, less sensitive to vibrational factors. They also asked for the detection of smaller targets at increasing range, which called for improved performances for the electronic system such as a lower vibration sensitivity and noise reduction.

Until now most of the RADAR UHF frequency synthetizers used bulk acoustic wave resonators, these devices have an upper frequency limitation and the need for electronic frequency multiplication which increases the noise of the final output signal; however they benefit from improvements brought about through 50 years of research and developement on this topic and they are still competitive.

Surface acoustic wave resonators which allow higher fundamental operating frequency (greater than 600 MHz) and reduce frequency mutiplication by at least a factor of two with a corresponding reduction of the electronic flicker noise. They also exhibit at least a 10 dB improvement in resonator noise floor because of the higher power which can be used in the oscillator loop.
Until recent times this advantage was counteracted by other technological limitations such as a high sensitivity to environmental conditions: surface contamination by moisture or absorption of contaminants even with the use of hermetic sealing, difficulties in obtaining a stress free mounting of the device when a high level of acceleration is present (typically a RADAR environment). Frequency trimming was also a problem until reactive ion etching technology become available to solve it. A selective etching of the quartz material was used to adjust the frequency of the devices, by increasing the mass loading affect of the electrodes [2].

A solution to most of the environmental problems was probably brought out by PARKER et al.[1], [3] at the end of a long study that led to the concept of an "all quartz" package for the resonator (see fig.1). In this stucture a quartz cover is sealed over the acoustic part of the surface acoustic wave resonator, by use of a low temperature firing glass, to achieve a hermetically sealed device. Both cover and substrate belong to the same quartz orientation to avoid expansion coefficient mismatch. The electrical connections from the outside to the transducers are made via conductor lines which extend out of the cover to the exposed part of the substrate.

We began this study following the path of PARKER et al., with two main applications in mind, resonators for oscillators and accelerometric sensors. For the second application the advantages of the all quartz package became obvious as the resonator structure needed to be mounted in a complex mechanical fixture which is generally not well adapted to a clean environmental conditions. Of course the resonator structure was quite different from that used in simple oscillator applications as is shown in fig 2. The resonator is situated in the middle of a plate with the quartz cover right above the sensitive part of the device, the accelerometric mass is attached to one end of the plate while the other end is used for the electrical connections and mechanical support.

2.- MANUFACTURING OF THE STRUCTURE

Two main problems had to be solved for manufacture of an all quartz package structure: the choice of the sealing glass and the deposition method.

Choice of a glass

As cover and substrate belong to the same material the seal should avoid adding mechanical stress on the structure. Most of the stress will arise during the cooling phase from thermal gradients after the high temperature sealing. As the quartz has anisotropic thermal expansion coefficients we first investigated the values of the linear thermal expansion coefficient in the two principal directions i.e. parallel and perpendicular to the X axis for three values of the cut angle: 33, 36, 42 degrees from Y. The values are given in the table of fig. 3. The values fall between 13.7 and 9.27 ppm/°C leading to an average value of 11.6 ppm/°C. We limited our search to sealing glasses with a firing temperature ranging from 350 to 500 °C and with a linear expansion coefficient of 12 ppm/°C in the -50 to +500 °C temperature range. Such a material was hard to find since most of the materials available on the market are used for CERDIP packaging sealing and have linear expansion coefficients of about 6.7 ppm/°C. We eventually found two glasses which exhibit the required properties (see table 1).

* This work was supported by D.R.E.T.

The glasses are available in powdered form with a granulometry of 345 MESH. Next we had to choose a deposition method.

Deposition mode

After a few experiments and an analysis of the problem which took into account the geometries of the sealing shape we decided to use the screen-printing technique. We prepared the glass in a serigraphic paste, choosing a mixture of a resinoid compound, α Terpineol and glass according to the flow chart of fig 4.

Manufacturing of the wafers

In a previous study [4] we have undertaken a research on resonator's aging using 320 MHz devices as a test vehicle, we decided to reuse the same device design for our first tests on "all quartz" packaging. We have processed a few wafers of 42° rotated Y cut quartz posessing an intrinsic Q of $2x10^6$. The final device dimensions were chosen for use as an accelerometric sensor (see §1) being physically longer than a standard resonator.
The covers and resonator substrates were made in two parts according to the flow chart given in fig. 5. The final assembly of cover and substrate was made under high vacuum (10^{-7} Torr) in a specially designed apparatus which has the capability of allowing substrate and cover alignment under vacuum.

Control of the hermeticity

The first control, gross leak, was made with a bubble tester which identified most of the defective devices. The second control, fine leak, was impossible to perform with a standard He leak tester. Because of the small size of the cavity under vacuum, the test must be performed by electrical measurement of the unloaded Q factor of the resonator which is very sensitive to the vacuum quality. The results will be presented in the next section.

3.- ELECTRICAL MEASUREMENT

Using the technology described above we have manufactured a group of devices. On each device we have performed three measurements : the unloaded and loaded Q values, the turn over temperature and aging over a period of 10 days.

Transmission curves

A typical transmission curve is shown on the fig. 6. Below the curve are the numerical values which were used for the Q determination. We obtained an unloaded Q value (Q_u) of 19841. The measurement was made a few weeks after sealing. As the cavity is very small we conclude that the leaking rate was well below 10^{-9} Torr.l.s^{-1}, otherwise we would have reached atmospheric pressure inside the cavity within this time resulting in a Q_u value well below that obtained. The loaded Q factor Q_l was 7525 and the center frequency of the devices was 320.338 MHz, the insertion loss being in the 4 dB range.

Turnover measurement

Turnover temperature was measured for two devices by measuring the center frequency of each resonator at 10° increments from -10 to +50 °C. A parabolic curve was then fitted to the data for each device and a turn over temperature calculated (see fig. 7). The turnover values of 16.35 and 14.86 °C were obtained, well in accord with 42 degrees rotated Y cut quartz, the 1.5 °C difference betwen the two devices being a result of the fabrication dispersion.

Long term aging

Measurements were also performed on two devices mounted in an oscillator situated in a temperature controled environment. Some difficulties were encoutered, as the turn over point was at a relatively low temperature resulting in the need for a low temperature thermostat. The curves showing the frequency variation, in parts per million versus time (in weeks) are given in fig. 8. They show clearly a very slow drift with calculated slopes of 0.042 ppm and 0.0083 ppm per day. When we extrapolate these measurements we obtain 15.3 and 3.0 ppm per year. As these results refer only to the first 10 days of aging we may expect a stabilisation in the next few weeks and expect to obtain smaller rates in 1 or 2 months. In any case these first results were better than those we obtained in our previous experiments, they were probably due to the thermal step during the sealing procedure which resulted in a relaxation of mechanical stresses.

4.- CONCLUSIONS

Results on all quartz package resonators have been presented. Excellent hermeticity and aging rates were obtained. The result was low insertion loss and high Q factors.
The next stage of our resonator development will be the design and fabrication of devices for oscillator application. The new designs will additionally take account of the anisotropic expansion factors which will result in a reduction of the accelerometrc sensitivity of the resonator [5], and we can expect to obtain vibration sensitivity below 10^{-9} per G in the direction perpendicular to the plane of the resonator in the 100 to 10000 Hz vibration frequency range. This will lead us to the realisation of an oscillator for RADAR applications.

ACKNOWLEDGEMENT

The authors wish to acknowledge in particular M. SANDOZ from Thomson B.C.S. for helpful discussions on all the technological problems.

BIBLIOGRAPHY

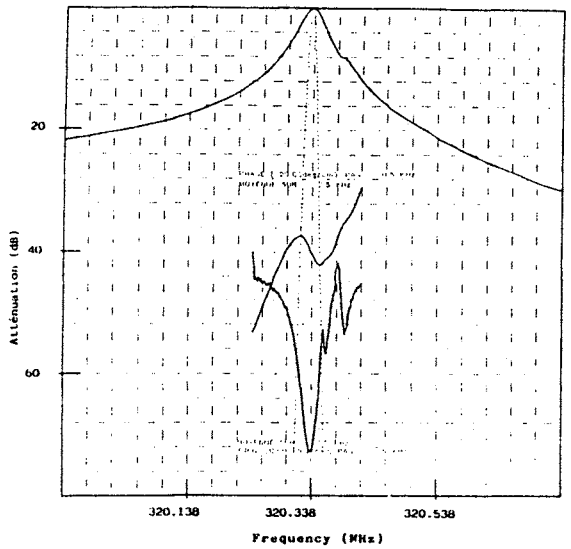
[1] T.E. PARKER, J. CALLERAME and J.K. MONTRESS "A New All Quartz Package For SAW Device" Proc. of the 1985 IEEE Ultrasonic Symposium

[2] W.J. TANSKI "Surface Acoustic Wave Frequency Trimming of the Resonant and Traveling-Wave Devices on Quartz", Appl. Phys. Lett. 39(1) July 1981

[3] T.E. PARKER and G.K. MONTRESS "Precision Surface-Acoustic-Wave (SAW) Oscillators", IEEE Transactions on Ultrasonic Ferroelectric and Frequency Control, Vol.35 N° 3 May 1988

[4] S. CALISTI Ph. DEFRANOULD D. HAUDEN "Influence des paramètres technologiques sur la stabilité en fréquence des résonateurs à ondes de surface", Proc. of the 1st European Time and Frequency Forum, March 1987

[5] H.F. TIERSTEN and D.V. SHICK "An Analysis of the Normal Acceleration Sensitivity of the ST-Cut Quartz Surface Wave Resonators with the Substrate Extending Beyond the Supports", Proc. of the 2nd European Frequency and Time Forum, March 1988



BANDE A -XX,XX dB	FREQUENCE INF. DE LA BANDE	FREQUENCE SUP. DE LA BANDE	LARGEUR DE BANDE	FREQUENCE CENTRALE
1.000dB	320.324 MHz	320.345 MHz	0.021 MHz	320.334717 MHz
2.000dB	320.319 MHz	320.350 MHz	0.031 MHz	320.334229 MHz
3.000dB	320.313 MHz	320.354 MHz	0.041 MHz	320.333740 MHz
4.000dB	320.308 MHz	320.360 MHz	0.052 MHz	320.334015 MHz
5.000dB	320.302 MHz	320.365 MHz	0.063 MHz	320.333501 MHz
6.000dB	320.296 MHz	320.370 MHz	0.074 MHz	320.333009 MHz

FIGURE 6: Transmission curve of the n°3 resonator (ambient temperature)
The unloaded Q factor Q_u is 19841
The loaded Q factor Q_l is 7525
Insertion loss is 4 dB

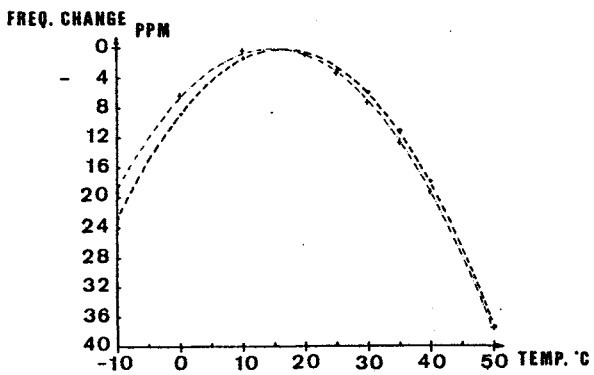


FIGURE 7: Turn over measurement of two devices

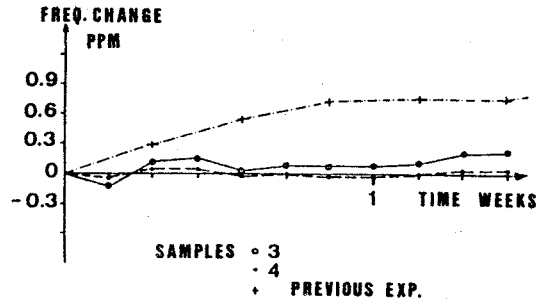


FIGURE 8: Long term aging, comparision between two "all quartz" packaged resonators and a standard hermetically packaged resonator from a previous experiment (data available for 1 week duration)

	GLASS		NEEDS
	SG 201	SG 309A	
LINEAR THERMAL EXPENSION COEFFICIENT in PPM per °C	11.09	10.5	12
TEMPERATURE RANGE °C	30 - 250	25 - 310	25 - 400
SEALING TEMPERATURE °C	450	480	LOWEST

TABLE 1: A compararison between two available glasses and our need

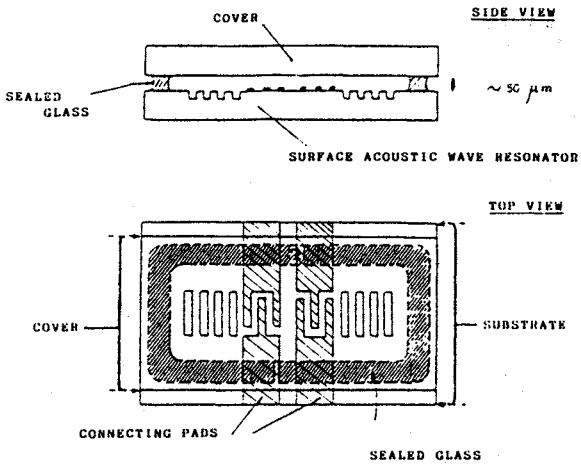


FIGURE 1: "All quartz" packaged resonator

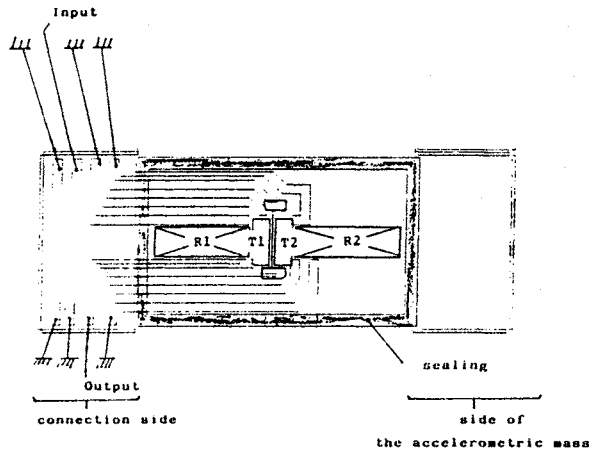
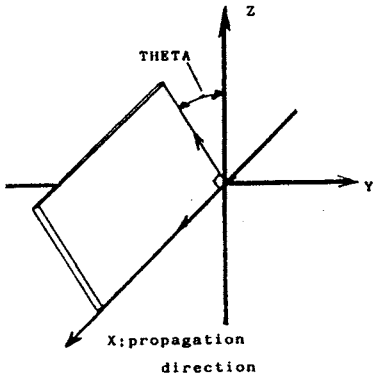


FIGURE 2: Accelerometric design for an "all quartz" packaged accelerometric sensor



ROTATION ANGLE THETA DEGREES	LINEAR THERMAL EXPANSION COEFFICIENT PARRALLEL TO X AXIS (PPM)	PERPENDICULAR TO X AXIS (PPM)
42	13.7	10.2
36	id.	9.58
33	id.	9.27

AVERAGE VALUE 12 PPM

FIGURE 3: Orientation of quartz plates and linear thermal expension coefficients

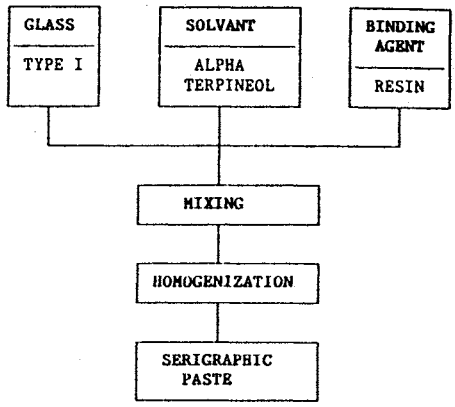


FIGURE 4: Simplified preparation method tor the serigraphic paste

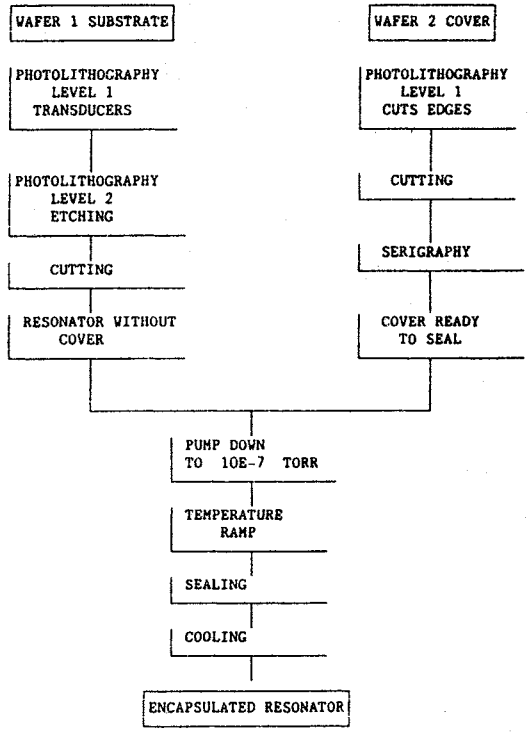


FIGURE 5: Simplified manufacturing flow chart

A NEW TECHNIQUE FOR THE TEMPERATURE COMPENSATION OF SAW OSCILLATORS

M.P.Cracknell, A.M.Harrison, D.J.Sharpe

STC Components, Quartz Crystal Division, England

ABSTRACT

A new method for the compensation of Surface Acoustic Wave (SAW) oscillators is described. This work is a development of a technique first used for the compensation of oscillators employing AT-cut bulk wave resonators. The benefits that result for SAW oscillators are high compensation performance, low power consumption, low complexity and relatively small physical size. Compensation is achieved by the application of a control voltage to the oscillator, such that the oscillators frequency remains constant for variations in temperature. A custom bipolar integrated circuit has been designed which generates a power series representation of the control voltage. Prototype oscillators have been constructed at 600 MHz and show a compensated performance of +/- 3ppm over a 120 deg.C operating temperature range. The potential to improve the performance to better than +/-2ppm exists.

INTRODUCTION

In recent years there has been a demand for oscillators of increased stability, higher frequency, reduced size, lower power consumption and good phase noise. To meet the demand for higher frequencies, SAW resonators and delay lines have been used to produce oscillators to well over 1 GHz. The major drawback of SAW oscillators compared with lower frequency bulk wave devices is the temperature performance of the oscillator. In both cases the overall performance is dominated by the temperature performance of the resonant element. A bulk wave resonator can be optimised by selecting the angle of cut used for the quartz blank to give an overband performance of +/-15ppm over -40 deg.C to +80 deg.C, a similarly optimised SAW device will achieve +/-60ppm at best.

Bulk wave oscillators have traditionally been given improved stability by the use of temperature compensation or ovenisation. Compensation is achieved by the use of an oscillator whose frequency can be varied by the application of a d.c. voltage being combined with a circuit which generates a voltage that varies with temperature (typically a network of thermistors and resistors). This voltage varies in such a way as to maintain the oscillators frequency at a near constant value throughout the operating temperature range.

SAW resonators and delay lines are used to produce voltage controlled oscillators with linear pulling characteristics. Such devices are highly suited to temperature compensation. Various schemes for compensation have been proposed and show that compensation to better than +/-10 ppm can be achieved over a 100 deg.C temperature range. The subject of this paper is the application of IC derived compensation, developed initially for lower frequency bulk wave oscillators, to a SAW delay line oscillator at a frequency of 600 MHz, the objective being to improve on previously presented results.

INTEGRATED CIRCUIT

The major reasons for the development of an IC based temperature compensation scheme for bulk wave oscillators, were that both improved compensation performance and a significant reduction in overall size may be achieved when compared to thermistor based designs.

An integrated circuit has been designed which generates a series of five voltages, from one which is constant with temperature through to one having a fourth order variation with temperature. These are represented graphically in fig.1. These curves are generated in the form of chebychev polynomials, the equations of which are:

$$V0(t) = 1$$
$$V1(t) = t$$
$$V2(t) = 2t^2 - 1$$
$$V3(t) = 4t^3 - 3t$$
$$V4(t) = 8t^4 - 8t^2 + 1$$

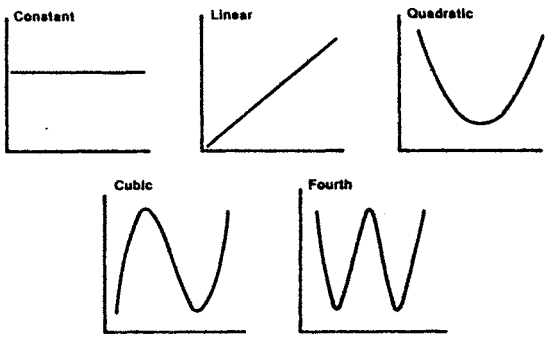


Fig.1 IC VOLTAGES

The desired compensation voltage can be obtained by applying suitable weighting coefficients.

$$V(t)=a0V0(t) + a1V1(t) + a2V2(t) + a3V3(t) + a4V4(t)$$

For a bulk wave oscillator using AT-cut crystals the dominant term will be V3(t). With a SAW device the dominant term will be V2(t).

A block diagram illustrating the various elements of the IC is shown in fig.2. The constant voltage is generated by a voltage stabiliser which incorporates a bandgap reference. The linear voltage is generated using the temperature coefficient of the base-emitter junction voltage of a transistor that is operated at constant emitter current. Conventional analogue multipliers comprising four transistors with the collectors cross-coupled are used to generate the quadratic, cubic and fourth order voltage variations with temperature. The summing amplifier is also included in the IC, along with the active devices necessary to fabricate a rudimentary low frequency bulk wave oscillator.

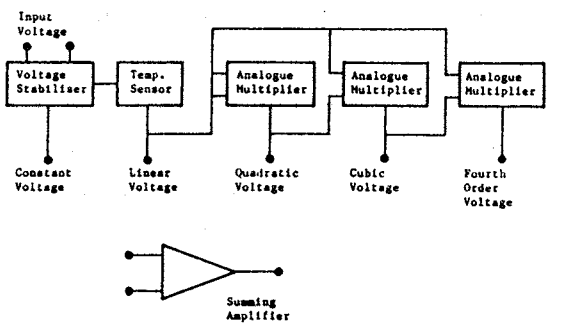


Fig.2. INTEGRATED CIRCUIT SCHEMATIC

INTEGRATED CIRCUIT APPLICATION

In order to compensate an oscillator using this method, the values of six resistors have to be determined. These resistors set the value of coefficients a0 to a4. A computer program is used to measure the five output voltages from the IC, and the compensation voltage required to keep the oscillator at the correct frequency. The measurements are repeated at a number of points throughout the operating temperature range. The program then calculates the values of the resistors required. This program uses the exchange algorithm to derive the minimax linear approximation for the set of measurement points. The problem is similar to that of solving a set of simultaneous equations with no exact solution; the coefficients are adjusted until the errors between the synthesised compensation voltage and the required compensation voltage are of a specified form. This approximation minimises the maximum error between the synthesised compensation voltage and the desired compensation voltage. Once the resistors are fitted the performance of the oscillator is verified by measuring the frequency variation of the unit over the operating temperature range.

A more detailed description of the chip and its operation is given by Wilson (2).

With this technique, better than +/-0.5ppm is being achieved over military temperature bands for bulk wave crystal oscillators. A typical result, including hysteresis is shown in fig.3.

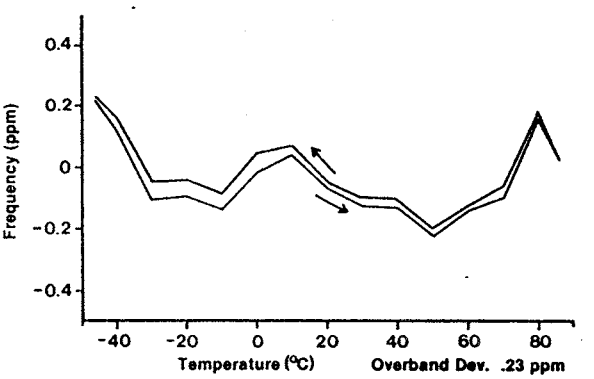


Fig.3. COMPENSATED 5 MHz OSCILLATOR

SAW VCO

STC has developed a range of SAW Voltage Controlled Oscillators (VCO's) around 600 MHz with a frequency pulling range of up to +/-350ppm for a control voltage range of 0 to 12V. The overband performance of the oscillator is shown in fig.4. The typical parabolic characteristics of the SAW have been offset slightly by the oscillator circuitry to give a 150ppm maximum variation over -40deg.C to +80deg.C.

A SAW delay line with an insertion loss of 15 dB and a Q of 550 has been used. The maintaining amplifier is a Thickfilm hybrid and the buffer amplifier is a silicon MMIC. The phase shifting network consists of three variable capacitance diodes with associated inductors and capacitors. The use of three diodes in this design was necessary to achieve wide pulling and the high linearity that was required for another application. In a VCO designed specifically for temperature compensation the pulling range could be considerably reduced, leading to a simplification of the circuit design and yielding reduction in overall size.

The output from this VCO is a sinewave having an amplitude of 0 dBm into a 50 ohm load. The device requires a current of approximately 55mA from a 12v supply.

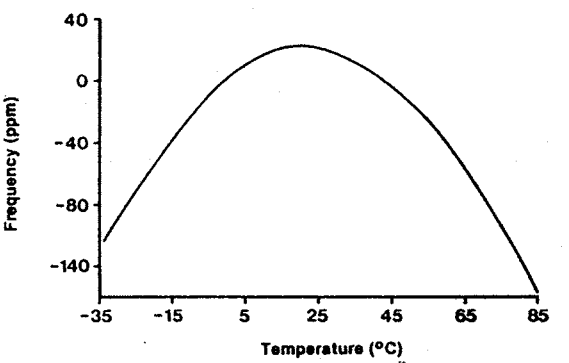


Fig.4. 600 MHz SAW VCO TEMPERATURE PERFORMANCE

SAW OSCILLATOR COMPENSATION

The results obtained from the investigation of the SAW VCO indicated that it should be possible to compensate this device to achieve a high degree of stability. Before carrying out a practical experiment however, the characteristics of the oscillator were combined with those of a typical IC and a theoretical result obtained. This result is shown in fig.5 and indicates that a frequency tolerance of better than +/-2ppm may be achieved.

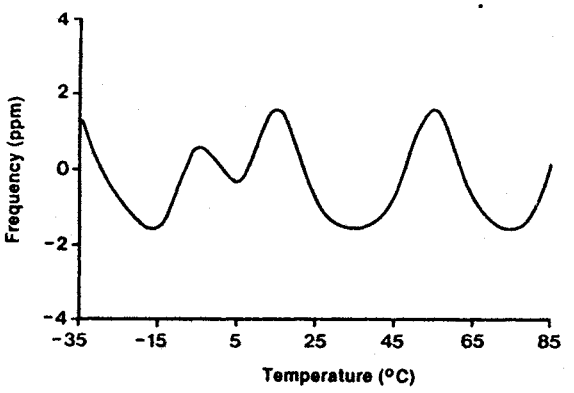


Fig.5. 600 MHz OSCILLATOR THEORETICAL COMPENSATION

One of the available SAW oscillators was incorporated with an IC in a module and a compensation run carried out. As previously explained this temperature run characterises the whole circuit, determines the resistor requirements and predicts the best performance that can be achieved provided all components fitted are ideal. The result is shown in fig.6 giving a compensation stability prediction of just over +/-2ppm. The resistors were then fitted and a verification run carried out resulting in a practical performance of +/-3ppm as shown in fig.7 It is expected to be able to bring this result nearer to the compensation result by slight alteration to the resistor values. This procedure is sometimes necessary to compensate for resistor tolerances and the loading effects of the resistors due to the finite impedance of the voltage sources. This technique is already applied to bulk wave devices when it is necessary to meet very tight tolerances.

The phase noise performance of the oscillator has yet to be evaluated. However, we believe that if the oscillator without compensation were to be compared to the same oscillator with the addition of temperature compensation circuitry, no degradation of the phase noise performance would be noted.

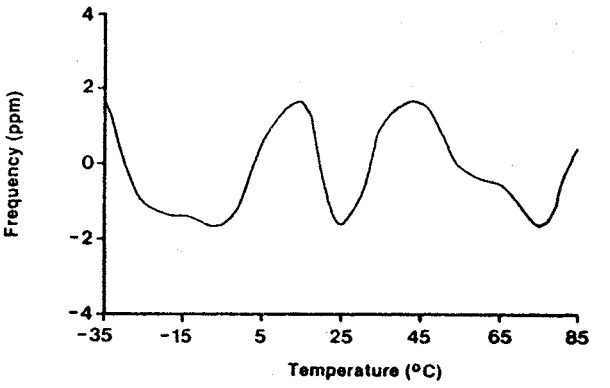


Fig.6. 600 MHz OSCILLATOR COMPENSATION RESULT

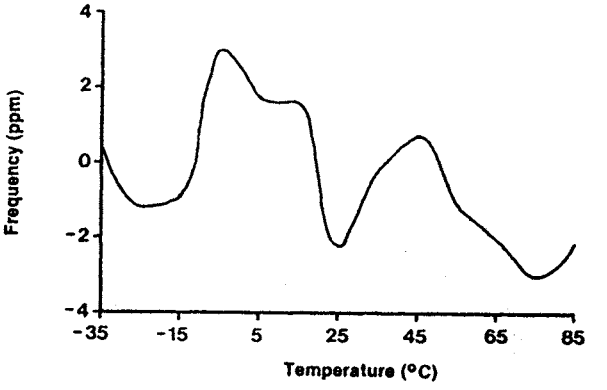


Fig.7. 600 MHz COMPENSATED OSCILLATOR;

PRACTICAL RESULT

CONCLUSIONS

It has been shown that it is possible to apply this compensation technique, developed for bulk wave oscillators, to much higher frequency SAW oscillators.

Prototype oscillators have been constructed at 600 MHz and show a compensated performance of +/-3ppm over a 120 deg.C operating temperature range.

The oscillator circuit used was not optimum for this application and improvements can be expected from a redesign to reduce the pulling sensitivity. This will increase the inherent stability of the device and reduce the susceptibility of the device to errors in the compensation voltage produced by the IC.

FURTHER WORK

Based on this initial work STC is undertaking a development program to produce a range of temperature compensated SAW oscillators at various frequencies between 400 MHz and 1 GHz.

A provisional specification follows:

Frequency Range	400 MHz to 1GHz
Temperature Range	-35 deg.C to 85 deg.C
Size	35 * 35 * 15mm
Supply Voltage	12V +/- 5%
Supply Current	<75 mA
Frequency Adjustment	>+/- 20ppm
Frequency Stability with:	
1. Temperature	<+/-3ppm
2. Supply Voltage	<+/-0.5ppm
3. Ageing	<+/-5ppm in first year
	<+/-2ppm/year thereafter
Output Power	> +7 dBm into 50 ohms
Harmonics	<-20 dBc
Spurious	<-70 dBc

ACKNOWLEDGMENTS

The authors wish to thank Dr N D Hardy for his help and advice during the preparation of this paper.

REFERENCES

1. James S Wilson "Precision Crystal Oscillator Design Thesis for Degree of Doctor of Philosophy". University of Aston, October 1983

2. J S Wilson "An Improved Method for the Temperature Compensation of Crystal Oscillators". Proceedings of the 37th Annual Frequency Control Symposium (1983) p.442-447.

3. R G Kinsman "Temperature Compensation of Crystals with Parabolic Temperature Coefficients". Proceedings of the 32nd Annual Frequency Control Symposium (1978). p.102-107.

4. R K Asatourian, M K Kilcoyne and M E Montamedi "Performance Characteristics of Temperature Compensated UHF SAW Oscillators for Satellite Communications". Proceedings of the 1984 Ultrasonics Symposium. p.186-190.

5. J Ladd, C Abdallah and T O'Shea "A Temperature Compensated L-Band Hybrid SAW Oscillator and Resonator Filter". Proceedings of the 1984 Ultrasonics Symposium. p.191-196.

6. A J Slobodnik,Jr., R D Colvin, G A Roberts and J H Silva "A Digitally Compensated SAW Oscillator. Proceedings of the 1981 Ultrasonics Symposium. p.135-138.

7. D L Lee "Design Considerations for Electronically Compensated SAW Delay Line Oscillators". Proceedings of the 1979 Ultrasonics Symposium. p.849-854.

8. T I Browning and M F Lewis "A Novel Technique for Improving the Temperature Stability of SAW/SBAW Devices". Proceedings of the 1978 Ultrasonics Symposium. p.474-477.

INVESTIGATION OF SAW OSCILLATOR STRUCTURES

P. Leblois, J. Groslambert

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

At present, more and more devices require an UHF source with good temporal and spectral characteristics. SAW quartz resonators are very attractive for these low phase noise oscillators because the excitation power applied, before nonlinear effects appear, can be much greater than for BAW resonators. Because the frequency stability of an oscillator is of the greatest importance for many applications one must study the different phase noise sources of each component in open loop before making an oscillator.

In the first part of this paper an experimental method of characterizing S.A.W. resonators is presented. Such measurements must be done carefully because the results obtained vary strongly with the measurement conditions. Next, phase noise and gain compression measurements on loop amplifiers are reported as a function of input power level. These results are used to compare various amplifiers and to select the best. From the amplifier gain compression curve and the loss in each electronic component, it is then possible to choose the optimum excitation level in the loop. Another difficulty in S.A.W. oscillators is adjustement of the frequency. To address this problem, several ways of adjusting the oscillation frequency have been tested, and their influence on the phase noise measured. A number of oscillators will be built and tested, especially for their frequency stability and phase noise as a function of power level in the loop, following the procedures described here.

Introduction

The demand for Surface Acoustic Wave (SAW) oscillators is increasing rapidly because their spectral purity is better than for Bulk Acoustic Wave (BAW) oscillators. However, to satisfy certain stringent requirements, the short term stability and aging of these oscillators must be improved. Many studies have shown the manufacturing process and the technology used in fabricating the resonator play a large part in determining oscillator stability [1, 5]. In this paper, the effect of electronic components on the oscillator characteristics is demonstrated by investigating each separate part of the oscillator. Several specific points, such as the electrical model of a SAW resonator, the power level in the loop, and frequency trimming, have been studied.

Characterization of the SAW resonator

A characterization of the SAW resonator must be performed because the frequency stability of the oscillator depends strongly on that of the SAW resonator. Moreover, to analyse the oscillator, all the parameters which model the SAW resonator must be known. The equivalent electrical diagram of a SAW resonator is built on a series resonant circuit R_1, L_1, C_1 which means that impedance is minimal at the resonance (Fig. 1). A capacitor C_0 must be added to this resonant circuit in parallel with the resonator input. This capacitor represents the capacity of the interdigital electrode transducer.

This work has been supported by the Direction des Recherches Etudes et Techniques (DRET).

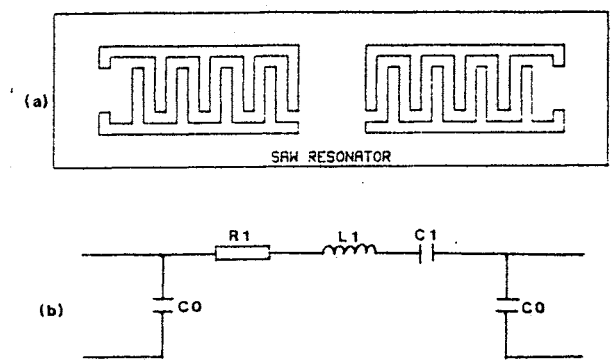


Fig. 1 : The electrical equivalent circuit of a SAW resonator (a) is a series resonant circuit R_1, L_1, C_1 and two inter-electrode capacitors C_0 (b)

Usually, the inter-electrode capacitor C_0 is small and its influence is negligible near the resonance frequency. Thus, measuring the resonance frequency f_0, the resonance attenuation $S_{21}(f_0)$ [6] and the -3 dB sideband Δf is sufficient to calculate the motional components :

$$R_1 = 2R_0\left(\frac{1}{S_{21}(f_0)} - 1\right)$$

where R_0 is the characteristic impedance of the system of measurement, and

$$L_1 = \frac{R_1 Q_0}{\omega_0} \quad and \quad C_1 = \frac{1}{R_1 \omega_0 Q_0}$$

where

$$\omega_0 = 2\pi f_0 \quad and \quad Q_0 = \frac{f_0}{\Delta f}\left(1 + \frac{2R_0}{R_1}\right)$$

For small C_0, the phase shift is equal to zero at the resonance ; but for measurements of 320 MHz SAW resonators, a negative phase shift appears. In that case, the transfer characteristic S_{21} must be studied as a function of the inter-electrode capacitor C_0

$$S_{21} = 2R_0 / \left(2R_0 + R_1(1 - \tau_0^2\omega^2) + \frac{2\tau_0(1 - L_1C_1\omega^2)}{C_1}\right.$$

$$\left. + j\left|\frac{(L_1C_1\omega^2 - 1)(1 - \tau_0^2\omega^2)}{C_1\omega} + 2(R_0 + R_1)\tau_0\omega\right|\right)$$

where $\tau_0 = R_0.C_0$.

When R_1, L_1, C_1, C_0 are known, the resonance frequency occurs when the modulus of the transfer function S_{21} is maximum. Motional components R_1, L_1, C_1 have been calculated for the SAW resonator, assumed to be perfect, and resonance curves drawn for several values of the inter-electrode capacitor C_0 (Fig. 2). These curves confirm the negative phase difference observed experimentally at the resonance frequency. A small increase in the resonance frequency, the loss and the loaded quality factor have also been observed. It is possible to explain the last two observations. First, the loss in the resonator increase because the capacity of the interdigital electrodes shunt a part of the energy. Second, these capacitors decrease the load seen by the motional circuit of the resonator and the loaded quality factor increases. These results confirm that for high frequencies the capacity C_0 of the interdigital electrodes in SAW resonators is not negligible.

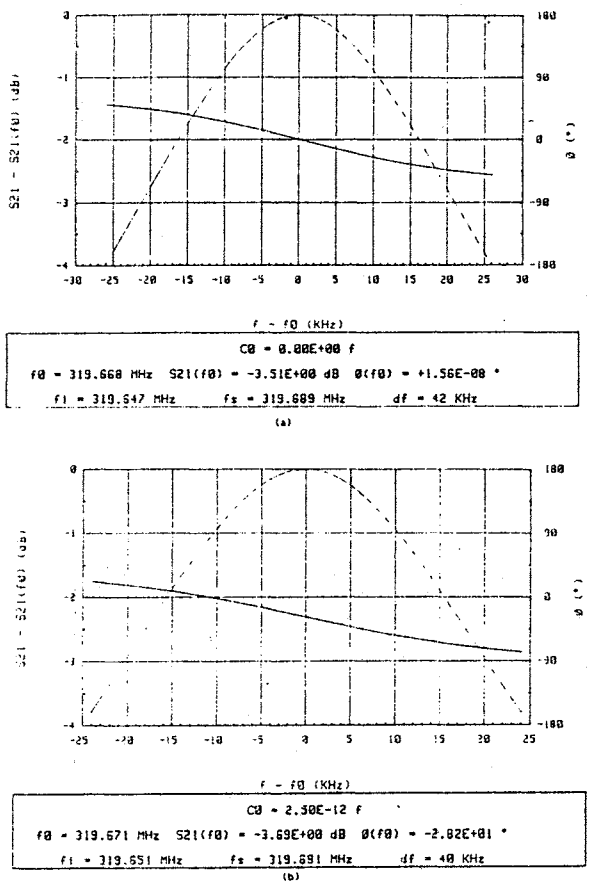


Fig. 2 : Resonance curves for a 320 MHz SAW resonator :
(a) when the resonator is perfect, that is to say when Co is negligible
(b) when Co = 2.5 pF, DC measured value for the actual resonator. The phase difference of the resonance is -28°

Realization of a SAW oscillator

An oscillator is made up of a resonant component for which the loss is B and the phase shift is Φ ; plus an integrated feedback loop including an amplifier for which the gain is G_1 and the phase shift is ϕ, to compensate the loss (Fig. 3). The oscillation conditions are :

$$|G_1| . |B| \geq 1$$
$$\phi + \Phi = 2 k \pi$$

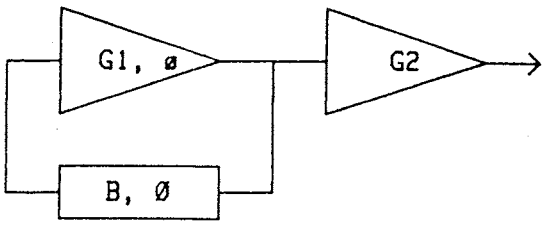


Fig. 3 : Oscillator structure :
- resonant component (B, Φ)
- sustaining amplifier (G_1, ϕ)
- decoupling amplifier (G_2)

When the sustaining amplifier has a large band, the oscillation frequency is determined by the resonant component because the amplifier does not produce a phase shift and has the same gain in a large frequency beach around the oscillation frequency (Fig. 4).

Since an oscillator is a frequency source, a part of the signal must be coupled out without disrupting the oscillation conditions. To do this, a decoupling amplifier G_2 must be placed at the oscillator output.

For high frequency oscillators, other additional precautions must be taken since all connections cause additional losses and phase shifts in the loop, and a phase converter must be placed in the circuit. To have a good output insulation, at these frequencies the decoupling amplifier must have a very small inverse transfer coefficient S_{12}. More, with a directional coupler the output decoupling is improved and it becomes easier to optimize the oscillator performance (Fig. 5).

Comparison between two amplifier technologies

Two lumped element amplifier technologies have been studied : the same amplifier has been built in a classic technology and in a SMD technology (Fig. 6). Measurements of the transfer characteristic show a 500 MHz band-pass for the one in SMD technology and a 300 MHz band-pass for the other (Fig. 7). The increase of the band-pass in the first case is probably due to partly reduction of the size of the components but mainly to the diminution of the size of the surface of the micro-strip conductors. Indeed, the micro-strip structure can be modeled by a parasitic capacitor per unit surface between the line and the ground (Fig. 8). That is to say the micro-strip structure is a low-pass filter in which the cut-off frequency decreases when the size of the connections increase. Since a SMD circuit has smaller conductors than a classic circuit, the band-width can be larger.

Power level in the oscillation loop

The power level of the signal in the loop is an important parameter because it plays an important part on the short term stability and on the aging in oscillators [7, 9]. Furthermore, the output level of an oscillator is usually fixed by the system specifications. Consequently, once the power level is known, it is possible to design the decoupling amplifier to operate in a linear range. Its gain is equal to the ratio of the output level to the loop level modulo the coupling of the directional coupler.

The power level in the loop is chiefly determined by the losses in the loop and occurs the gain compression curve of the sustaining amplifier. The oscillation occurs when the losses in the circuit are equal to the amplifier compression gain ; so that it is sufficient to estimate the losses and project them back onto the compression curve to determine the power level of the sustaining amplifier output (Fig. 9). For amplifiers with weak compression, the measured oscillation level is close to the theoretical value with errors probably due to an incorrect estimate of losses because of the impedance matching.

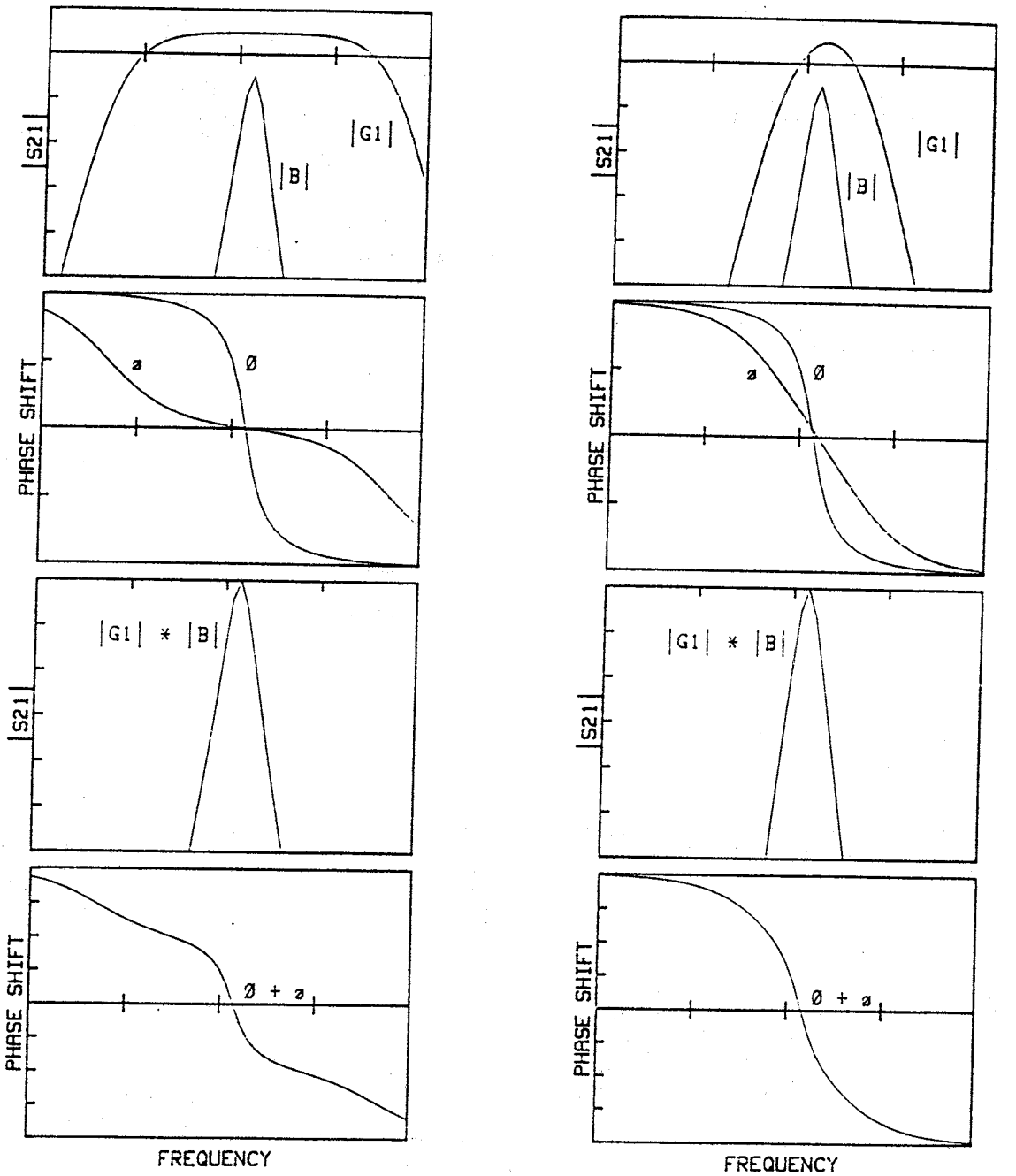


Fig. 4 : Determination of the oscillation frequency (a) for a wideband amplifier and (b) for a narrowband amplifier. In the second case, the phase slope of the amplifier is added to the phase slope of the resonator, thus the phase noise at the oscillator is higher

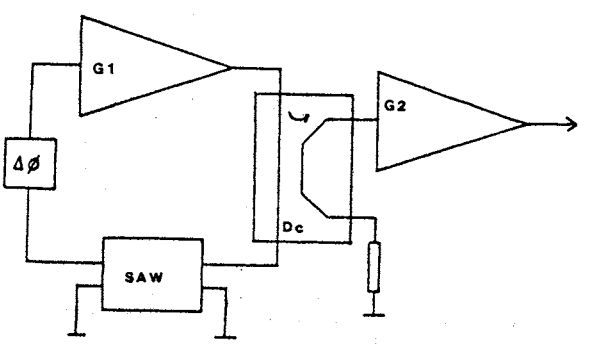


Fig. 5 : UHF SAW oscillator :
- G_1 : sustaining amplifier
- SAW : SAW resonator
- D_c : directional coupler
- $\Delta\Phi$: phase converter
- G_2 : decoupling amplifier

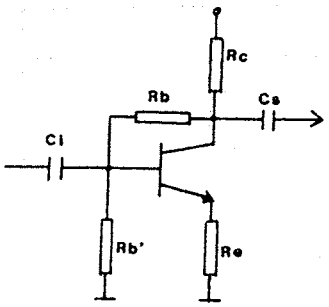


Fig. 6 : Lumped component amplifier used to compare SMD technology and classic technology

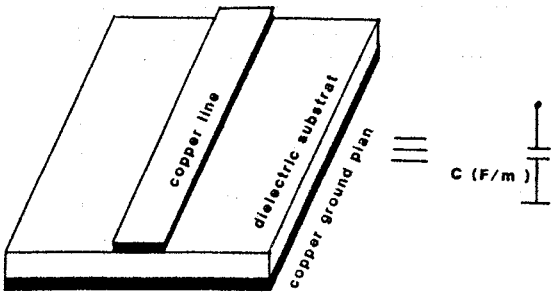


Fig. 8 : A micro-strip conductor is modeled by a capacitor per unit surface

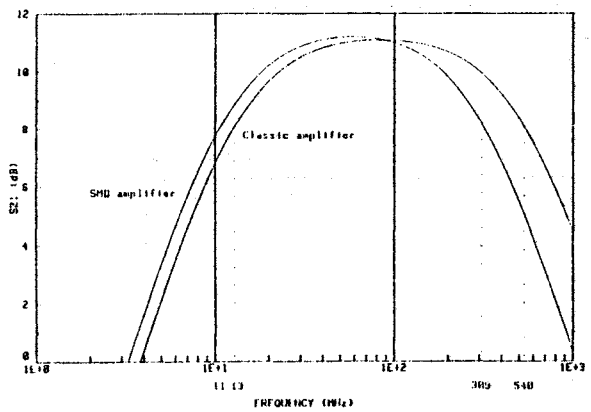


Fig. 7 : Transfer characteristic of a SMD amplifier and a classic amplifier. For the SMD technology, the gain-bandwidth product is improved by about 200 MHz

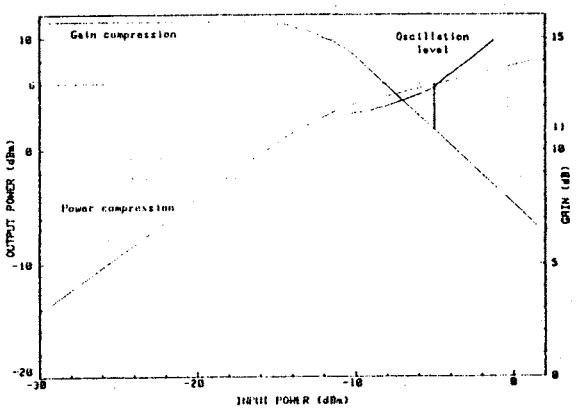


Fig. 9 : To determinate the oscillator output level, the losses in the loop must be equal to the amplifier gain compression, then the gain gives the amplifier output level. In this example, for 11 dB losses, the output level is 6 dBm. The measured output level is also represented on these curves

A first measurement of the spectral density of phase noise of a 320 MHz oscillator has shown that the flicker noise level increases with the power level in the loop (Fig. 10). This is the behaviour of an oscillator in which the sustaining amplifier is the principal phase noise source [10]. Indeed, the flicker noise has a tendency to increase with the compression gain [11].

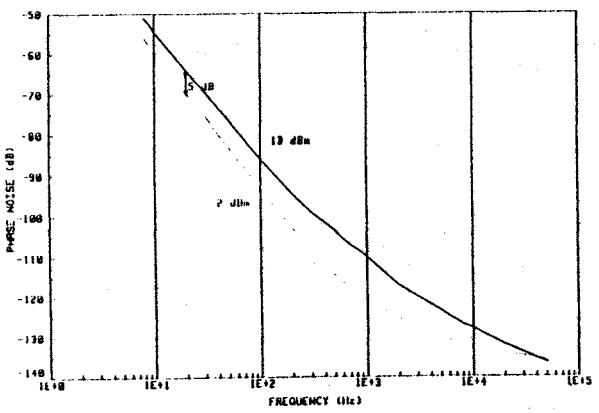


Fig. 10 : Measurement of the phase noise spectrum for a 320 MHz oscillator when the output level is 13 dBm and 2 dBm

Frequency trimming of SAW oscillators

The inter-electrode capacitor C_0 of a SAW resonator makes it difficult to use the usual trimming procedure and a phase converter is often used for the purpose [12].

The phase converter that was built is a symetrical tee filter made up by two coils of reduced impedance x_1 and by a variable capacitor of reduced impedance x_2 (Fig. 11). These impedances were chosen to have values $x_1 = 1$ and $x_2 = -1$ at the resonance frequency. When the capacitor varies between 7.7 and 2.0 pF, the phase of the transfer function can be pulled by about 90° for a 320 MHz SAW resonator (Fig. 12). But over this range, the attenuation can vary strongly.

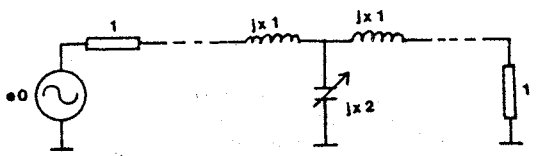


Fig. 11 : Tee phase converter used to pull the oscillation frequency

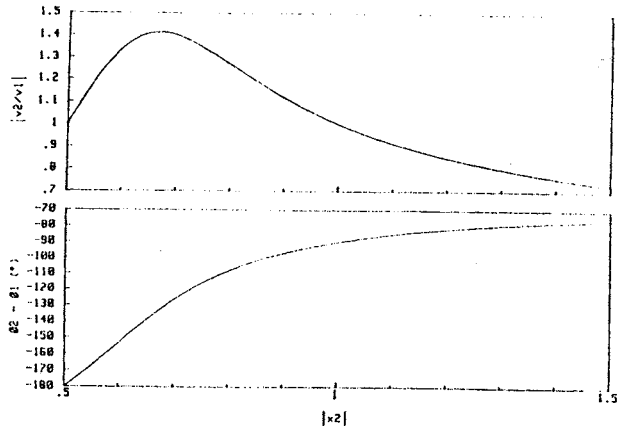


Fig. 12 : Transmittance of the phase converter for $x_1 = 1$ and x_2 varying from -.5 to 1.5. This is equivalent to a capacitor varying from 6.6 to 20 pF for a 320 MHz frequency, with a characteristic impedance of 50 Ω

The big disadvantage of this system is that on the Bode diagram, the phase curve moves towards the amplitude curve (Fig. 13). Thus the oscillator can operate off the peak of the resonance curve, with non optimal characteristics.

The solution often used for a BAW oscillator has also been studied for SAW oscillator. It consists of putting a variable capacitor C in series with the resonator (Fig. 14). Observation of the resonance frequency and the losses at resonance show that this method is not satisfactory either. To stay in a correct range of attenuation (superior to -6 dB), the frequency trimming must be limited to 8 kHz. Moreover, the phase shifts with increase of the variable capacitor, the same problem as with the phase converter.

To reduce losses in the capacitive divider C/C_0, a coil L has been placed across the interdigital electrode capacitor C_0 to provide a parallel resonance close to the oscillation frequency (Fig. 15). The quality factor of the coil must be negligible compared that of the active part R_1, L_1, C_1 of the SAW resonator. Otherwise the phase noise of the oscillator is increased. The behaviour of the various performance parameters affected by the variable capacitor illustrate the effectivenes of this solution to the trimming problem :

- larger range (15 kHz) for trimming the oscillation frequency (Fig. 16a)
- losses lower than 6 dB (Fig. 16b)
- small variations (2°) of the phase shift (Fig. 16c) and the loaded quality factor (Fig. 16d).

Conclusion

On one hand, the interdigital electrode capacitor of SAW resonator has been shown to have a significant effect for operating frequencies in the UHF range. To calculate the components of a SAW oscillator, this capacitor must be known to a good precision.

On the other hand, it is possible to trim the oscillation frequency with a variable capacitor mounted in series with the SAW resonator providing there is a coil mounted across the resonator to compensate for the inter-electrode capacitor. In the future when all-quartz package SAW resonators are to be built another way will be used to trim the resonance frequency [13, 14]. In these structures it is possible to adjust the frequency by evaporating a metal layer onto the inside of the all-quartz package, using a laser beam.

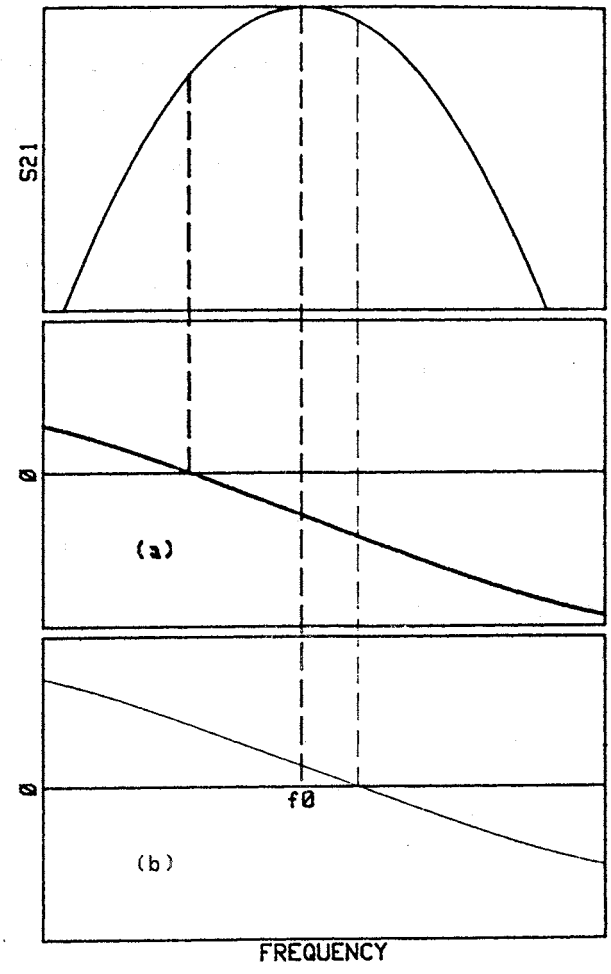


Fig. 13 : Open loop transmittance for a SAW oscillator with a phase converter. Only the value of the variable capacitor has changed between curves (a) and (b)

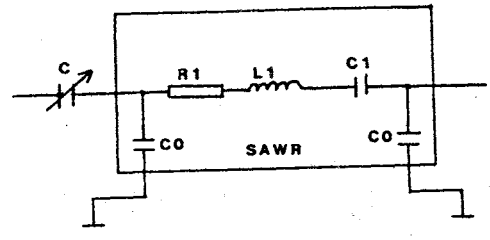


Fig. 14 : The utilization of a variable capacitor C in series with the SAW resonator allow to adjust the resonance frequency over a small range

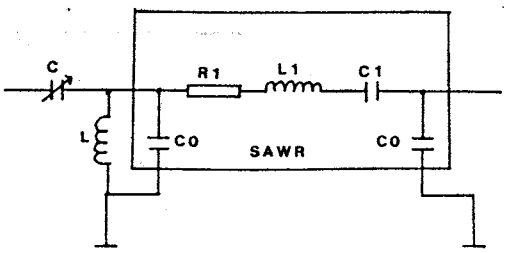


Fig. 15 : To improve the system of frequency adjustment, a coil L has been put in parallel with the SAW resonator input such that $LC\omega^2 = 1$

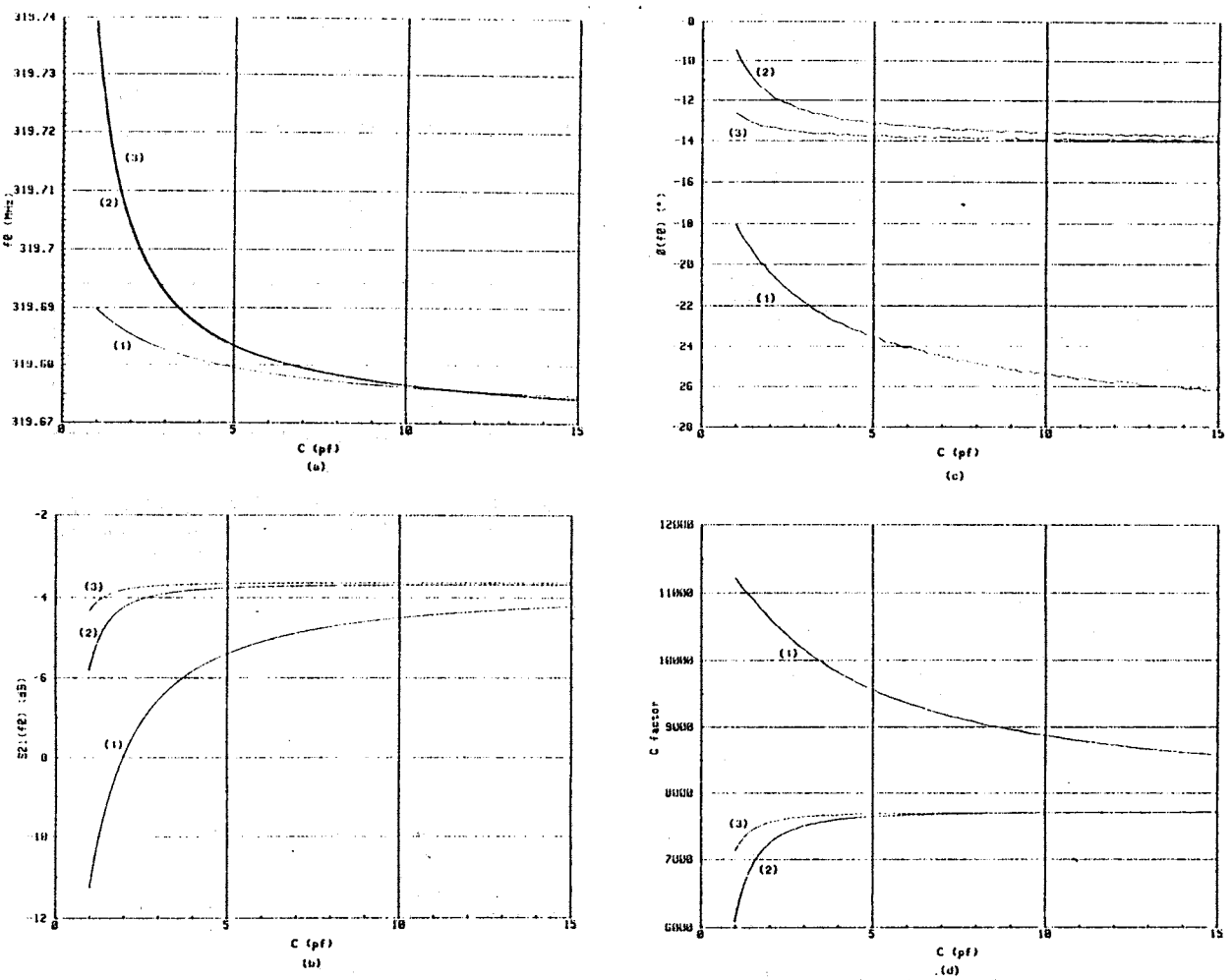


Fig. 16 : Variation of (a) the resonance frequency, (b) the resonance attenuation, (c) the phase shift, and (d) the quality factor as a function of the variable capacitor C for a 320 MIIz SAW resonator
- (1) no coil
- (2) a coil with a Q-factor of 30
- (3) a coil with a Q-factor of 100

References

[1] T.E. Parker, "Analysis of aging data on SAW oscillators", Proc. of the 34th Ann. Freq. Cont. Symp., 1980.

[2] W.R. Shreve, "Active aging of SAW resonators", Proc. of the Ultrasonics Symp., 1980.

[3] T.E. Parker, "Random an systematic contributions to long term frequency stability in SAW oscillators", Proc. of the Ultrasonics Symp., 1983.

[4] S. Calisti, P. Defranould, D. Hauden, "Influence des paramètres technologiques sur la stabilité en fréquence des résonateurs à ondes de surface", 1st E.F.T.F., Besançon, 1987.

[5] A. d'Almeida, P. Leblois, D. Hauden, S. Calisti, J.Y. Duquesnoy, L. Penavaire, "Aging in SAW oscillators", 2nd E.F.T.F., Neuchâtel, 1988.

[6] C. Gentili, "Amplificateurs et oscillateurs micro-ondes", Masson, 1984.

[7] W.R. Shreve, R.C. Bray, S. Elliot, Y.C. Chu, "Power dependence of aging in SAW resonators", Proc. of the Ultrasonics Symp., 1981.

[8] R.B. Stokes, M.S. Delaney, "Aging mechanisms in SAW oscillators", Proc. of the Ultrasonics Symp., 1983.

[9] T.E. Parker, "Random walk frequency fluctuations in SAW oscillators", Proc. of the 40th Ann. Freq. Cont. Symp., 1986.

[10] T.E. Parker, "Characteristics and sources of phase noise in stable oscillators", Proc. of the 41th Ann. Freq. Cont. Symp., 1987.

[11] M.M. Driscoll, "Low noise crystal oscillators using 50 ohms modular amplifier sustaining stages", Proc. of the 40th Ann. Freq. Cont. Symp., 1986.

[12] T.E. Parker, "Precision surface acoustic wave (SAW) oscillator", Proc. of the Ultrasonics Symp., 1982.

[13] T.E. Parker, J. Callerame, G.K. Montress, "A new all quartz package for SAW devices", Proc. of the 39th Ann. Freq. Cont. Symp., 1985.

[14] J.A. Greer, T.E. Parker, "A novel technique for trimming the frequency of a sealed surface acoustic wave resonator", Proc. of the 41th Ann. Freq. Cont. Symp., 1987.

THE DESIGN OF VHF QUARTZ CRYSTAL OSCILLATORS FOR SEVERE VIBRATION ENVIRONMENTS

M. K. Hobden
Marconi Electronic Devices Ltd.
Doddington Road - Lincoln LN6 OLF - U.K.

Modern systems specifications place increasing demands on the reference oscillator designer. The use of radar in tracked vehicles and small fast boats has highlighted the problems of maintaining a coherent output where the master reference oscillator is subjected to very high levels of rancom and pseudo-periodic mechanical excitation.

This paper addresses the problem of designing a vibration isolation system for such an environment and details the design rules to be followed to ensure success with state of the art quartz resonators. The acceleration sensitivity of existing T05 enclosure quartz crystals is analysed and the limitations of existing mounting techniques is shown. The design of a multi-pole anti-vibration mounting is presented, using modern materials to maximise inherent damping under large displacements and a non-linear model is developed to explain the advantages of the configuration used. The relative sensitivities of each section of the oscillator circuiting to vibration is quantified with particular emphasis on upconverted acceleration induced sidebands via oscillator non-linearities.

Finally, the measured results for an oscillator developed using the above criteria are presented showing that it is possible to reach levels of 150 dBc/Hz at 1 kHz from the carrier under conditions of a random excitation spectrum of 8.9 GRMS with superimposed swept sine components.

OSCILLATEURS"CHAUDS"

Emmanuel GIRARDET - Louis BIDART

QK PIEZOELECTRONIQUE

RESUME

Les gammes de température généralement exploitées tendent à s'élever vers les températures chaudes.

Si, dans la pratique l'on rencontre parfois des températures de - 55°C et le plus souvent des températures de - 40°C et de - 20°C, la tendance moderne conduit à utiliser le matériel au-delà de + 80°C et jusqu'à 95°C.

Si cela ne pose pas de gros problèmes pour la plupart des oscillateurs, en revanche la réalisation de matériels haute stabilité doit faire face à de nouvelles difficultés car il faudra thermostater à des températures pouvant atteindre 105°C.

Nous proposons certaines solutions expérimentées sur des oscillateurs de la classe 1.10^{-10}/jour à 5 MHz et 10 MHz.

1. LES OSCILLATEURS "CHAUDS" - APPLICATIONS - CONTRAINTES

Les oscillateurs de haute stabilité trouvent leur application dans les domaines de la métrologie (mesure du temps, références), les systèmes de localisation, télémesures, systèmes asservis, etc...

Dans la plupart des cas, les conditons d'environnement associées sont du type "laboratoire" : pas ou peu de contraintes mécaniques, consommation et encombrement peu critiques, plage de température d'utilisation limitée, faible gradient thermique.

La stabilité des oscillateurs utilisés est recherchée sur des intervalles de temps longs (fluctuations journalières, vieillissement mensuel, annuel) ainsi que sur des périodes plus réduites (quelques minutes à quelques dizaines de minutes). On définit alors la stabilité court terme sur un temps d'intégration donné (quelques millisecondes à quelques secondes).

Mais le besoin de sources haute stabilité se retrouve aussi dans le domaine de la métrologie embarquée (applications militaires, avioniques). Les critères d'environnement deviennent prépondérants. Aux perturbations mécaniques, acoustiques, magnétiques s'ajoute une plage de température d'utilisation souvent très étendue, de - 40°C ou - 20°C à + 95°C voire + 100°C. Cette exigence est liée à la compacité des systèmes développés où une pseudo régulation thermique n'est pas envisageable (gradients thermiques trop rapides, encombrement disponible réduit, énergie mise en oeuvre trop importante).

Ces exigences fonctionnelles induisent de nouveaux critères liés à la stabilité, en particulier le temps de mise en oeuvre, la vitesse de stabilisation, la retrace.

A partir d'oscillateurs haute stabilité standards actuels (5 et 10 MHz), l'étude a porté sur leur adaptation au fonctionnement en environnement thermique sévère(température d'utilisation jusqu'à 100°C), et sur la mesure comparative de leurs performances. Cette comparaison est particulièrement significative. En effet, de par leur constitution, il a été possible de caractériser le même lot d'oscillateurs en configuration "standard" et "haute température".

Les extensions aux autres conditions d'environnement (vibrations mécaniques, acoustiques) font l'objet d'essais et de développement parallèles en cours.

2. CONCEPTION TECHNIQUE

Support de l'étude, l'oscillateur haute stabilité "MOT" est conçu pour des gammes de température d'utilisation de - 20 + 60°C ou - 40 + 70°C. Oscillateur et quartz sont thermostatés à une température de 70 à 85°C.

Sans modification de conception, le MOT a été adapté à une utilisation en gamme de température étendue (-40 + 95 et + 105°C).

Le choix du résonateur va permettre, à règlages électriques de l'oscillateur rigoureusement identiques, de caractériser en stabilité un même lot de MOT en configuration "standard", puis "haute température" et inversement.

2.1. RESONATEUR

a) Exigences

Conçues pour équiper des oscillateurs dans la gamme de fréquence 5 - 10 MHz, les performances de stabilité recherchées imposent de prendre en compte les critères suivants :

- coéfficient de surtension le plus élevé possible (stabilité court terme, vieillissement)
- dérive de fréquence autour de la température de régulation du thermostat minimale (stabilité dans la gamme de température)
- choix d'un boitier compatible avec les performances de stabilité à long terme.

REMARQUE : Le thermostat est "calé" à un point d'inversion de la caractéristique Fréquence-température du quartz. Ce point doit être situé au minimum 10°C au-dessus de la température haute de la gamme d'utilisation pour assurer une régulation thermique correcte.

b) Solutions

* Coefficient de surtension

Les résonateurs de coupe AT (simple rotation) et SC (double rotation) présentent les coefficients de surtension suivants :

	5 MHz	10 MHz
AT Partiel 3	2,3 - 2,5 10E6	1,2 10E6
SC Partiel 3	2,1 - 2,4 10E6	1,2 10E6

* Courbes Fréquence-température

. Coupe AT

La caractéristique F(T) est une courbe du 3ème degré présentant un point d'inflexion situé à 25°C. Le point d'inversion "haut" des courbes F(T) peut évoluer en fonction de l'angle de coupe, de 50°C à 100°C (limite).

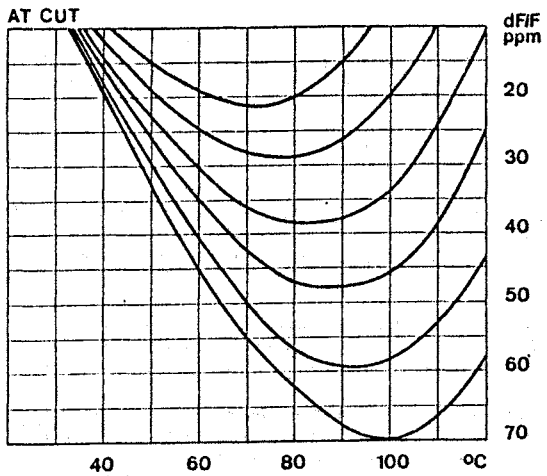


Fig. 1

. Coupe SC

Cette coupe est caractérisée par deux modes de vibration:

Mode b particulièrement sensible à la température (variation relative de fréquence de l'ordre de 30.10E-6/°C)

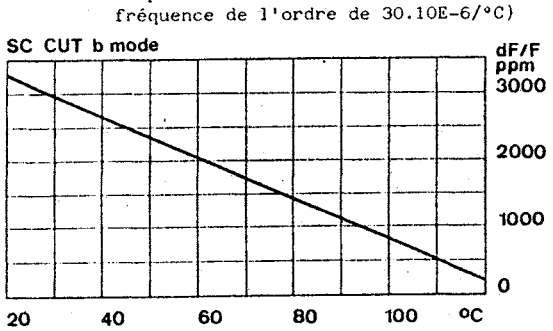


Fig. 2

Mode c courbes F(T) du 3ème ordre, dont le point d'inflexion est situé aux environ de 90°C.

Pour des gammes de température standards, on réalisera une régulation thermique au point d'inversion "froid", compris entre 60°C et 80°C suivant les paramètres de coupe .

Cette même famille de courbe présente un point d'inversion "chaud", situé entre 90°C et 115°C susceptible d'être exploité pour la réalisation d'oscillateurs "haute température".

REMARQUE : On notera la faible variation relative de Fréquence entre les points d'inversion "haut" et "bas" d'une même courbe caractéristique SC. (quelques 10-7 à quelques 10-6)

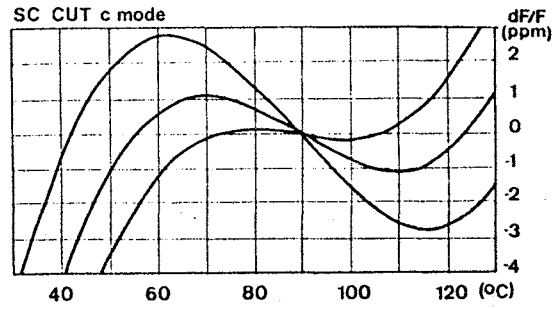


Fig. 3

. Caractéristique des points d'inversion

La figure 4 présente les courbes Fréquence-température comparées aux points d'inversion.

A variation de température identique autour des points d'inversion, la fréquence évolue environ cinq fois moins rapidement dans le cas d'une coupe SC (P.I. hauts et bas) que sur une coupe AT (P.I. hauts).

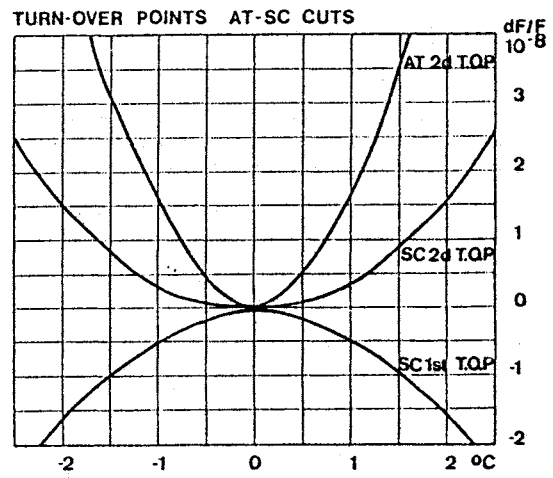


Fig. 4

* Autres paramètres

. Anisochronisme

Les caractéristiques électriques des résonateurs se dégradent lorsque la puissance d'excitation augmente. Le seuil est situé très sensiblement plus haut dans le cas de coupe SC.

. Sensibilité accéléromètrique

Facteur important dans le cas de matériel soumis à des contraintes de vibrations mécaniques. Là encore, les meilleurs résultats sont obtenus sur des résonateurs de coupe SC.

c) Choix

Compte tenu de l'ensemble de ces performances, les résonateurs SC 5 et 10 MHz sont utilisés dans le développement des oscillateurs haute température.

Un soin tout particulier est apporté à la fabrication des résonateurs et à la fermeture des

boitiers verre (type HC 27/U) (cycles thermiques de stabilisation), de manière à garantir les performances optimales de stabilité à long terme.

2.2. OSCILLATEUR

L'oscillateur développé autour des quartz 5 et 10 MHz (SC partiel 3) est du type Clapp (Fig. 5). Le point de fonctionnement de l'oscillateur est déterminé de manière à affecter le moins possible le facteur de qualité du résonateur (coefficient de surtension en charge).

La puissance d'excitation du quartz peut être règlée de quelques dizaines de μW à plusieurs centaines de μW.

Comme indiqué précédemment (§ 2.1., Fig.2) l'écart relatif de fréquence entre les deux points d'inversion est faible. Ainsi, à règlage oscillateur identiques, il est possible de choisir la température de fonctionnement de l'oscillateur à l'un de ces deux points.

Oscillateur et amplificateur de sortie sont réalisés en technologie hybride (encombrement, rendement de la boucle de chauffage). Seuls les éléments série Quartz (règlage) et le sous ensemble régulation de tension font appel à des composants discrets.

Les gammes de température d'utilisation des éléments constitutifs, les température de jonction des composants actifs (régulateur de tension) permettent l'emploi de cet oscillateur à des températures de fonctionnement supérieures à celles de la gamme visée.

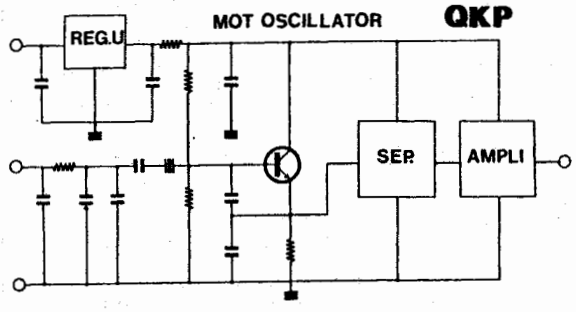


Fig. 5

2.3. THERMOSTAT

a) Description

Afin d'atteindre les caractéristiques optimales de stabilité de fréquence en fonction de la température, le résonateur, mais aussi l'oscillateur et l'étage régulateur de tension seront thermostatés à température fixe (point d'inversion du quartz).

Ces éléments sont montés dans un boitier métallique d'un volume d'environ 15 cm3.

Un transistor de puissance assure le chauffage du boitier par l'intermédiaire d'un thermostat de type proportionnel (Fig. 6 et 7).

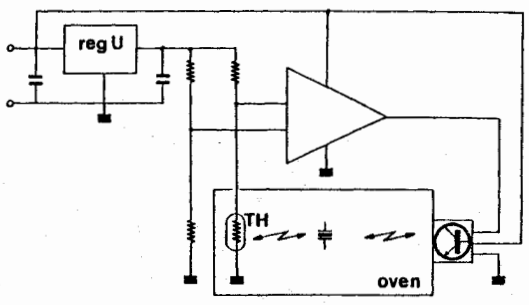


Fig. 6

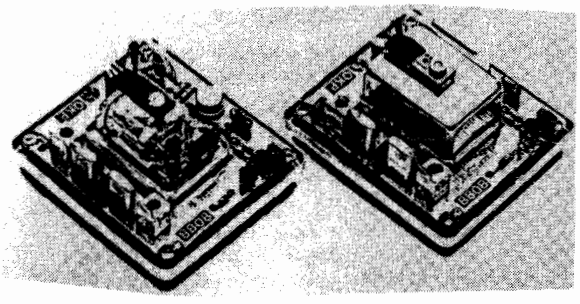


Fig. 7

De la même manière que pour l'oscillateur, les éléments contitutifs sont choisis dans des listes compatibles avec l'utilisation à haute température.

Le transistor de puissance est défini de manière à ce que la température de jonction en fonctionnement extrême (démarrage à basse température) reste en deçà des limites admissibles (calcul de sécurité).

Ce thermostat de structure très simple est utilisé sur tous les oscillateurs de type MOT. Il présente un rapport mise en oeuvre/performances intéressant.

b) Caractérisation

* Mesures thermographiques

Le dispositif est constitué d'une caméra infra-rouge et d'un système d'analyse et traitement numérique de l'image. Sa mise en oeuvre a permis deux types d'analyse :

- vérification à postériori sur les éléments actifs, transistor de puissance, des températures de jonction maximales atteintes dans les conditions d'utilisation extrêmes (démarrage - régimes stabilisés)

- topographie du gradient de température à l'extérieur de l'enceinte chauffée, en régime établi. Cet essai est effectué pour diverses orientations géomètriques de l'oscillateur, et pour différentes températures de consigne. Malgré la totale asymètrie de chauffage, la répartition de température extérieure atteint une homogénéité satisfaisante.

Les documents exploités sont du type de celui présenté figure 8.

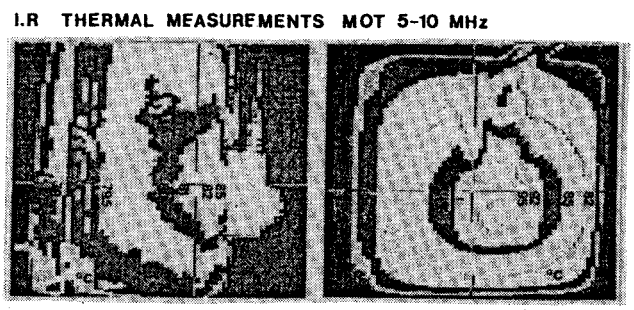


Fig. 8

* Régulation thermique au niveau du quartz

La connaissance du gradient thermique à l'extérieur du boitier ne suffit pas.

La stabilité en température de l'oscillateur est essentiellement fonction de la variation de température résiduelle au coeur de la lame de quartz. Il n'est, bien sûr, pas possible de placer un capteur à ce niveau.

Par contre, les résonateurs de coupe SC présentent la caractéristique de posséder un mode de vibration très sensible à la température (§ 2.1., Fig. 2).

L'oscillateur est règlé de manière à ce que le résonateur fonctionne sur ce mode.

Le thermostat étant opérationnel, la mesure de la fréquence de sortie de l'oscillateur en fonction de la température extérieure permet de caractériser la qualité de la régulation thermique au point le plus sensible, c'est à dire au niveau du résonateur lui-même.

Les mesures sont effectuées pour différentes orientations géomètriques de l'oscillateur dans la totalité des gammes de température d'utilisation.

Température extérieure	Température de consigne	Variation au niveau du quartz
- 20 + 70°C	85 ° C	$\leq$ 1°C
- 40 +110°C	115 ° C	$\leq$ 1,2 °C

Le coefficient de qualité de la régulation thermique est de l'ordre de 100.

A partir du même principe de base de thermostat, il est possible d'atteindre, pour des produits répondant à des cahiers des charges différents, un coefficient de qualité de l'ordre de 300.

* Temps de stabilisation thermique

La mesure de ce paramètre est effectuée en utilisant le même principe (mode b). On compare alors la durée d'établissement de l'équilibre thermiquedans le cas d'une configuration "standard" (température du four 85°C) et " haute température" (115°C).

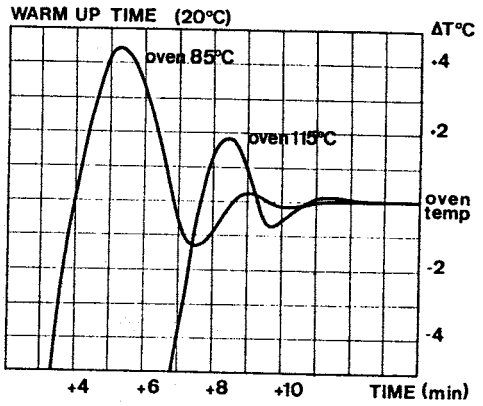


Fig. 9

3. MESURES - COMPARATIF

3.1. ECHANTILLONNAGE

Les essais sont réalisés sur un lot de plus de 30 oscillateurs décomposé comme suit :

10 MHz 10 oscillateurs
5 MHz 20 oscillateurs

Certaines pièces sont caractérisées successivement pour deux conditions de fonctionnement (voir § 2.1, 2.2.)

Type de fonctionnement	Désignation	Température du four
Standard	B.T.	65°C $<\Theta<$ 75°C
Haute température	H.T.	100°C $<\Theta<$ 110°C

Le chronogramme d'essais suivant est établi :

Fréquence	n°lot	Qté	Type de fonctionnement
10 MHz	1	5p	HT ↘ BT
	2	5p	BT ↗ HT
			t ►
5 MHz	3	5p	HT ↘ BT
	4	5p	BT ↗ HT
	5	10p	← HT →

3.2. MESURES

Certaines caractéristiques électriques sont indépendantes du mode de fonctionnement de l'oscillateur (HT ou BT) ; c'est le cas des caractéristiques de sensibilité à la charge ou à la tension d'alimentation. Bruit de phase et stabilité court terme ($t \leq 3$ s) sont également constants.

Les valeurs typiques sont présentées figure 10 (stabilité court terme) et figure 11 (bruit de phase.

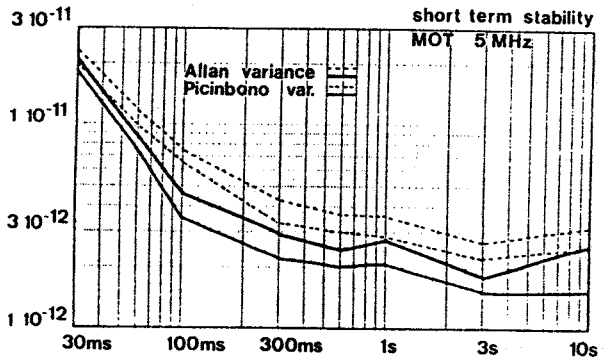


Fig. 10

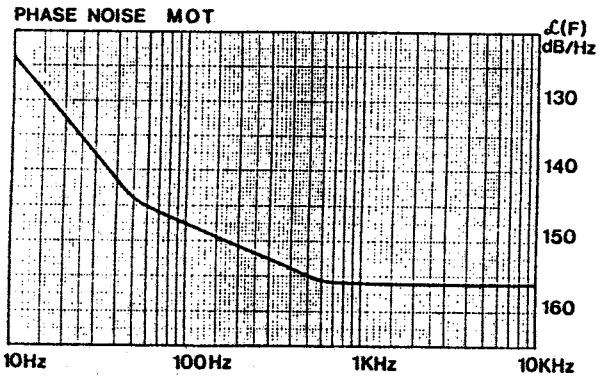


Fig. 11

L'essentiel des essais a porté sur les mesures susceptibles d'évoluer en fonction des régimes de fonctionnement.

Il s'agit en particulier des mesures de :

- temps de démarrage - vitesse de stabilisation
- vieillissement long terme
- retrace

3.2.1. Démarrage (Fig. 12)

A température de fonctionnement constante, temps de démarrage et vitesse de stabilisation dépendent essentiellement du couplage thermique Quartz/Source de chauffage et du type de thermostat (§ 2.3) Les caractéristiques mesurées répondent à la majorité des conditions d'utilisation.

Si nécessaire, on pourra faire appel à des thermostats à chauffage rapide (§ 4 - Extensions).

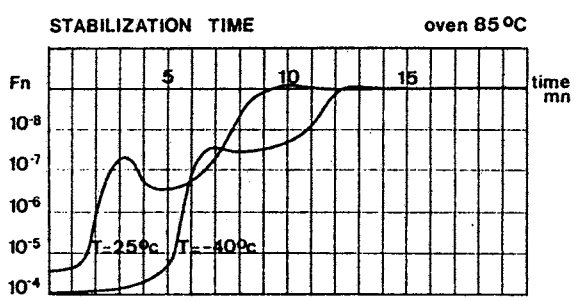


Fig. 12-1

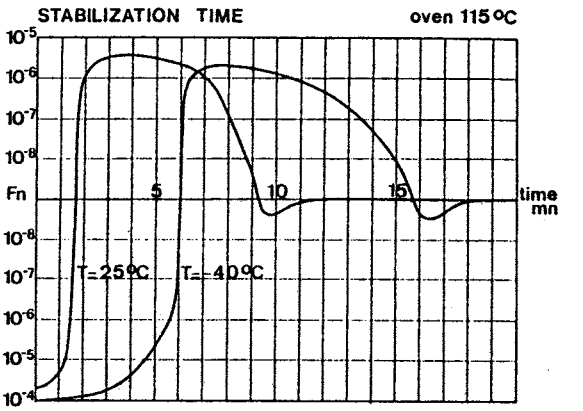


Fig. 12-2

3.2.2. Stabilité long terme

Les courbes typiques de vieillissement des oscillateurs des lots(1),(2) et (5) sont présentées figures 13, 14, 15. Les stabilités sont calculées à partir de pentes moyennes de vieillissement sur les derniers jours.

Dans tous les cas, les résultats mesurés sont obtenus après des durées de fonctionnement relativement courtes (quelques semaines).

Une première comparaison relative entre oscillateurs de lots différents, après une période identique de fonctionnement, est possible.

Les performances absolues sont données à titre indicatif. Naturellement,ces chiffres s'améliorent après une période de stabilisation supérieure (mesures en cours).

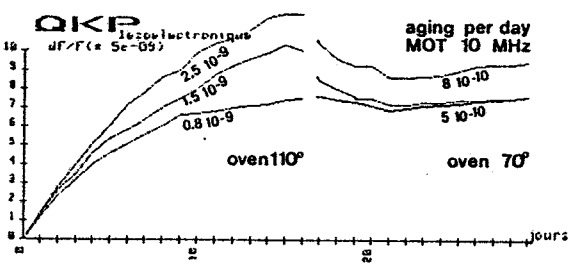


Fig. 13

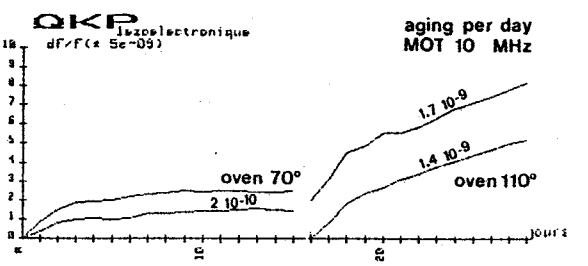


Fig. 14

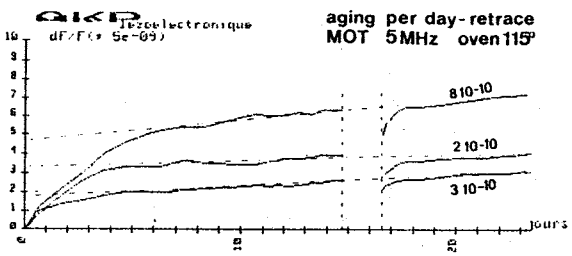


Fig. 15

La synthèse des résultats actuels est indiquée dans le tableau suivant :

N° Lot	Fréquence	Régime de Fonctionnement	Antériorité	Stabilité par jour
①	10 MHz	HT	néant	1-2. 10^{-9} / jour
①	10 MHz	BT	3s. HT	5. 10^{-10}/ jour
②	10 MHz	BT	néant	2. 10^{-10}/ jour
②	10 MHz	HT	3s BT	1-2. 10^{-9} / jour
③⑤	5 MHz	HT	néant	3-5. 10^{-10}/ jour
④	5 MHz	BT	néant	1.10^{-10}/ jour

INTERPRETATION

En régime basse température les oscillateurs MOT 5 et 10 MHz atteignent très rapidement et dans une grande proportion une stabilité de quelques 10E-10/jour puis 1.10E-10/jour après un mois de fonctionnement continu.

En régime haute température, les performances de stabilité diffèrent en fonction de la fréquence.

A la fréquence de 10 MHz, un vieillissement moyen de 1,5.10E-9/jour est obtenu après une quinzaine de jours de fonctionnement.

A durée identique, les oscillateurs 5 MHz présentent une stabilité moyenne de 5.10E-10/jour. Ces deux chiffres s'améliorent sensiblement sur des durées d'observation plus importantes.

3.2.3. Retrace

Les caractéristiques de retrace prennent toute leur importance dans le cas d'utilisations interrompues de références haute stabilité.

Après une période de fonctionnement suffisamment longue pour permettre de connaître la pente de vieillissement par jour de l'oscillateur, le fonctionnement de celui-ci est interrompu pendant une durée de 48 Heures.

Pendant cette période, l'oscillateur peut être stocké dans deux types de conditions :

- stockage à température ambiante (25°C)
- cyclage thermique (8 cycles -40 + 85°C)

La reprise en fréquence après coupure (ou retrace) est comparée à la fréquence initiale de l'oscillateur avant coupure (vieillissement pris en compte).

Les oscillateurs du lot 5 (10p) ont été mesurés en retrace (Fig. 15). L'influence du type de stockage n'induit pas de différence sensible au niveau des résultats.

	Mini	Moy.	Maxi
To + 15 mn	3.10^{-9}	5.10^{-9}	1.10^{-8}
To + 60 mn	1.10^{-9}	3.10^{-9}	5.10^{-9}
To + 12 h	1.10^{-9}	$1,5.10^{-9}$	3.10^{-9}

4. SYNTHESE - EXTENSIONS

Les essais réalisés autour des oscillateurs de type MOT permettent de caractériser de manière significative l'évolution des performances dans le cas d'une extension vers l'utilisation haute température.

L'élévation de la gamme de température d'utilisation ne présentent pas de difficultés technologiques particulières et ne remet pas en cause la fiabilité du matériel.

Dans ces conditions, les meilleures performances de stabilité ont été mesurées sur des oscillateurs 5 MHz; les résultats à 10 MHz restent d'un bon niveau.

Le mesures électriques sont réalisées sur des lots conséquents. Les résultats présentent une faible dispersion à l'intérieur de chaque lot et une très bonne reproductibilité. Les essais se poursuivent actuellement, en particulier en ce qui concerne la stabilité à long terme.

Ces essais sont essentiellement liés à l'évolution des performances face aux contraintes d'environnement thermique. En ce qui concerne d'autres aspects d'environnement durci liés à une utilisation embarquée (vibrations mécaniques, acoustiques, électromagnétiques), plusieurs études sont menées actuellement à QKP. Elles conduiront en particulier au développement d'une famille d'oscillateurs haute stabilité embarqués (MOT.EM)(Figure 16).

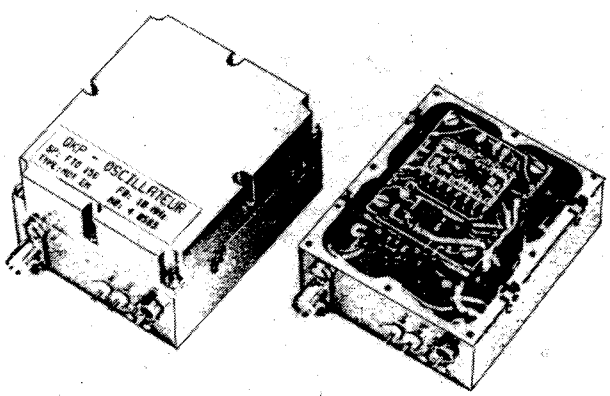


Fig. 16

INFLUENCE DES SENSIBILITES ACCELEROMETRIQUES ET ACOUSTIQUES SUR LE BRUIT DE PHASE DES OSCILLATEURS PIEZOELECTRONIQUES

MM. Patrick RENOULT - Emmanuel GIRARDET - Louis BIDART

QK PIEZOELECTRONIQUE

L'Electronique moderne doit faire face à des contraintes de plus en plus sévères générées par un environnement hostile d'une grande complexité. Le matériel doit conserver toutes ses caractéristiques, à tout moment, en dépit des fluctuations climatiques ou mécaniques.

Parmi les différentes conditions d'exploitation, le fonctionnement en présence de vibrations et de bruit phonique retiendra notre attention car les précautions à prendre interviennent dès la conception du produit.

Dans les années écoulées, il a été largement débattu de l'importance de la sensibilité accélérométrique des résonateurs piézoélectroniques et des mesures à prendre pour améliorer ce paramètre, mais il faut désormais tenir compte conjointement de la sensibilité acoustique d'une source de fréquence soumise à un bruit phonique.

Il est évident qu'un matériel électronique équipant une fusée ou un avion de combat doit supporter en même temps des vibrations et un bruit acoustique aléatoires qui influent sur le bruit de phase d'un oscillateur piézoélectrique.

Le développement des oscillateurs piézoélectriques est axé sur l'obtention d'un bruit de phase minimum à la fois près et loin de la fréquence porteuse et à un stade industriel de 1 à 200 MHz.

Soumis à des vibrations, un oscillateur verra son bruit de phase se dégrader près de la porteuse tant que s'exercera la sensibilité accélérométrique du résonateur, cependant que sous l'action conjointe ou séparée d'un bruit phonique, le palier de bruit sera perturbé.

Si donc l'on veut conserver une qualité spectrale obtenue avec beaucoup de mal, il faut trouver des solutions.

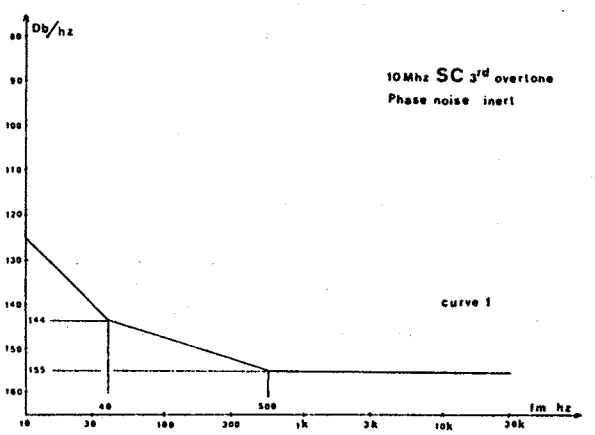
SENSIBILITE ACCELEROMETRIQUE ET BRUIT DE PHASE

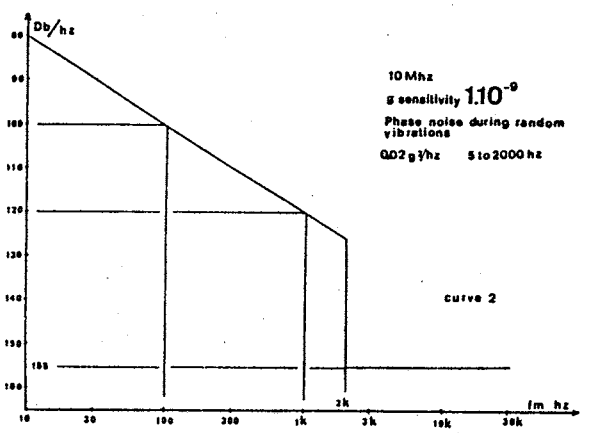
Pour conserver un bruit de phase durant les vibrations il faut soit réduire la sensibilité accélérométrique du résonateur employé, soit la neutraliser. Prenons un spectre de vibration de 0,02 g^2/Hz de 0 à 2000 Hz. Ceci est une valeur expérimentale. En réalité, un spectre réel est beaucoup plus chaotique.

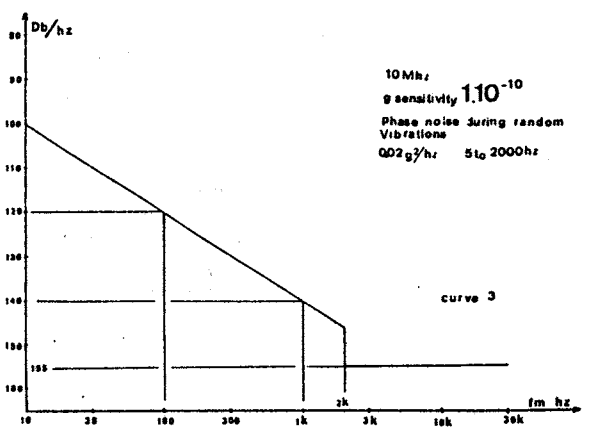
Soit, par ailleurs, un oscillateur ayant un bruit de phase à 10 MHz suivant la courbe n°1.

Si le résonateur de cet oscillateur possède une sensibilité accélérométrique de 1.10^{-9}/g il induira un bruit de phase durant les vibrations selon la courbe n°2.

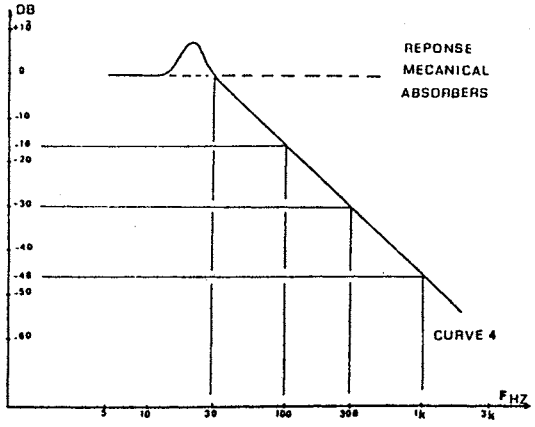
De même pour 1.10^{-10}/g, on aura en vibrations la courbe n°3.



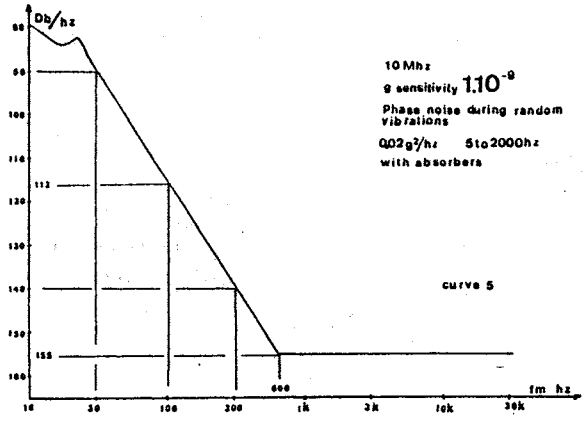




On peut équiper le système d'un ensemble d'amortisseurs agissant comme un filtre mécanique ayant une courbe de réponse suivant la courbe n°4.

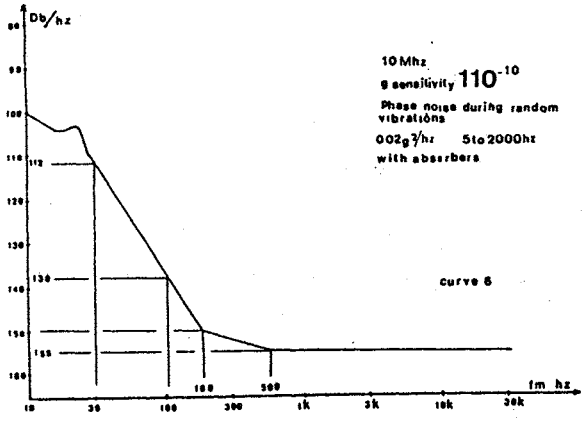


Dans ces conditions, un oscillateur à quartz comprenant un résonateur avec une sensibilité accélérométrique de 1.10^{-9}/g et des amortisseurs mécaniques aura un bruit de phase durant les vibrations suivant la courbe n°5.

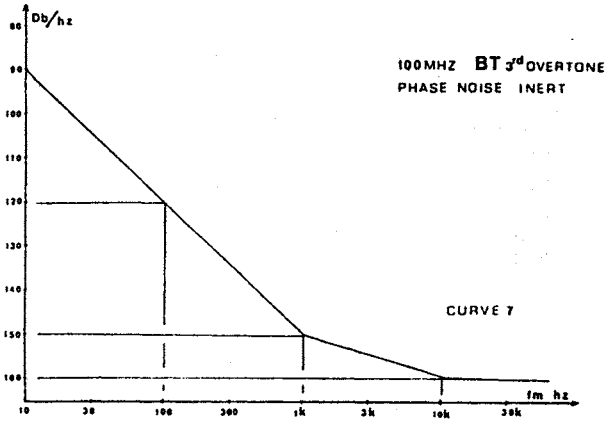


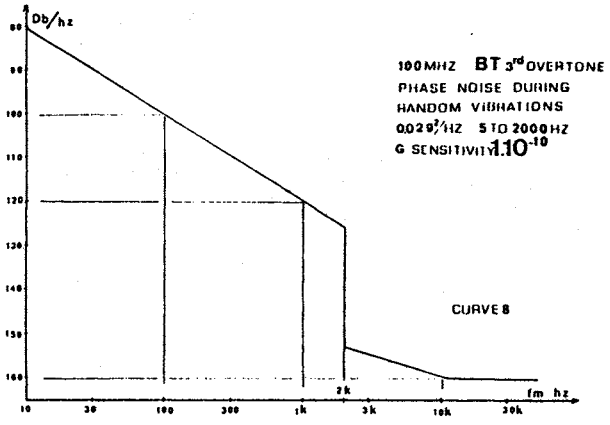
L'examen de cette courbe montre qu'entre 1 et 30 Hz, le bruit de phase dépend uniquement de la sensibilité accélérométrique. Après 30 Hz, le bruit de phase est la résultante des effets conjugués de la sensibilité accélérométrique et du filtrage mécanique. Ce n'est qu'à 600 Hz de la porteuse que le bruit de phase en inerte et en vibration sera équivalent.

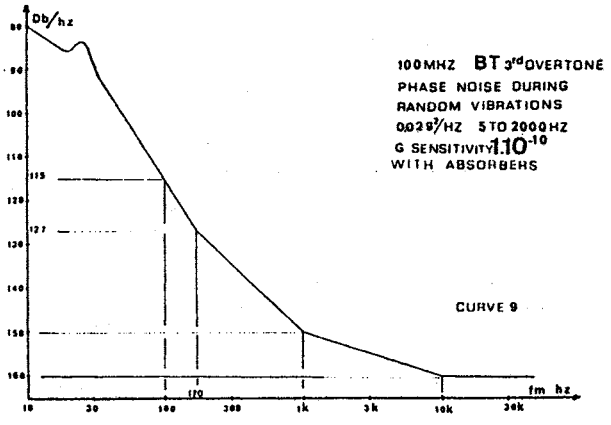
Le même raisonnement appliqué à une sensibilité accélérométrique de 1.10^{-10}/g donnera la courbe n°6.



Si maintenant l'on considère le bruit de phase en inerte suivant la courbe n°7 pour un oscillateur à quartz à 100 MHz, on aura en vibration une altération due à une sensibilité accélérométrique figurée en courbe n°8. L'ensemble étant suspendu par des amortisseurs ayant toujours la réponse de la courbe n°4, on obtiendra le bruit de phase de la courbe n°9 en vibrations. On aura encore en dessous de 30 Hz une portion de courbe dépendant uniquement de la sensibilité accélérométrique, au delà une partie transitoire figurant l'action conjuguée des amortisseurs et à partir de Fo + 170 Hz une équivalence entre le bruit de phase en vibrations et en inerte.



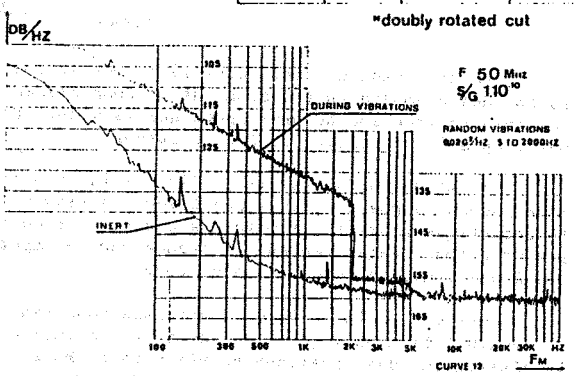


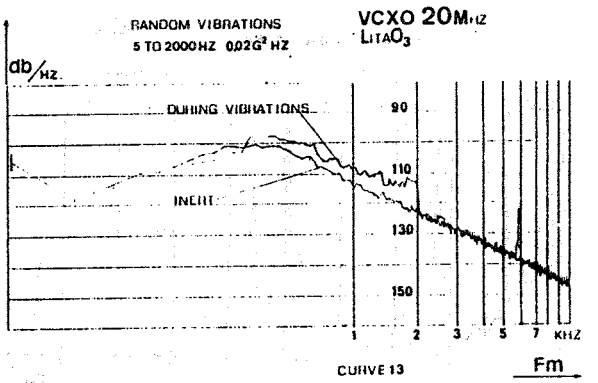


g sensitivity (dF/F / g) — table 10

			5 MHz	10 MHz	20 MHz	50 MHz	100 MHz
QUARTZ	AT cut	P 1	$3\,10^{-9}$	$3\,10^{-9}$	$1\,10^{-9}$	$1\,10^{-8}$	$1\,10^{-8}$
		P 3			$5\,10^{-10}$	$5\,10^{-10}$	$3\,10^{-10}$
		BVA	$5\,10^{-11}$				
		P 5	$1\,10^{-9}$	$1\,10^{-9}$		$5\,10^{-10}$	$3\,10^{-10}$
	SC cut	P 3	$7\,10^{-10}$	$1\,10^{-9}$	$1\,10^{-10}$		
		BVA		$5\,10^{-11}$			
		QAS		$1\,10^{-10}$			
	BT cut	P 3		$1\,10^{-10}$	$1\,10^{-10}$	$1\,10^{-10}$	$1\,10^{-10}$
		P 5				$3\,10^{-10}$	$1\,10^{-10}$
$LiTaO_3$		P 1			$2\,10^{-9}$	$1\,10^{-8}$	$5\,10^{-8}$
		P 3 *					$5\,10^{-9}$

*doubly rotated cut





Ainsi on peut voir qu'un filtrage mécanique aura atténué les effets néfastes de la sensibilité accéléromètrique des résonateurs. L'éfficacité du procédé est d'autant plus grande si l'on considère que les normes de vibrations tendent à s'étaler non plus seulement de 0 à 2000 Hz mais de 0 jusqu'à 4000Hz dans un avenir proche.

Chaque fois que l'utilisateur du matériel n'aura d'intérêt que pour un bruit de phase relativement éloigné de la porteuse (à partir de plusieurs centaines de Hz) cas de l'avionique, il sera interessant de limiter les efforts de sélection sur la sensibilité accéléromètrique (1.10^{-9} plutôt que 1.10^{-10}) et de se tourner vers l'emploi d'amortisseurs appropriés.

A l'inverse, près de la porteuse la tenue en vibration dépendra toujours de la qualité de la sensibilité accéléromètrique.

Le tableau 10, sans être exhaustif, donne un ensemble des valeurs vérifiées pour la sensibilité accéléromètrique à différentes fréquences pour des coupes et des matériaux divers.

On notera la dégradation des performances pour les résonateurs usinés ioniquement tant en quartz qu'en Tantalate de lithium.

Les mesures de sensibilité accéléromètrique ont été réalisées sur des résonateurs d'après le synoptique du tableau n°11 figurant en annexe. La figure n°12 montre le comportement d'un oscillateur thermostaté à 50 MHz avec la détérioration du bruit de phase en vibration.

La figure n°13 représente le comportement en vibration d'un oscillateur commandé en fréquence utilisant un résonateur en tantalate de Lithium et sans amortisseur. L'effet de la sensibilité accéléromètrique est clairement exprimé. On peut voir aussi que dans ce cas le bruit de phase est conservé après 2000 Hz, ce qui laisse présager que le Tantalate de Lithium reste insensible au bruit phonique.

SENSIBILITE ACOUSTIQUE ET BRUIT DE PHASE

On peut constater que les sensibilités accéléromètriques et acoustiques apparaissent à leur maximum pour des sollicitations suivant des axes différents. Il s'ensuit que les moyens mis en oeuvre pour limiter leurs effets dans une configuration donnée s'exercent souvent en sens contraire.

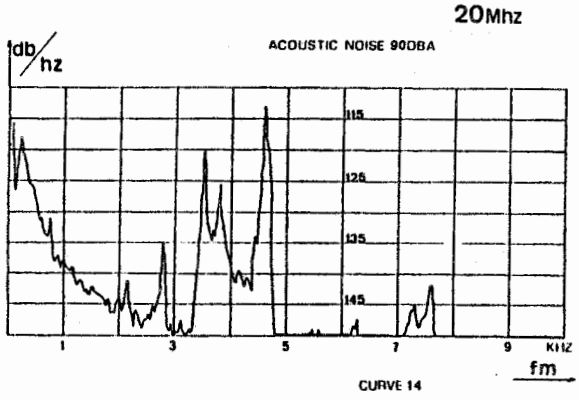
Une protection extérieure contre le bruit phonique étant en principe malaisée, on cherchera en priorité l'amélioration de la microphonie du résonateur.

Divers paramètres influent sur le comportement d'un résonateur soumis à un bruit acoustique intense, entre autres :

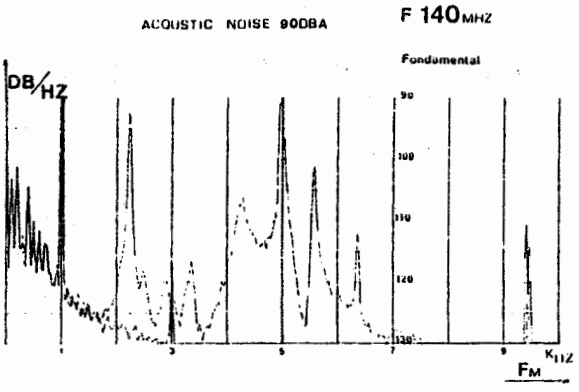
- le matériau piézoélectrique
- la coupe
- l'état de surface et la métallisation
- le nombre et l'importance des fixations.

Le quartz présente une sensibilité acoustique croissante avec la fréquence comme on peut le voir sur les figures n°14 et n° 15.

En première approximation, on peut affirmer que tout résonateur à quartz ayant une épaisseur inférieure à 50 microns démontrera une grande sensibilité acoustique ce qui laisserait penser que les résonateurs usinés ioniquement ou chimiquement seraient à proscrire.

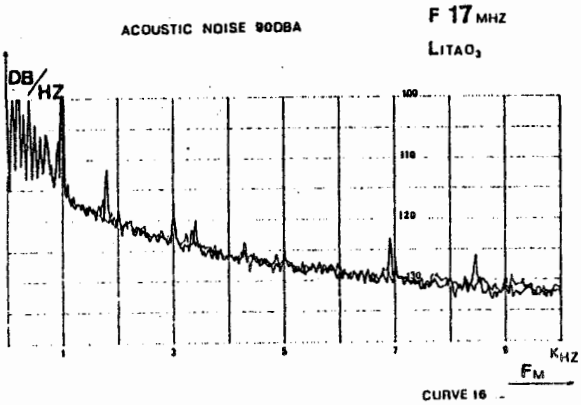


CURVE 14

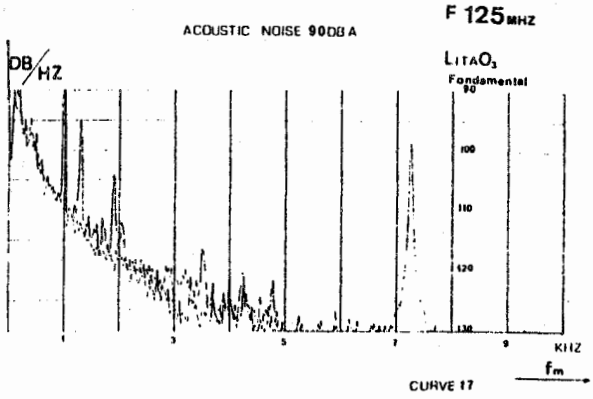


CURVE 15

A l'inverse et paradoxalement, le Tantalate de lithium démontre une parfaite insensibilité au bruit phonique : voir courbes n°16 et n°17. On en déduira immédiatement que dans le cas du Tantalate de Lithium seule l'amélioration ou la neutralisation de la sensisibilité accéléromètrique reste une préoccupation.



CURVE 16

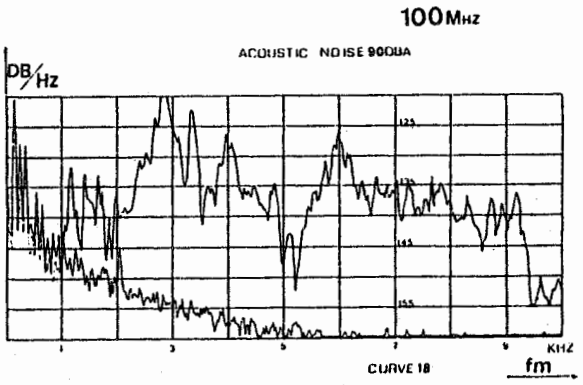


CURVE 17

Les différentes coupes employées pour la réalisation des résonateurs à quartz ont un comportement variable en fonction du bruit phonique mais cela doit beaucoup à la conception de l'oscillateur.

L'état de surface et la qualité de la métallisation peuvent augmenter considérablement la sensibilité acoustique comme en témoigne la courbe n°18 où la perturbation enregistrée est en grande partie imputable à une mauvaise tenue des électrodes.

Enfin, la méthode la plus efficace pour réduire la sensibilité acoustique tient dans la conception et la réalisation des fixations du résonateur. Mais comme la sensibilité accélérométrique dépend elle aussi largement de ces mêmes fixations, il faudra arbitrer entre deux solutions pouvant donner des résultats contradictoires.

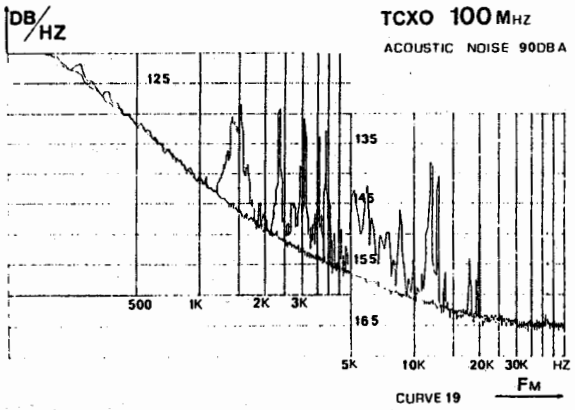


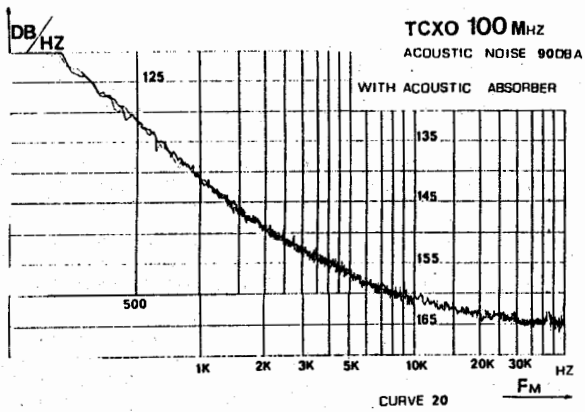
CURVE 18

Toutes les mesures ont été faites d'après le synoptique de la figure 11. Il ne s'agit encore, pour l'instant, que d'une méthode plus qualitative que quantitative. Il reste encore à codifier, sinon à normaliser le processus.

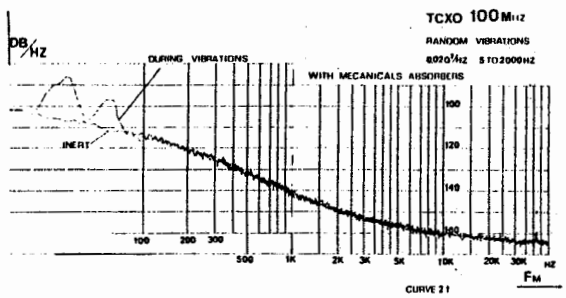
Enfin, après avoir fait appel aux mécaniciens pour pallier la sensibilité accélérométrique par un système d'amortisseurs, on pourra interpeller avec profit les chimistes pour la mise au point d'absorbants acoustiques.

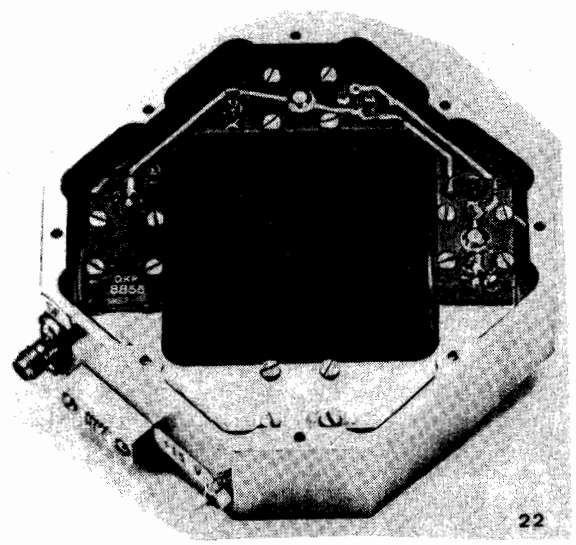
Les courbes n°19 et n°20 montrent des résultats obtenus en bruit acoustique par l'emploi d'élastomères chargés, simplement collés autour de la partie sensible de l'oscillateur.

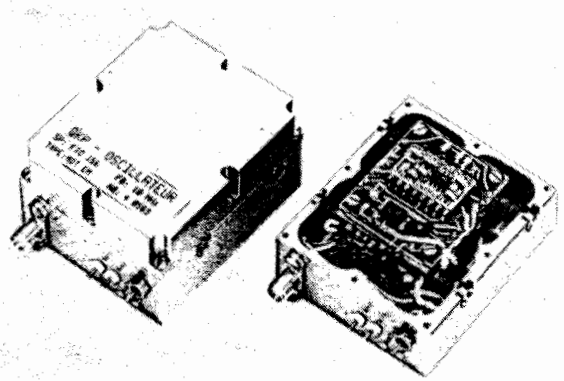




La courbe n°21 montre le comportement en vibration du même oscillateur. Les photos 22 et 23 sont des exemples de réalisation.







CONCLUSION

Après de longues expérimentations, nous sommes arrivés à un certain nombre de conclusions.

La recherche d'une sensibilité accélérométrique sans cesse améliorée est une quête longue et onéreuse.

Il est préférable, à partir d'un état de l'art acceptable industriellement, de sauvegarder le bruit de phase en vibration par l'emploi de structures mécaniques agissant comme un filtre.

Pour obtenir des résultats du même ordre sans amortisseurs, il faudrait atteindre des sensibilités accélérométriques de quelques 10^{-12}/g.

Cependant, très près de la porteuse (10 Hz), le bruit de phase en vibration continuera à dépendre de la sensibilité accélérométrique du résonateur.

Tout ceci est valable pour le quartz et le Tantalate de lithium.

La recherche devrait s'orienter maintenant vers l'amélioration de la sensibilité acoustique des résonateurs à quartz à l'inverse des résonateurs au Tantalate de Lithium pour lesquels les résultats apparaissent naturellement acquis.

ANNEXES

MESURE DE LA SENSIBILITE ACCELEROMETRIQUE DES RESONATEURS.

Moyens utilisés :

Excitation . Générateur sinusoïdal
. Amplificateur de puissance
. Pot vibrant
. Accéléromètre - amplificateur de charge

Moyens de mesure :

. Analyseur de spectre
(dynamique 120 dBc)

Le résonateur monté sur un oscillateur est soumis à une excitation vibratoire de fréquence connue et d'accélération donnée. Le calcul de la sensibilité accélérométrique est effectué à partir de la mesure des raies de modulation sur le spectre de l'oscillateur.

Connaissant l'amplitude des raies de modulation nous en déterminerons l'indice m et le ΔFc.
Excitation par des vibrations sinusoïdales.

$$\mathcal{L}(f)_{(dBc)} = 20 \log \left[\frac{\Delta F}{F_o} \cdot g \cdot \frac{F_o}{2 f_m} \right]$$

d'où $S = \frac{\Delta F}{F_o} = \left[\frac{2}{F_o} \cdot \frac{f_m}{g} \right] \cdot 10^{\frac{\mathcal{L}(f)}{20}}$

$S : \frac{\Delta F}{F_o}$ = Sensibilité du résonateur par g (accélération)

g = Accélération (nombre de g)

Fo = Fréquence de l'oscillateur

fm = Fréquence de vibrations

Ces formules sont à utiliser pour des indices de modulation : $m \leq 0{,}01$.

Pour des indices de modulation $> 0{,}01$ il faut utiliser les fonctions de Bessel.

Excitation par des vibrations aléatoires

L'accélération exprimée en g^2 /Hz à calculer en grms dans une bande de 1 Hz

$$\mathcal{L}(f)\ db/Hz = 20 \log \left[\frac{\sqrt{2}}{2} \cdot g_{rms} \cdot \frac{\Delta F}{F_o} \cdot \frac{F_o}{f_m} \right]$$

d'où $S = \frac{\Delta F}{F_o} = 10^{\frac{\mathcal{L}(f)}{20}} \left[\frac{2}{\sqrt{2}} \cdot \frac{f_m}{g_{rms} \cdot F_o} \right]$

MESURE DE LA SENSIBILITE ACOUSTIQUE

Moyens utilisés :

Excitation
- . Générateur de bruit
  Bande de fréquence 20 Hz à 20 KHz
- . Amplificateur de puissance
- . Chambre anéchoïque
- . Microphone - amplificateur

Moyens de mesure :

- . Banc de mesure de bruit de phase des oscillateurs.

Les moyens mécaniques (pots vibrants) suivant leur taille peuvent être utilisés à des fréquences maximum de 2000 à 10.000 Hz. A accélération constante, l'amplitude ou le déplacement diminuant en fonction du carré de la fréquence ($g \approx 0{,}02 . f^2 . d$) certaines limites sont atteintes (asservissement avec précision), enrichissement du spectre d'excitation, résonances des structures mécaniques.
Pour une accélération de 10 g, de fréquence 10 000 Hz, le déplacement C à C est de 0,05 μm.

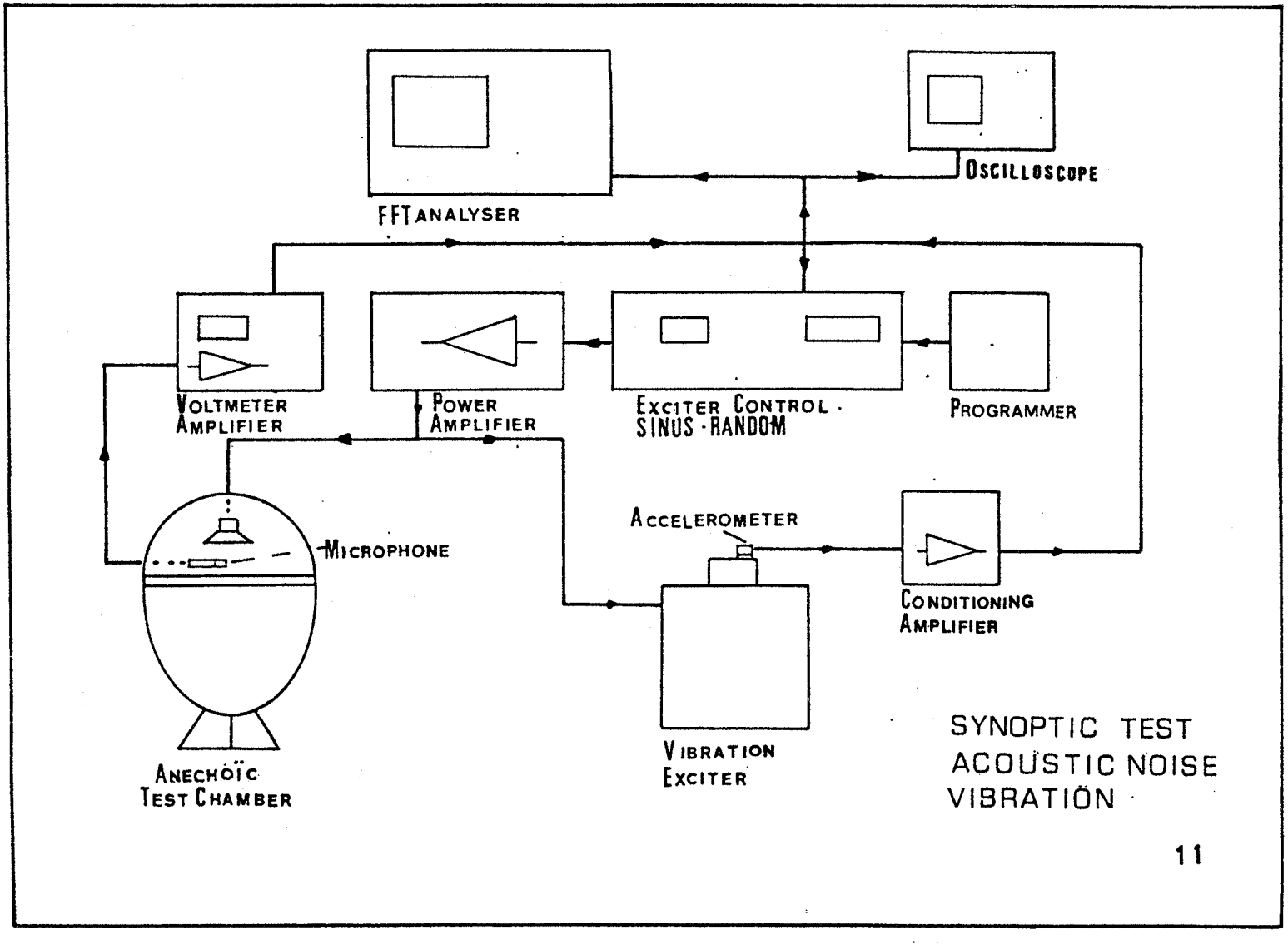
La méthode d'excitation acoustique nous permet d'obtenir un spectre régulier en bruit blanc d'une puissance de 90 dB(A) à $\pm$ 3 dB dans une bande de fréquence de 20 Hz à 20.000 Hz.

Une petite chambre anéchoïque nous permet de tester les résonateurs en Laboratoire.

La reproductibilité des pectres d'excitation permet de caractériser les structures mécaniques de la tenue de la lame du résonateur et de connaître l'influence de ces résonances mécaniques sur le spectre de phase de l'oscillateur.

BIBLIOGRAPHIE

- The effect of vibration on Frequency standards and clocks - R.L. FILLER - 35th Frequency Control Symposium.
- Influence of Environment conditions on a quartz resonator - M. VALDOIS, J. BESSON, J.J. GAGNEPAIN - Proceedings 28th Annual Symposium on Frequency control.
- Low 'g' Sensitivity Crystal Units and their testing A. WARNER, B. GOLDFRANK, M.MEIRS and M.ROSENFELD - 33rd Annual Symposium on Frequency Control.
- Design of a Bulk Wawe Quartz Resonateor Insensitive to Acceleration - R. BESSON, J.J. GAGNEPAIN, D. JANIAUD and M. VALDOIS - 33rd Annual Symposium on Frequency Control.
- Analyse des mécanismes de sensibilité accélérométrique et barométrique des résonateurs à quartz de type QAS - J.P. AUBRY, J.C. CRAVEUR, F. DEYZAC, M.FRAGNEAU Revue Technique THOMSON-CSF - Vol. 19 n°2, Juin 1987.


FFT ANALYSER
OSCILLOSCOPE
VOLTMETER AMPLIFIER
POWER AMPLIFIER
EXCITER CONTROL · SINUS · RANDOM
PROGRAMMER
MICROPHONE
ACCELEROMETER
CONDITIONING AMPLIFIER
ANECHOÏC TEST CHAMBER
VIBRATION EXCITER
SYNOPTIC TEST
ACOUSTIC NOISE
VIBRATION
11

TEMPERATURE DEPENDENCE OF AN UHF OSCILLATOR USING A DIELECTRIC RESONATOR

R.E. Elcheikh, M. Valentin

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

An UHF Dielectric Resonator Oscillator (DRO) using a rectangular partly metallized Dielectric Resonator (DR) is presented. In this paper are studied the oscillation characteristics as a function of the circuit parameters, with special attention to the effect of temperature on the frequency stability. One interesting characteristic of the device is the temperature variation of the oscillation frequency with the bias voltage of the amplifier. This temperature characteristic shows a common turning point for the particular resonator positions.

Introduction

Oscillators represent the basic energy sources for all microwave communication systems such as radar, radio-telephone and future land mobile satellite communication systems. For these applications, a local oscillator is an essential device which must operate at constant frequency, generate as little noise as possible and be compatible with the miniaturization of Microwave Integrated Circuits (MIC) in which waveguides, coaxial lines and cavities are replaced by microstriplines and dielectric resonators.

Stabilized microwave oscillators using dielectric resonators in the 4-25 GHz frequency band have been reported in the literature by many authors [1,2,3,4].

Size reduction of electronic circuits is now realised with the development of recent semiconductor technologies. However in systems operating in the 800 MHz band the dielectric resonator is not small enough. In the proposed oscillator, a size reduction of dielectric resonator is achieved by using a quarter wave dielectric rectangular resonator with three metallized faces [5]. The dielectric (Zr,Sn) TiO_4 offers an adjustable temperature coefficient [6], low loss and high dielectric constant (36 to 38).

Experiments were performed to investigate the DRO behavior with respect to several parameters such as the DR resonant frequency, the d-c bias voltage and temperature variations.

DRO circuit

The dielectric resonator, which is the frequency determining element in the DRO configuration, is used in the feedback circuit. The amplifier stage is a broad band monolithic, cascadable amplifier (21 dB gain, noise figure : 5,5 dB, 50 Ω input/output impedance).

Fig. 1 shows two oscillator configurations, fabricated on a dielectric substrate (RT Duroid 6010). The configuration of Fig. 1b was used to build both DRO and voltage-controlled oscillators (VCO's). A buffer amplifier is integrated to provide a better load isolation and to minimize the effects of frequency pulling.

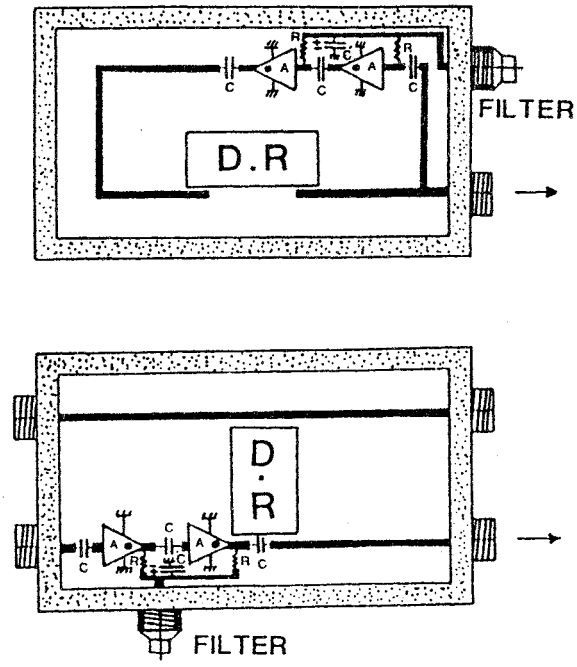


Fig. 1 : Configurations of dielectric resonator oscillator
a : Two-port DR b : Four-port DR

The DR magnetically coupled to two microstriplines operates as a band-transmission filter in the configuration of Fig. 1a ; in the configration of Fig. 1b the DR operates both as a band-rejection filter between the ports 1 and 2 and as a transmission filter between the ports 1 and 3 or between the ports 1 and 4.

The resonant frequencies of the two dielectric rectangular resonators tested (dimensions : 18 × 10 × 10 mm) and (18 × 6 × 6 mm) are 873 MHz and 925 MHz respectively, the temperature range for experiments is between +10°C to +50°C and the maximum d.c. bias voltage of the DRO is 11.5 V, corresponding to an output power of 17 dBm. A pushing figure of -575 kHz/V is obtained for bias voltage ranging from 8 V to 11.5 V.

Temperature dependence of frequency oscillation

The relative importance of all factors affecting the resonant frequency stability must be considered when a high stability is required, especially in a local oscillator application [7]. The temperature dependence of the frequency f_r of electromagnetic standing waves in a dielectric resonator is expressed by the temperature coefficient τ_{fr}, which is a function of the thermal expansion coefficient τ_a and of the temperature coefficient $\tau_{\varepsilon r}$ for the permittivity ε_r. A low τ_{fr} can be obtained if the quantities $\tau_{\varepsilon r}$ and τ_a compensate one another [8].

The amount of coupling between the DR and the microstripline has a direct effect upon output power and frequency stability, as well as resonant frequency. In the DR oscillator circuits the thermal expansion coefficient of the microstrip line and the contribution from the other components in the circuit must be compensated by the contribution from the resonator. Metallic shielding required to minimized the radiation losses affects the frequency also. With a suitable design the effect of the shielding package can be made small compared with the effect of the active element and the resonator. The coefficient τ_{fr} of the resonator can be made ajustable between -10 ppm/°C and +10 ppm/°C by means of a dielectric spacing "e" between the resonator and the substrate.

The parameters controlling the oscillating characteristics are :

- resonator dimensions
- distance ℓ between the center position of theDR and a reference plane
- distance d between the microstrip edge and the DR edge
- gap e between the DR and the substrate
- bias voltage of the amplifier.

The principal DRO frequency drift with the temprature is the phase deviation between the DR and the active circuit. The temperature coefficient of the oscillator τ_{fo} can be expressed as :

$$\tau_{fo} = \frac{F(\beta)}{Q_R} \frac{d\Phi_d}{dt} + \tau_{fr}$$

where $F(\beta)$ is function of the coupling factor β of the DR with the microstripline and the position of the DR along the transmission line, and ϕ_d the argument associated with the input impedance of the active element.

Measurements of oscillation frequency variations

When the temperature is changed, the metallic shielding and the electronic circuit respond rapidly to temperature changes and reach thermal equilibrium sooner than the dielectric resonator. When making measurements as a function of temperature, the time interval between successive temperature steps must therefore exceed the thermal time constant of the dielectric resonator oscillator.

Frequency variations with time

Fig. 2 shows a typical plot of the oscillation frequency as a function of time at a fixed temperature ; a fractional frequency change below to 3.6 ppm is observed during 12 days.

Bias dependence

The position of the DR along the transmission line is varied experimentally until the resonator achieves the proper phase angle to support oscillations. The oscillation frequency characteristics of the two-port oscillator are shown in Fig. 3, as a function of the bias voltage V_B for two DR positions with the temperature T (°C) taken as parameter. These figures show a common point of oscillation frequency, corresponding to a bias voltage V_p = 10.35 V for one DR position (Fig. 3a) and V_p = 10.6 V for another position (Fig. 3b).

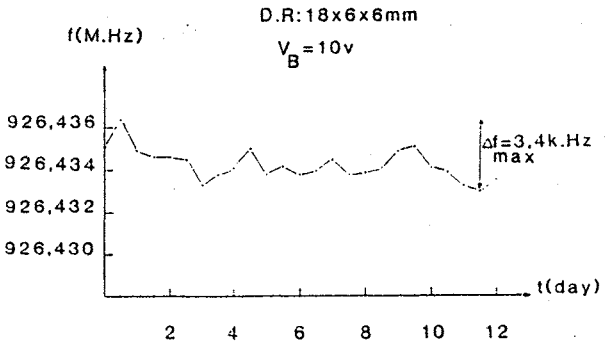
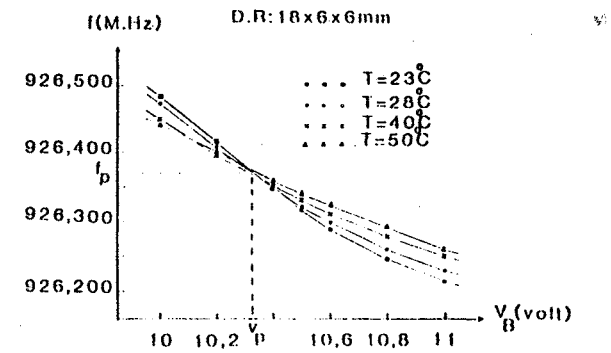


Fig. 2 : Oscillation frequency as a function of time



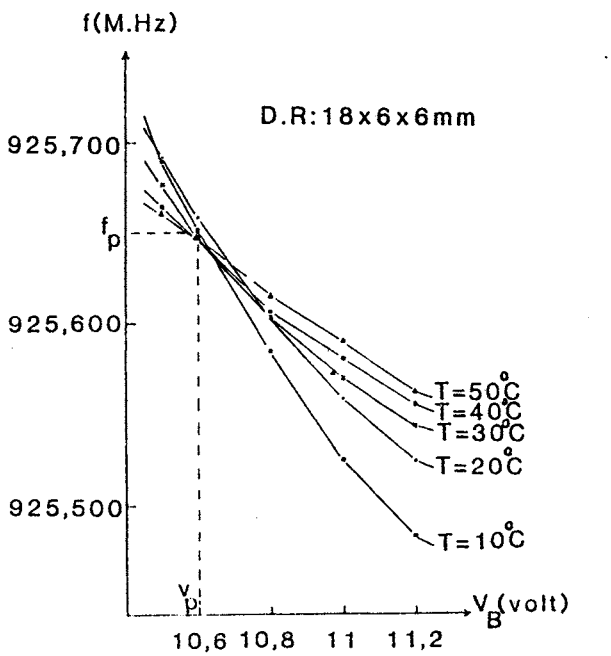


Fig. 3 : Oscillation frequency as a function of the bias voltage for two DR positions

As the temperature increases, the oscillation frequency increases for the voltages superior to V_p and decreases for those inferior to V_p. Therefore, it is expected that the frequency stability is improved at this bias voltage.

For an other DR of different dimensions (18×10×10 mm), disposed in the same environment, the characteristic is linear and does not present a common point of oscillation frequency.

Temperature dependence of oscillation frequency

Fig. 4 shows change of the oscillation frequency versus temperature T ranging from +10°C to +50°C with bias voltage as a parameter.

A frequency variation of less than -13 kHz over the temperature range from 10°C to 50°C is obtained at the bias voltage of 10.6 V.

The measurements show the effect of the environment humidity for temperature inferior to 15°C ; water vapor is liquefied on any part of the device and affects the DRO stability. For stability improvement at a temperature inferior to 15°C, precautions must be taken.

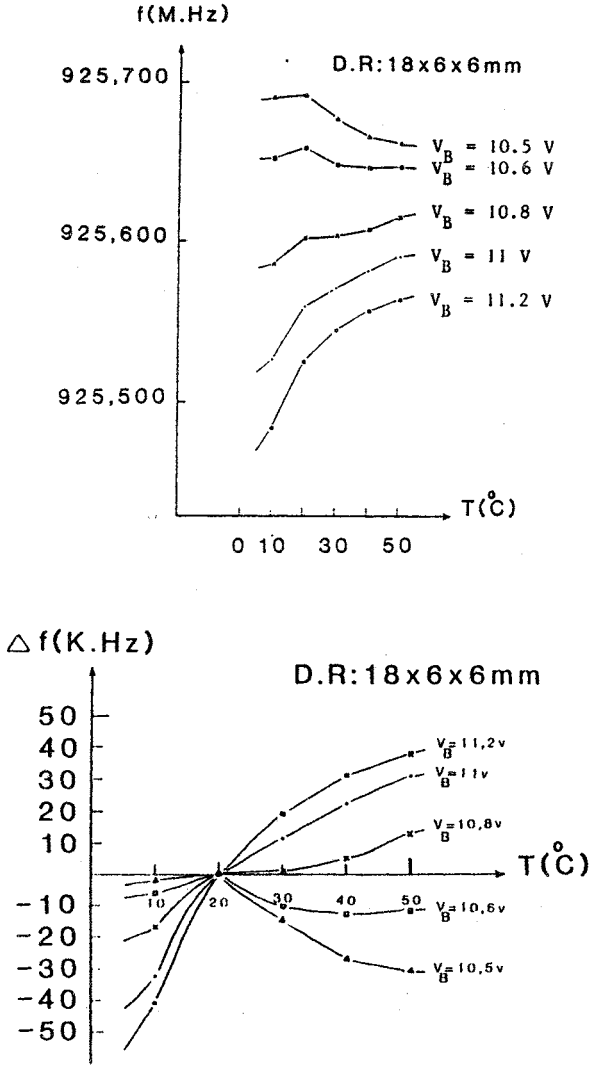


Fig. 4 : oscillation frequency and frequency deviation as a function of temperature

Short term stability

Oscillator stability has been measured in the time domain by evaluating the fractional frequency fluctuations $<\sigma_y(\tau)>$ versus the averaging time τ and expressed as Allan's variance. Fig. 5 shows the short term stability of the two-port DR oscillator : an improvement for $\tau > 1$ s is obtained for bias voltage V_B = 10.6 V.

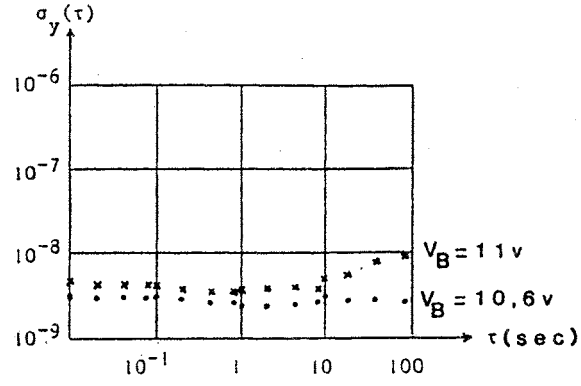


Fig. 5 : Short term frequency stability

Conclusion

Dielectric resonators have brought significant improvement to the design of microwave oscillators. In the range of UHF frequencies, below 1 GHz, the large size of the final oscillator is determined by the size of the dielectric resonator. Compactness, light weight, temperature stability and relative low costs are achieved with partly metallized λ/4 dielectric resonators.

The factors affecting the frequency stability with temperature have been investigated. It is shown that the circuit parameters and the resonator position contribute considerably to the stability. As a result it is found that the oscillation frequency-voltage characteristics with temperature taken as parameter exhibit a common point of oscillation frequency for a bias voltage V_p. The temperature stability and short term stability in the range of 1 s to 100 s are improved ; a fractional frequency change of -0.35 ppm/°C is obtained.

References

[1] H. Abe, Y. Takayama, A. Higashisaka, H. Takamizawa, "A highly stabilized low-noise GaAs FET integrated oscillator with a dielectric resonator in the C-band", IEEE Trans. Microwave Theory Tech., vol. MTT-26, pp. 156-162 (1978).

[2] J.K. Plourde, C.L. Ren, Application of dielectric resonators in microwave components", IEEE Trans. Microwave Theory Tech., vol. MTT-29, pp. 754-770 (1981).

[3] Z. Galani *et al.*, "Analysis and design of a single-resonator GaAs FET oscillator with noise degeneration", IEEE Trans. Microwave Theory Tech., vol. MTT-32, pp. 1556-65 (1984).

[4] M.J. Loboda, T.E. Parker, G.K. Montress, "Frequency stability of L-band, two-port dielectric resonator oscillators", IEEE Trans. Microwave Theory Tech., vol. MTT-35, pp. 1334-38 (1987).

[5] M. Valentin, R.E. Elcheikh, "Microwave oscillator using partly metallized dielectric resonator", Proc. of the 2nd E.F.T.F., Neuchâtel, 1988.

[6] J.C. Mage, B. Marcilhac, "Amélioration de la stabilité en température des matériaux pour résonateurs diélectriques", Actes des 5ème Journées Nationales Microondes, Nice, pp. 41-43 (1987).

[7] M. Valentin, "Bruit de phase et de fréquence de résonateurs diélectriques", In "Interaction of radiation with matter", a volume in honour of A. Gozzini, Scuolo Normale Superiore, Pise, pp. 174-185 (1987).

[8] T. Higashi, T. Makino, "Resonant frequency stability of the dielectric resonator on a dielectric substrate", IEEE Trans. Microwave Theory Tech., vol. MTT-29, pp. 1048-52 (1981).

ACCURATE DELAY CALIBRATION OF SATELLITE GROUND STATIONS FOR TWO-WAY TIME TRANSFER

G. de Jong

Van Swinden Laboratory VSL
Delft
The Netherlands

Summary.

The present state-of-the-art intercomparison of International Atomic Time Scales using the US Global Positioning System (GPS) is limited to about 30 ns systematic uncertainty and 10 ns statistical uncertainty. A recently proposed two way method using Phase Modulation with Pseudo Random Noise signals (Mitrex) and small satellite ground stations that were located at the sites of the participating laboratories, has demonstrated a statistical uncertainty (precision) of less than 1 ns; however, the determination of the absolute uncertainty (accuracy) could not be performed due to unknown asymmetry in the ground station delays. These delays should become known to an uncertainty of 5 ns or less.

The calibration method described in this paper is focussed towards the separate measurement of the delay in the ground station from the 11 GHz antenna to the 70 MHz IF input of the receiver/correlator, and from the 70 MHz IF transmitter output to the 14 GHz antenna; both including the antenna feeds, the upconverter, cables, amplifiers and downconverter. The measurements are performed by using the 70 MHz transmit/receive unit (modem) itself. A simulator of the satellite transponder is placed in front of the antenna. First the combined receive delay and transmit delay of the ground station is measured. Then by using a calibrated cable to the satellite simulator the total receive delay is measured. Finally the transmit delay is calculated by subtracting the receive delay from the combined delay. Both delays are thus determined to an absolute uncertainty in the order of 1 ns. Also a measurement and calculation to determine the internal delay of the IF transmitter and the IF receiver is reported.

Introduction.

Clocks are providing the time of the day. The most accurate clocks are using cesium beam frequency standards as their source of accuracy. International agreement in the CGPM has established the definition of the duration of the second based on the physical property of cesium 133 atoms, and the atomic time scale TAI is made of these SI seconds by adopting a definition of the origin. To have a practical time scale close to the astronomic time, UTC has been introduced, which is equal to TAI, but kept in close agreement with solar time by addition or subtraction of so-called leap seconds. UTC and TAI are computed by the BIPM from the time comparison data of laboratory and commercial cesium clocks. The accuracy of comparing clocks all over the world has been improved recently by the use of the GPS common view method. This is a one way method capable of precision around 8 ns and accuracies of about 30 ns.
A further improvement can be realized by using two way methods, as demonstrated by experiments using ground stations at each clock site capable of reception and transmission. This paper is addressing the calibration of the ground stations to achieve a high accuracy in such two way time comparisons.

General principle of time comparison.

To measure the difference of the time scales of two clocks a time interval counter is used. One clock output is connected to the start of the counter, and the other clock output (1 pps) to the stop input. State of the art counters measure the difference with a resolution of 20 ps, and an accuracy of 100 ps or 0.1 ns.
If the clocks are not close to each other a cable can be used to transport the output signal of the second clock to the stop input of the time interval counter. However, the time interval reading is not correct because of the propagation delay of the 1 pps signal in the cable. This delay is to be measured and corrected for. A 50 Ω coaxial cable has typically a delay of about 5 ns/m.
If the cable is relative long then distortion of the pulse occurs which gives an extra delay. In a cable the higher frequencies are more attenuated than the lower frequencies, and the cable acts as a low pass filter. To avoid such a group velocity distortion, modulation of the 1 pps signal on a carrier frequency can be done. So the additional delay from distortion can be minimized.

In stead of using a cable for transportation of such a modulated carrier the space can be used by means of transmitters and receivers with the appropriate antennas. This is done with low frequency transmitters for time signals such as DCF77 at 77.5 kHz, MSF60 at 60 kHz, HBG at 75 kHz; also Loran-C and HF standard frequency and time transmitters. The maximum reception distance is several thousands of kilometers.
A problem for accuracy is the determination of the distance from transmitter to receiver and to correct for propagation properties of the medium: troposphere, ionosphere, (barometric pressure, humidity, degree of ionization).

Satellites are now being used. One problem is the very long distance and the accuracy of the determination of this distance. Another problem is the ionospheric delay; the latter however is diminished by the square of carrier frequency and is less than 1 ns at 12 GHz. The two-way method using satellites eliminates the need to determine separately the exact distance from each station to the satellite, as is shown in the next chapter.

Two way time comparison.

From fig 1 we can see that the difference of the clocks at station 1 and 2 can be determined.

TA(k) is the time scale at station k
TI(k) is time interval reading
TT(k) is transmitter delay
TR(k) is receiver delay
TU(k) is uplink delay
TD(k) is downlink delay
TS(k) is satellite delay
TC(k) is correction for relativistic effects

The difference TA(1)-TA(2) is to be determined. The reading at station 1 is:
TI(1) = TA(1) - TA(2) + TT(2) + TU(2) + TS(2) + TD(1) + TR(1) + TC(1).
The reading at station 2 is:
TI(2) = TA(2) - TA(1) + TT(1) + TU(1) + TS(1) + TD(1) + TR(2) + TC(2).
The difference gives:
TI(1) - TI(2) = 2(TA(1) - TA(2)) + TT(2) - TT(1) + TU(2) - TU(1) + TS(2) - TS(1) + TD(1) - TD(2) + TR(1) - TR(2) + TC(1) - TC(2)
or:
TA(1) - TA(2) = ½ { TI(1) - TI(2) + TT(1) - TT(2) + TU(1) - TU(2) + TS(1) - TS(2) + TD(2) - TD(1) + TR(2) - TR(1) + TC(2) - TC(1)}.

If the same transponder in the satellite is used, and the transmit antenna and the receive antenna are on the line of symmetry between station 1 and 2, then:
TS(1) = TS(2)
If the satellite is not moving too fast then :
TU(1) = TD(1) TU(2) = TD(2), and their delays cancel.
The TC(1) en TC(2) can be calculated accurate enough from the station coordinates and the satellite position.

The remaining terms are now:
TA(1) - TA(2) = ½{TI(1) - TI(2) + TT(1) - TT(2) + TR(2) - TR(1) + C}

If the transmit delays TT and the receive delays TR at both stations were equal, then they would cancel. But that can only be proved by measuring the difference by locating two ground stations close enough together, which is not a simple action. Even then the sum of the transmit difference and the receive difference is determined, not their values separately.
Therefore here a method to determine the absolute value of the transmit- and receive-delays separately will be proposed.

The time transfer modem.

The MITREX modem (fig 2) has a modulator section in which the 1 pps signal is impressed on a bi-phase modulated carrier. The PRN generator is clocked at a 2,5 MHz rate and every 4 ms a bit sequence is generated. As soon as the 1 pps signal appears, then one sequence of 4 ms is inverted. The sequence is re-clocked together with the 1 pps transmitter pulse and then phase modulates a 70 MHz carrier frequency. The signal is band pass filtered and is fed as IF signal to the upconverter and is transmitted.

The demodulator section receives the 70 MHz downconverted signal and provides a 1 pps received pulse as soon as it detects the 4 ms inverted PRN sequence.

The sum of the internal delays in the modem can be measured by connecting the 70 MHz output with the 70 MHz input and connecting the 1 pps transmitted output to the start input of a time interval counter and the 1 pps received output to the stop input. The delay of any external cable can be measured by connecting the cable also between modem output and input. The cable delay is found by subtracting the internal modem delay.

The calibration of the ground station delays.

A typical earth station is shown in fig. 3. The delay between the reference plane at the time interval counter and the plane at the antenna is to be determined. First the delay between the antenna and the modem reference plane will be measured. For this purpose a passive satellite transponder simulator is constructed. It is consisting of two double balanced mixers connected in series between two SMA to wave guide transitions acting as antennas. The mixers convert the transmit frequency (14 GHz) to the receive frequency (11 GHz). So one local oscillator is tuned to 2.93 GHz and the other is at 70 MHz.
The modem is used to measure the sum (TT + TR) of the transmit delay TT and the receive delay TR through the satellite simulator at the modem reference plane.
This is similar to the test loop translator which in some ground stations is installed between the output of the upconverter and the downconverter input, but now the delay in the antenna feeds is also included. The internal delay of the satellite simulator is estimated 2 ns: the cable length between input and output is 0.27 m and the mixers are very wide band (2 - 18 GHz).

In order to determine the values of the transmit and receive delay separately, the 70 MHz cable CS to the satellite simulator is used. The delay of this cable CS is measured first. It is disconnected from the simulator and so are the 70 MHz cables CT and CD to the up- and the downconverter respectively.
First cable CT is interconnected to cable CD at the far end and the sum delay (CT + CD) is measured.
Then cable CS is substituted for CT and the sum delay (CS + CD) is measured.
Finally (CS + CT) is measured. From the obtained 3 measurements the delay of cable CS is calculated:
CS = ½{(CS + CD) + (CS + CT) - (CT + CD)}.
Also the other cable delays can be calculated.

The calibrated cable CS is reconnected to the satellite simulator and the other cables are also reconnected. The next step is to connect cable CT to the 70 MHz continuous carrier and cable CS to the modulated 70 MHz modem output. Now the CW carrier of 14 GHz is PRN modulated in the satellite simulator. The sum of CS and the receive delay TR from simulator to modem input is measured: (CS + TR).
The delay CS is known , so TR can be calculated. The sum (TT + TR) of transmit delay TT and TR was already measured, so TT can also be calculated.

It should be emphasized that TT and TR are measured to the reference plane at the modem 70 MHz in- and output. Still undetermined is the internal transmit delay Tt with respect to the 1 pps received output.
Let us look again at the modem in Fig. 2. The transmit part is the simpler, after the re-clocking a wide-band mixer is used (delay estimated <2 ns) followed by a band-pass filter. Filters exhibit always delays, increasing as the band-width decreases. Also the receiver has a band-pass filter at its input. So the transmit filter was taken out of the circuit and then the total internal delay was measured again. The difference between filter in and out is the delay of the filter. Further transmit delays are calculated from the measured cable lengths from the wide-band mixer to the 70 MHz output connector. Once the total internal transmit delay Tt being determined, the internal receive delay Tr is calculated by subtracting the transmit delay Tt from the total internal delay.
Now the internal delays can be added to the external ground station delays to find the total receive delay as well as the total transmit delay from the antenna reference plane to the time interval counter reference plane.

From Fig. 3 it is clear that the antenna reference plane is not coinciding with the intersection of the axis of rotation, the latter being the station coordinates; the distance between them can be measured and a correction using the speed of light should be applied.

The systematic uncertainty appearing from estimation of the internal delays of the mixer in the modulator and in the satellite simulator is estimated to be less the 2 ns. The statistical uncertainty, or precision is of the order of 1 ns but is depending upon the averaging time.

At VSL the antenna is connected with IF cables of about 120 m to the modem. The values found at VSL were: CS = 628.7 ns, and TT = 669.2 ns, TR = 648.6 ns,

Tt = 319.3 ns, Tr = 845.3 ns; so the total transmit delay is 988.5 ns and the receive delay 1493.9 ns. So an asymmetry of 505.4 ns exists.

Conclusion.

In this paper it has been shown that with the described calibration method using a satellite simulator all transmit and receive delays can be determined to the nanosecond level accuracy. It is important to know the total transmit and receive delays at each earth station for two-way time comparison, symmetry should not be assumed at the nanosecond accuracy level.

References.

Ph. Hartl et al., "High accuracy global time transfer via geosynchronous telecommunication satellites with Mitrex", Journal Flight Science & Space Research, vol. 7, no. 5, pp. 335 - 342, 1983.

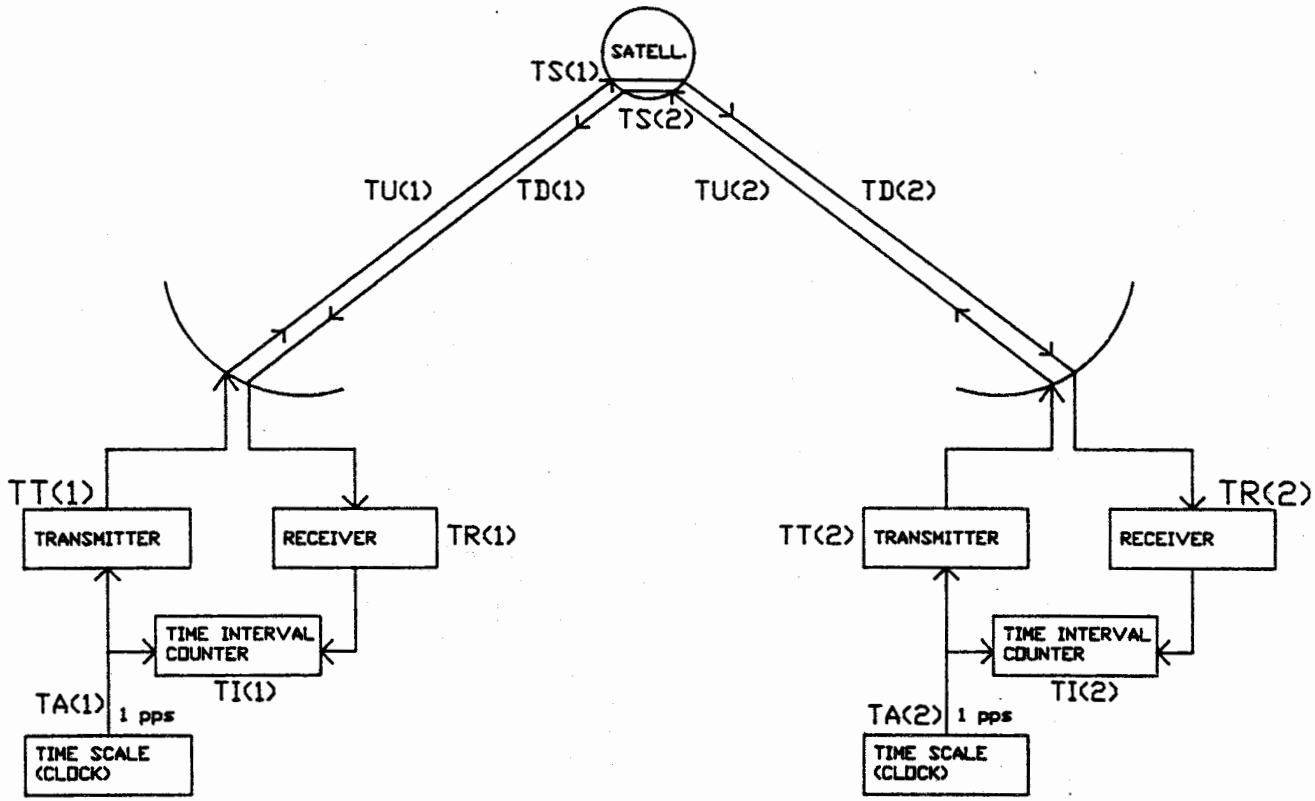
D. Kirchner et al., "A two-way time transfer experiment via ECS-1 using the MITREX modem", IEEE Trans. on Instr. and Meas., vol. 37, no. 3, Sept. 1988.

G. de Jong, "Measuring the propagation time of coaxial cables used with GPS receivers", Proceedings 17th Annual PTTI Applications and Planning Meeting, Dec. 3-5, 1986, Washington DC, USA.

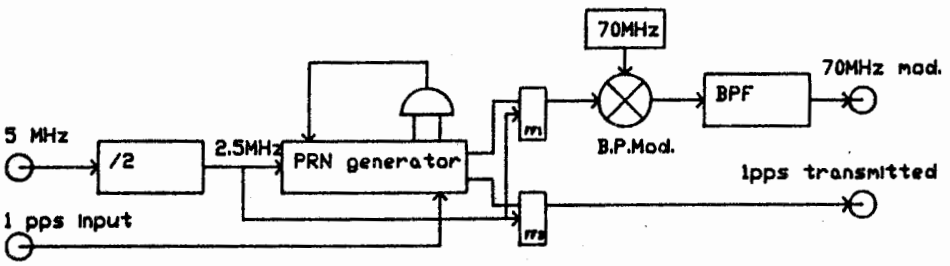
G. de Jong, "Results of several years of comparison of European Time scales by means of a receive only satellite method", Conference on Precision Electromagnetic Measurements, June 1984, Delft, the Netherlands.

W. Klepczynski et al.,"Preliminary comparison between GPS & two-way satellite time transfer", Proceedings 42nd Annual Symposium on Frequency Control, 1-3 June 1988, Baltimore, USA, IEEE Cat. no. 88CH2588-2, pp. 472-477.

Fig. 1 Two-way Time Comparison Principle



Mitrex Modulator section



Mitrex Demodulator section
(receiver)

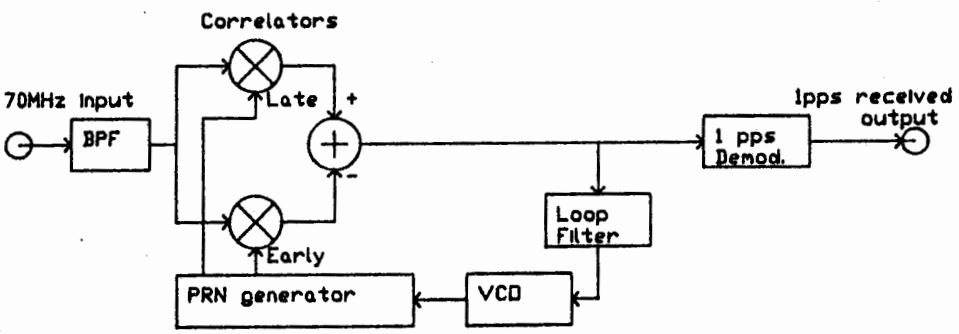


Fig. 2 Mitrex Modem

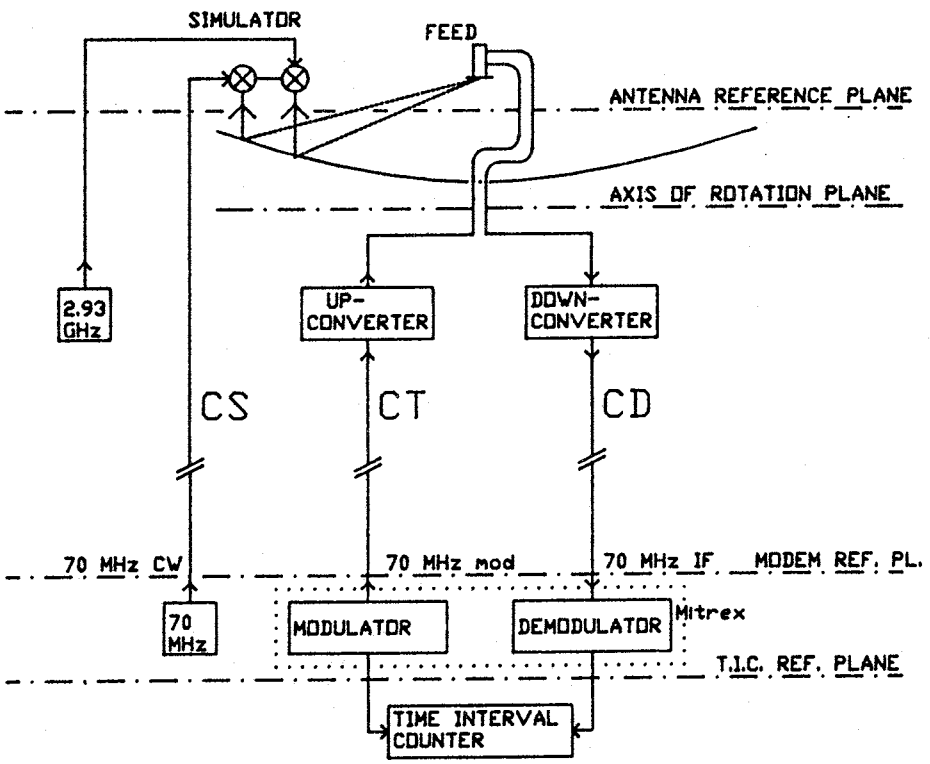


Fig. 3 Earth Station Calibration

Synchronization of time scales by television method using ECS satellites
Preliminary results

O. Buzek, J. Čermák*, J. Vondrák**, F. Cordara, P.G. Galliano, V. Pettiti***, P. Tavella****

* IREE-CSAS, Institute of Radio Engineering and Electronics, Prague, Czechoslovakia
** Astronomical Institute, Prague, Czechoslovakia
*** Istituto Elettrotecnico Nazionale Galileo Ferraris, Turin, Italy
**** Politecnico di Torino, Italy

This work has been done in the frame of a scientific cooperation program between Czechoslovakia and Italy.

Abstract

A time synchronization experiment between the time scales of the Istituto Elettrotecnico Nazionale G. Ferraris (IEN) in Torino and of the Institute of Radioengineering and Electronics (IREE) in Prague, based on the use of the telecommunication satellite ECS, started in July 1988 and is in progress at present.

The aim of this experiment, based on the use of the well known television passive method applied to the signals broadcasted by the satellite, is to establish a system allowing high precision clock comparisons at moderate cost over Europe.

An evaluation of the precision obtained in the first months of comparisons between IEN, IREE and other European laboratories is presented.

The improvements obtainable taking into account the elements of the satellite orbit are also evaluated.

1 - Introduction

The synchronization of time scales by the television method [1] using the ground TV network has been experienced, in the course of time, by most of the laboratories contributing to the definition of UTC. It is known that, within the view of a common TV transmitter, this simple and inexpensive method can yield a timing precision of about 10 ns rms. The precision limits are affected primarily by the interference with the waves reflected from surrounding objects, particularly in urban areas, and by the propagation delay changes due to weather conditions. An obvious disadvantage of this method stems from the common view requirement which, for the wave propagated above the ground, restricts the time comparisons within relatively small areas. When extended to larger areas, i.e. when different transmitters interconnected by microwave links must be used, the timing precision drops by about one order [2] due to path delay instabilities of the microwave links, TV signal remodulations, different and unstable parameters of TV transmitters.
Thus the ground TV time comparisons are being gradually substituted by GPS, at least in those laboratories where the GPS instrumentation is available.

Nowadays there seems to be a new opportunity for the TV method in connection with the geostationary communication satellites which cover large territories with different national TV programs. The idea of using the geostationary satellite television signal (GSTS) for the synchronization of time scales is not new [3,4,5,6,7], however no attempt has been made to elaborate the method, both technically and organizationally, to such a degree that it could serve on a routine basis to various users.

The purpose of this paper is to inform the time and frequency community about a joint experiment which started in July, 1988, and which is carried out in collaboration between the IEN, Torino, the IREE, Prague, and the ASMW*, Berlin, later joined by the AOS*, Borowiec and the TUG*, Graz. The final aim of the experiment is to establish a system based on the GSTS method which would allow for high precision time comparisons at moderate costs all over Europe.

2 - Estimation of Timing Accuracy

The GSTS method possesses the advantage of signal source common view for all of the observers who can be, in contradiction to the ground TV, set apart over large territories. Advantageous properties of the signal coming from the satellite at 11 GHz are also evident: negligible overall propagation delay changes, no reflections and resulting interference.
Actually, system analysis of the GSTS timing potentials hints that there are no serious system drawbacks in comparison with the currently used timing via the GPS.

The key question of the GSTS method is how to determine the satellite instantaeous position in order to evaluate the differential delay with sufficient accuracy. Evidently, the differential delay error will be equal to the timing error. The problem is that the satellite in the geostationary orbit is not fixed with respect to its observers. Due to the natural forces it is in a permanent motion maintained within a so called "parking box" by occasional corrective manoeuvres initiated by the control tracking center on the Earth. Generally the satellite position may not deviate by more than 0.1° from its nominal position, both in longitude and latitude [7]. Though small the deviation seems, it may still cause unacceptable changes of the differential delay depending, for a given satellite, on the position of the observers.

---

* ASMW - Amt für Standardisierung, Messwesen und Warenprüfung, Berlin, DRG.
AOS - Astronomical Latitude Observatory, Borowiec, Poland.
TUG - Technische Universität Graz, Austria.

F. STHAL

2.1 - Time error distribution due to the satellite movement

The geostationary satellites chosen for this synchronization experiment belong to the family of European Communication Satellites (ECS) of the European Space Agency (ESA). In the followings we will deal in particular with ECS 5 satellite, that is positioned at 10 degrees East of Greenwich, to estimate the time synchronization error that one can expect over the European continent if the satellite coordinates are affected by unpredictable variations. The zone covered by the ECS satellites is shown in Fig. 1 where is outlined also the area served by the spotbeam west that has been used in our experiment.

Let us consider two receiving sites A and B and a geostationary satellite S, whose positions relative to a rotating equatorial geocentric reference frame can be expressed with the following vectors:

$$\begin{aligned} \vec{r}_A &= \vec{r}_A(\vartheta_A, \lambda_A, r_A) \\ \vec{r}_B &= \vec{r}_B(\vartheta_B, \lambda_B, r_B) \\ \vec{S} &= \vec{S}(\vartheta_S, \lambda_S, r_S) \end{aligned} \qquad (1)$$

where: ϑ_i is the geodetic latitude
λ_i is the longitude referred to Greenwich
r_i is the distance from the center of the Earth

The differential propagation delay of the signals transmitted by the satellite and received at stations A and B, can be expressed as:

$$t_p = \frac{1}{c}\left(|\vec{S}-\vec{r}_A| - |\vec{S}-\vec{r}_B|\right) = \frac{1}{c}\left[\varrho_A - \varrho_B\right] \qquad (2)$$

where ϱ_A and ϱ_B are the distances of the receiving stations from the satellite and c is the mean value of the speed of light along the propagation path.

To compute the variations of the differential propagation delay t_p between the two stations that affect the synchronization results, we have considered the satellite movements both in longitude $\Delta\lambda_s$ and latitude $\Delta\vartheta_s$ neglecting any change in its range. This assumption is justified as a first approximation because in our case the variation in range is nearly one fifth compared with the variations in latitude and longitude.

Differentiating the term $(\varrho_A^2 - \varrho_B^2)$ versus $\Delta\vartheta_s$ and $\Delta\lambda_s$, assuming that the variations of $(\varrho_A + \varrho_B)$ due to the satellite motion are negligible with respect to those of the term $(\varrho_A - \varrho_B)$ and, being the latitude of the satellite ϑ_s very close to zero, we can write the approximate expression:

$$\Delta(\varrho_A-\varrho_B) = \frac{2Sr_A}{\varrho_A+\varrho_B}\left|\cos\vartheta_A \sin(\lambda_S-\lambda_A)\Delta\lambda_S - \sin\vartheta_A\Delta\vartheta_S\right| - \frac{2Sr_B}{\varrho_A+\varrho_B}\left|\cos\vartheta_B \sin(\lambda_S-\lambda_B)\Delta\lambda_S - \sin\vartheta_B\Delta\vartheta_S\right| \qquad (3)$$

Solving this equation for different "B" sites in Europe after having fixed station A at IEN - Torino, several values of $\Delta t_p = \Delta(\varrho_A - \varrho_B)/c$ have been found assuming the worst case of 0.1° of variation both in latitude ($\Delta\vartheta_s$) and longitude ($\Delta\lambda_s$) of ECS 5. In Fig. 2 are reported the contours of the time errors that can affect the synchronization values as obtained from equation (3).

2.2 - Satellite position and timing accuracy evaluation

In what follows an analysis is made of the hypothetical configuration of 3 participating laboratories (station 1 = IEN, station 2 = IREE, station 3 = ASMW) and the ECS 5 satellite as shown in Fig. 3.

Consider a geostationary satellite, bearing a TV transmitter and orbiting in a near circular orbit around the Earth; the inclination of the orbit to the equator is supposed to be very small. Then, if the small perturbations of the orbit due to the oblateness of the Earth and the attraction of the Moon and the Sun are neglected for the moment, the rectangular coordinates of the satellite in a non-rotating geocentric coordinate system (xy-plane coinciding with the Earth's equator, x-axis poiting towards a fixed equinox and z-axis coinciding with the rotational axis of the Earth) can be simply approximated as

$$\begin{aligned} x &= a\left|\cos L - 3k/2 + (k/2)\cos 2L + (h/2)\sin 2L\right| \\ y &= a\left|\sin L - 3h/2 - (h/2)\cos 2L + (k/2)\sin 2L\right| \\ z &= 2a\,(q \sin L - p \cos L), \end{aligned} \qquad (4)$$

in which the higher-order terms have been neglected. Here a denotes the semi-major axis of the orbit, $k = e\cdot\cos(\omega+\Omega)$, $h = e \sin(\omega+\Omega)$, $q = \sin(i/2)\cos\Omega$, $p = \sin(i/2)\sin\Omega$ are the constant elements of the orbit, depending on the classical Kepler's elements e (eccentricity), ω (argument of perigee), i (inclination to the equator) and Ω (right ascension of the ascending node). The only time-dependent element is the satellite's mean right ascension $L = \omega + \Omega + M$, where M denotes the mean anomaly of the satellite, which is a linear function of time.

Further consider station 1, which determines the distance to the satellite by measuring the time interval elapsed from the transmission of a synchronisation pulse to the satellite and reception of the same pulse coming back to the station. Let the coordinates of this station in the rotating equatorial coordinate system (with X-axis pointing towards Greenwich meridian) be X_1, Y_1, Z_1. Stations 2 (with coordinates X_2, Y_2, Z_2) and 3 (with coordinates X_3, Y_3, Z_3) compare their clocks by measuring the time difference of reception of a TV synchronisation pulse transmitted from the satellite to both stations. The distance ϱ_i between the i-th station and the satellite can be calculated from the formula

$$\varrho_i = (x - X_i\cos S + Y_i \sin S)^2 + (y - Y_i\cos S - X_i \sin S)^2 + (z - Z_i)^2 \qquad (5)$$

where S is the Greenwich sidereal time (i.e. the angle between the equinox and Greenwich meridian). Assuming that the elements of the satellite's orbit are not known with sufficient accuracy, the value ϱ_i is a function of the unknown corrections to these elements. Consequently, the measured time interval Δt_1 at station 1 is also a function of these unknowns. The same holds for the measured time interval Δt_{23} at stations 2 and 3, which depends also on the difference between the two compared time scales.

The corresponding partial derivatives of ϱ_i with respect to these orbital elements are approximately equal to

$$\begin{aligned} \partial\varrho/\partial a &\cong 1 + 3ntY_s/2\varrho \\ \partial\varrho/\partial L &\cong -aY_s/\varrho \\ \partial\varrho/\partial k &\cong -a\cos L \\ \partial\varrho/\partial h &\cong -a\sin L \\ \partial\varrho/\partial q &\cong -2(Za/\varrho)\sin L \\ \partial\varrho/\partial p &\cong 2(Za/\varrho)\cos L, \end{aligned} \qquad (6)$$

where Y_s denotes the distance of the station from the plane containing the satellite and Z-axis (i.e. approximately a constant for a given station and

satellite), n is the mean angular velocity of the satellite in the non-rotating system (approximately equal to 6.300 radians per day) and t is the time in days elapsed from an arbitrarily chosen epoch. The linear dependence of $\partial\varrho/\partial k$ and $\partial\varrho/\partial p$ on the one hand and $\partial\varrho/\partial h$ and $\partial\varrho/\partial q$ on the other hand is obvious; it is clear that these unknowns can be estimated independently only if the satellite is observed from stations sufficiently separated in the north-south direction. A similar situation holds for $\partial\varrho/\partial a$ and $\partial\varrho/\partial L$ because $\partial\varrho/\partial L$ is a constant and $\partial\varrho/\partial a$ changes with time only very slowly.

We try to estimate simultaneously, from a set of the measured time intervals Δt_1, Δt_{23}, the orbital elements of the satellite and the difference of the compared time scales at stations 2 and 3, using the method of least squares. The observation equations for the residuals v (whose sum of squares is to be minimized) generally read

$$v = 2/c \sum_{i}^{6} \partial\varrho_1/\partial E_i \, \Delta E_i + 2\varrho_1/c - \Delta t_1 \qquad (7)$$

for the measurements Δt_1 at station 1 and

$$v = 1/c \sum_{i}^{6} (\varrho_2/\partial E_i - \partial\varrho_3/\partial E_i)\Delta E_i + \\ + \Delta T_{23} + \Delta\dot{T}_{23}\, t + (\varrho_2-\varrho_3)/c - \Delta t_{23} \qquad (8)$$

for the measurements Δt_{23} at stations 2 and 3. Here c denotes the speed of light, E_i stands for the i-th orbital element and ΔT_{23}, $\Delta\dot{T}_{23}$ are the searched for time scale comparison and its rate, respectively. The solved-for parameters are ΔE_i, ΔT_{23} and $\Delta\dot{T}_{23}$. We already saw that there are some dependences between certain partial derivatives $\partial\varrho/\partial E$ that can lead to dangerous singularities in normal equations. These suspicions have been confirmed by numerical models based on Eqs (7) and (8) and simulated observations of the European TV satellites, from stations 1, 2, and 3. The conclusions drawn from these modelled situations are the following:

1) It is impossible to mutually separate the unknowns k and p as well as h and q; the correlations between them are close to unity (always between 0.95 and 1.00). The measured time intervals are influenced far less by p and q than by k and h. The elements p and q must therefore be treated as known orbital elements rather than solved-for parameters.
2) It is almost impossible to separate the determination of the semi-major axis from the longitude of the satellite within a reasonably long period of observations (which can be by no means longer than the time between two successive manoeuvres of the satellite, i.e. approximately one month. The coefficient of correlation is less than 0.9 only if the satellite is more than 15° in longitude apart from the stations. It is therefore recommendable to fix the value of semi-major axis to its geostationary value and to solve only for the longitude of the satellite.
3) It is necessary to make measurements at least twice a day, but preferably three or four times a day, in order to avoid further correlations.
4) It is better to use measurements covering a longer time interval (at least several days) in a single adjustment, in order to strengthen the solution.
5) It is preferable if station 1 (i.e. the one providing pseudo-ranging) is distant from stations 2 and 3 (the ones comparing their clocks). Otherwise relatively strong correlation between ΔT_{23} and L occurs.
6) A better solution is obtained if the satellite's longitude differs substantially from the longitudes of the stations. From this point of view, the observation of ECS 5 from middle-European stations does not yield the best solution.

As a typical demonstration of solution let us show the results for a set of simulated TV receptions of ECS 5 (longitude 10°E) at three stations (IEN, IREE, ASMW) made every six hours during six consecutive days (25 measurements). Only three orbital elements (k, h and L) have been included as the solved-for parameters, together with clock difference (ΔT_{ij}) and its rate ($\Delta\dot{T}_{ij}$) between the i-th and j-th station. The matrix of correlation coefficients is shown in Tab. 1. It is clear that the correlations are rather weak with the exception of L and T_{12} which is however not dangerous. Assuming the standard deviation of a single time measurement equal to 100 ns (comprising not only the errors of time measurements themselves but also all the neglected perturbations of the orbit) the model gives the precisions of adjusted unknowns as displayed in Tab. 2. The worse value obtained for ΔT_{12} reflects the fact that this parameter is correlated with L, if one of the stations compared is identical with (or near to) the one used for pseudo-ranging.

	k	h	L	ΔT_{12}	$\Delta\dot{T}_{12}$	ΔT_{23}	$\Delta\dot{T}_{23}$
k	1						
h	0.03	1					
L	-0.02	0.05	1				
ΔT_{12}	0.02	-0.04	-0.77	1			
$\Delta\dot{T}_{12}$	0.00	0.00	0.00	0.00	1		
ΔT_{23}	0.00	-0.01	-0.21	-0.16	0.00	1	
$\Delta\dot{T}_{23}$	0.00	0.00	0.00	0.00	0.00	0.00	1

Table 1. Correlation coefficients

Parameter	Precision
k	$\pm$ 1x10^{-7}
h	$\pm$ 1x10^{-7}
L	$\pm$ 0.0007°
ΔT_{12}	$\pm$ 30 ns
$\Delta\dot{T}_{12}$	$\pm$ 10 ns/day
ΔT_{23}	$\pm$ 20 ns
$\Delta\dot{T}_{23}$	$\pm$ 10 ns/day

Table 2. Precision of the estimated parameters based on the assumed precision of a single measurement $\pm$ 100 ns.

3 - Description of the experiment

To test the potentiality of a synchronization system based on a geostationary satellite using the passive television method, on July 1988 four European laboratories began an experiment that is still going on. As will be explained in the followings, at this first stage, no station performed range measurement as was hypothesized in the previous paragraph.

The family of ECS satellites of Eutelsat and, in particular, since mid-October 1988, ECS 5, placed in a geostationary orbit at 10° east longitude, have been used.

Among the television channels, available from this satellite, was chosen the RAI UNO program transmitted

at 11.009 GHz in horizontal polarization by a 20 W transponder.

The participating laboratories have been the following ones:

- Astronomical Latitude Observatory (AOS) - Poland
- Amt für Standardisierung, Messwesen und Warenprüfung (ASMW) - Dem. Rep. of Germany
- Istituto Elettrotecnico Nazionale (IEN) - Italy
- Institute of Radio Engineering and Electronics (IREE) - Czechoslovakia

In the course of the experiment, two other laboratories joined this group, the Technische Universität (TUG) - Austria and the Van Swinden Laboratorium (VSL) - Netherlands.

At each site was arranged a measurement setup of the kind shown in Fig. 4. Some technical details about the receiving stations are reported in Table 3 together with the evaluated antennas coordinates.

	LABORATORY					
Parameter	AOS	ASMW	IEN	IREE	TUG	VSL
Latitude	52°16'37.0" N	52°27'14" N	45°00'53.6" N	50°07'53" N	47°04'01.5" N	51°59'58.9" N
Longitude	17°04'23.7" E	13°37'01" E	07°38'20.1" E	14°27'09" E	15°29'35.5" E	04°22'50.7" E
Height	129 m	50 m	297 m	300 m	534 m	60 m
Antenna diameter	-	0.9 m	3 m	3 m	3 m	3 m
LNA - Noise figure	-	1.4 dB	1.9 dB	2 dB	1.4 dB	4.5 dB
IF Bandwidth	-	30 MHz	40 MHz	-	30 MHz	40 MHz
Station time reference	UTC(AOS)	UTC(ASMW)	UTC(IEN)	UTC(TP)	UTC(TUG)	UTC(VSL)

Table 3. Ground stations specifications

Every day two sets of 20 time interval measurements between the local 1 PPS reference and the trailing edge of the first field synchronizing pulse of the RAI UNO video signal were measured starting at 08:15:01 UTC and 20:15:01 UTC in each laboratory.

The measurement results were exchanged weekly in a concise form by telex or by General Electric - Mark III with files named TVXXX* updated each Tuesday. The complete set of measured data were sent monthly to the coordinating laboratory (IEN) by mail. Fig. 5 shows on a monthly base the percentage of synchronization data available for each laboratory.

In two sites, namely IEN and TUG, time differences between the local 1 PPS and the satellite TV signal were also measured every ten minutes from November 1988 to March 1989 to get information about the satellite movement. Moreover, at IEN has also been measured the time difference between the terrestrial television signal received from the Turin-RAI transmitter and the satellite signal.

All the participating laboratories collaborate to the international atomic time scale and are linked to BIPM using either the GPS time link (IEN, TUG, VSL) or the television link (AOS, ASMW, TP).

Satellite position data for November and December 1988 have been supplied by the ESA satellite control station at Redu (Belgium) that is charged of ECS satellites tracking. These data have been used to investigate the satellite movement and to correct the synchronization data.

4 - Measurement results

The evaluation of the results obtained in the ECS synchronization experiment, has been focused on the period November-December 1988 for which ESA-Redu orbital elements data were available. From these data, the variations of ECS 5 in latitude, longitude and range from the origin of the rotating reference frame have been computed and reported in Figure 6 where it can be seen that the variations around their mean values are within ± 16 km (range), ±0.07° (long.), ±0.08°(lat.). Some discontinuities due to orbital manoeuvres can also be observed.

The results of the measurements performed every ten minutes at IEN between the terrestrial television signals and the signals coming from ECS 5 relative to the beginning of November 1988 are reported in Fig. 7 where the diurnal variations due to the satellite movement are shown. Each point of Fig. 7 represents, in microseconds, with a constant value substracted, the sum of the distances of the satellite from Telespazio-Fucino station and from IEN-Torino, versus time. The interruptions in the curves are due to the lack of TV transmissions. For all the computations performed on the measured data whose results are reported in all the next graphs, the linear regression parameters, computed at every laboratory for each series of 20 time interval readings, have been used.

This solution was chosen to reduce the amount of information to be exchanged among the laboratories.

Together with the linear fit parameters, the standard deviation of the residuals was computed after the rejection of the outliers by means of a statistical filter.

In fig. 8 have been reported the time differences between UTC(IEN) and UTC(TP) obtained from ECS 4 and ECS 5(*) synchronization measurements from mid-September 1988 up to February 1989. Two curves have been traced to distinguish the two daily sets of measurements (08:15 and 20:15 UTC). The peak-to-peak variation of each curve, if the mean rate between the two time scales is removed, is of the order of 4 microseconds that is well within the time error estimated previously and reported in Fig. 2.

The results of the time comparisons performed between IEN and the other metrological laboratories (AOS, ASMW, TP-IREE) by means of ECS 5 in November-December 1988 together with those corrected for the ESA satellite position data and the BIPM data, from Circular-T, have been plotted in Fig. 9.

(*) The following ECS satellites were used in the course of this experiment:
ECS 1 from July 88 to mid-Sep. 88 (13°E)
ECS 4 from mid-Sep. 88 to mid-Oct. 88 (10°E)
ECS 5 from mid-Oct. 88 onward (10°E)

From these graphs it can be seen that the peak-to-peak time excursions are reduced to some hundreds of nanoseconds, if one corrects for the differential propagation delay variations, and that the mean rates of the time scales compared are in good agreement with those obtained from BIPM data. As regards the uncertainty of these comparisons, worst-case residuals at the microsecond level have been found.

This is mainly due to the following factors: the uncertainty in the satellite position determination and in the receiving antennas coordinates and the differential delays of the receiving stations, which have not been accounted for. Corrections for the ionospheric and tropospheric propagation delays and for the relativistic effect must also be applied to the synchronization data to improve the results.

In the case of the time comparisons between IEN and TUG, reported in Fig. 10, the curves obtained from ECS 5 (corrected for the satellite position) and the BIPM data are in closer agreement because the antenna coordinates are expressed in the same reference system and because the differential delay of the receiving equipment has been partially evaluated. The differential propagation delays corrections, about 1 ns in the case of the ionosphere and about 5 ns in that of the troposphere, have been neglected. Viceversa the relativistic correction due to Earth rotation (Sagnac effect), amounting to about 30 ns, has been applied to the IEN/TUG results.

In Fig. 11 are plotted the residuals of UTC(IEN)-UTC(TUG) obtained subtracting the ECS synchronization data from the GPS ones. The Allan variance computed on these residuals is reported in Fig. 12.

5 - Conclusions

The results obtained in the first months of the synchronization experiment between some European laboratories, based on the television signals received from ECS 5 geostationary satellite, have shown that an uncertainty ranging from some hundreds of nanoseconds up to one microsecond is achievable over a very large area, provided that the measurements are corrected for the satellite position parameters as determined by an ESA tracking station.

To establish an independent synchronization system relying only on the timing laboratories measurements, it is necessary nevertheless to perform also range measurements at the telecommunication ground station, where the television signals are broadcasted to the satellite, and this approach will be tested in the incoming months.

The possibility to determine the position of ECS 5 with four stations performing both GPS and ECS measurements by means of the time differences between clocks provided by the GPS link, will also be checked. This opportunity will be realized in the near future since a fourth station, ISPT in Rome (Italy), equipped with GPS is going to join this experiment.

A major problem affecting the uncertainty of this kind of time comparisons is the measurement of the receiving stations delay which could be performed by means of a calibration equipment visiting all the participating laboratories.

6 - Acknowledgement

The authors would like to thank Mr. B. Demelenne of ESA, Redu station, Dr. D. Kirchner of TUG, Dr. F. Kalau of ASMW, Dr. P. Fraczyk of AOS and Ing. G. De Jong of VSL for their indispensable cooperation.

References

|1| Tolman J., Ptáček V., Soućek A., Stecher R.: Microsecond Clock Comparison by Means of TV-synchronizing Pulses, IEEE Trans. Instr. vol. IM-16, pp. 247-254, Sept. 1967.

|2| Ptáček V.: The Long-Term Stability of Time Comparisons using the TV Method, Bull. Astron. Inst. Czechosl. 29/1978, pp. 372-375, No 6, June 1978.

|3| Kirchner D., Ressler H., Cordara F., Galliano P.G., Pettiti V.: A comparison of Time and Frequency measurements via OTS-2 with results obtained by NAVSTAR/GPS and Loran-C, IEEE Trans. Instr. vol. IM-34, No 2, June 1985.

|4| De Jong G., Kaarls R.: Time Synchronization via OTS-2, BCR Information, Applied Metrology, Report EUR 10554 EN, 1986.

|5| Kawase S.: Closed dynamical One-Way Time-Synchronization using geostationary Satellite, Journal of Radio Research Laboratory, vol. 34, No 142, pp. 43-53, June 1987.

|6| Banerjee P., Mithlesh Saxena, and Mathur B.S.: Passiva TV Technique for Time Transfer via Indian Satellite INSAT-1B, IEEE Trans. Instr. and Meas. vol. IM-36, No 2, June 1987, pp. 579-583.

|7| Kovačević B.Z., Dimitrijević B.R., Arsić M.Z., Kovacevic N.S.: Precise Real time signal dissemination over the TV broadcasting satellite, Radio Science, vol. 14, No 4, July-August 1979.

|8| Perek L.: The scientific and technical Aspects of the geostationary Orbit, Acta Astronautica, vol. 17, No 6, pp. 589-598, 1988.

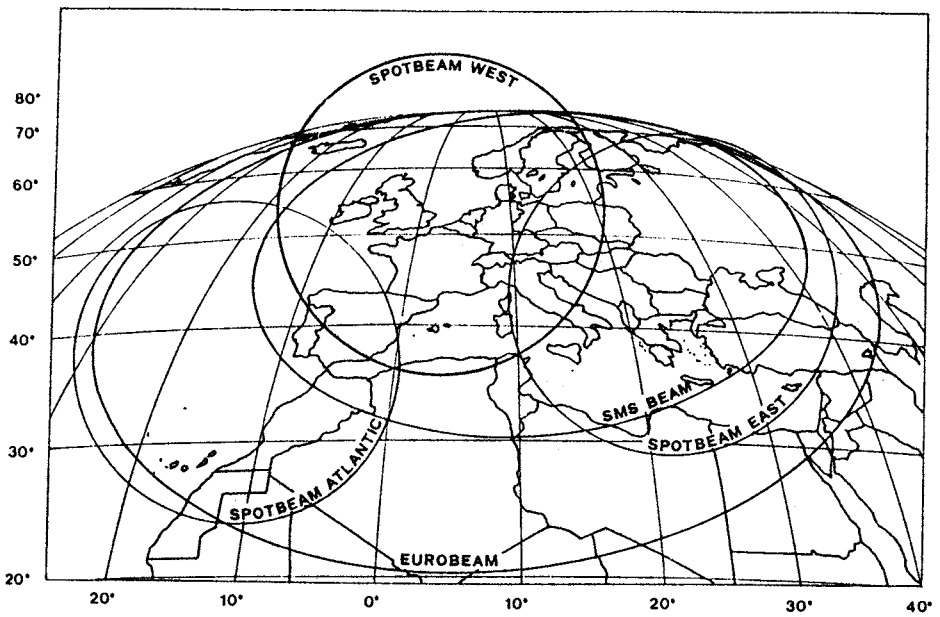


Fig. 1 - ECS satellites coverages.

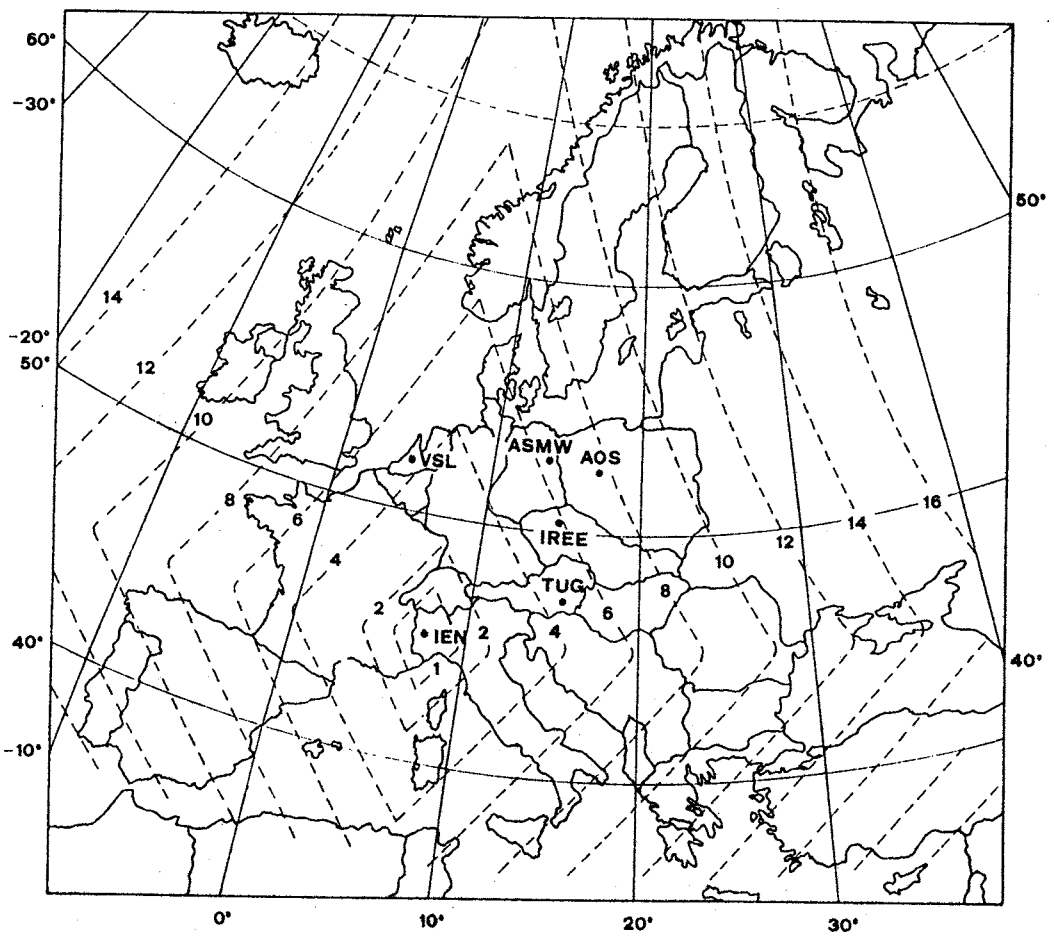


Fig. 2 - Map of the time error contours (in microseconds)

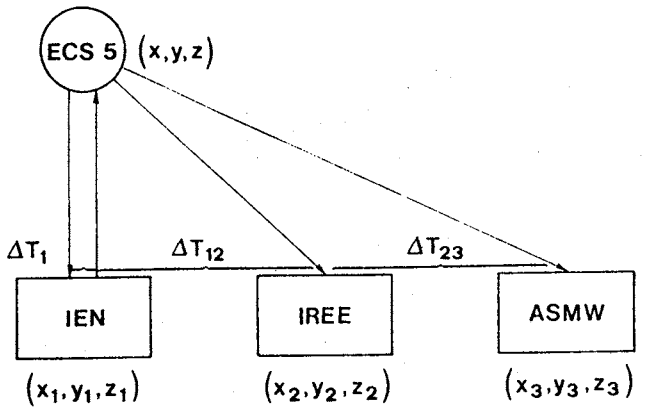


Fig. 3 - Schematic view of the experimental configuration

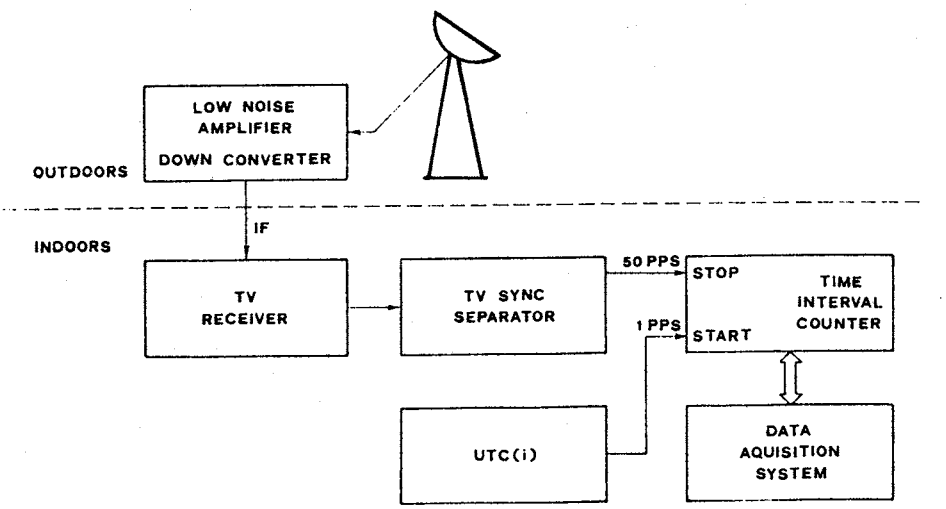


Fig. 4 - Measurement setup

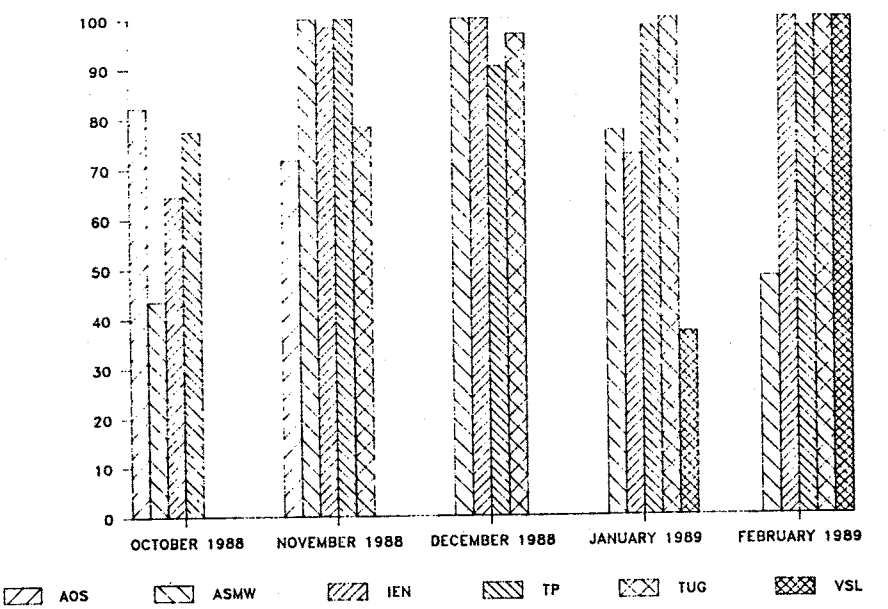
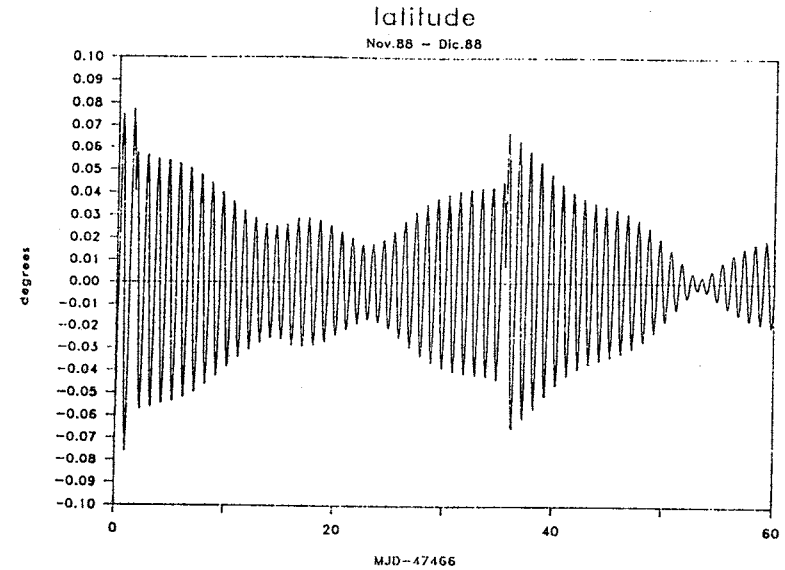
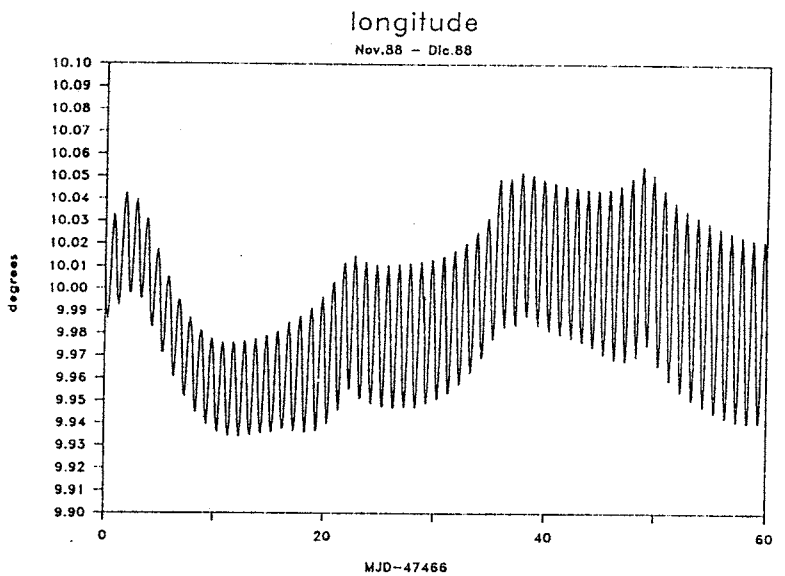


Fig. 5 - Data supplied by participating laboratories





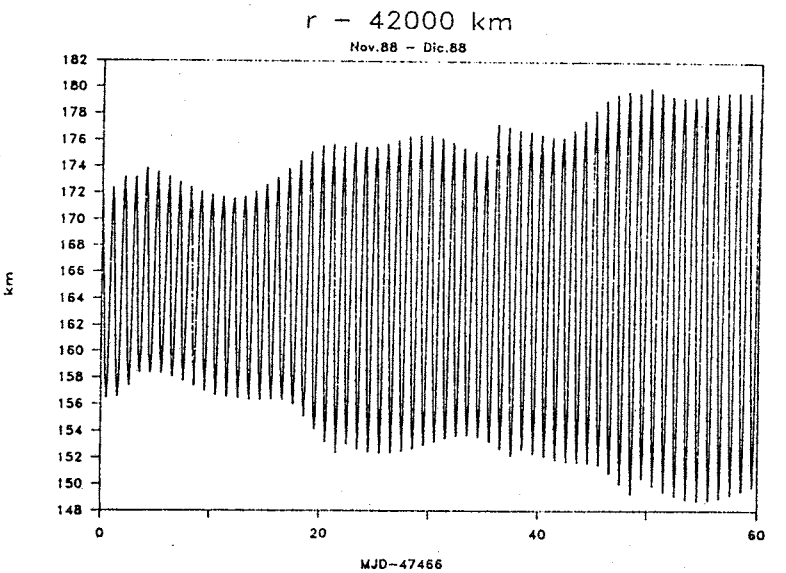


Fig. 6 - ECS 5 position from ESA-Redu data

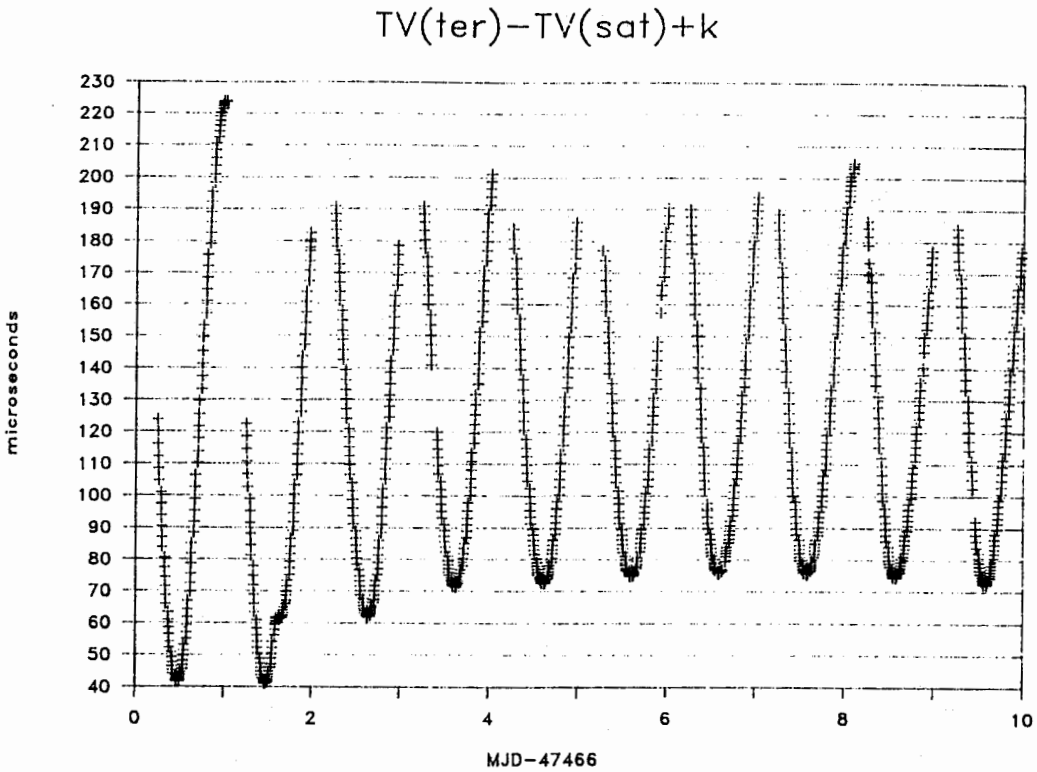


Fig. 7 - Ten-minutes time differences measured at IEN between the terrestrial and the ECS TV signal

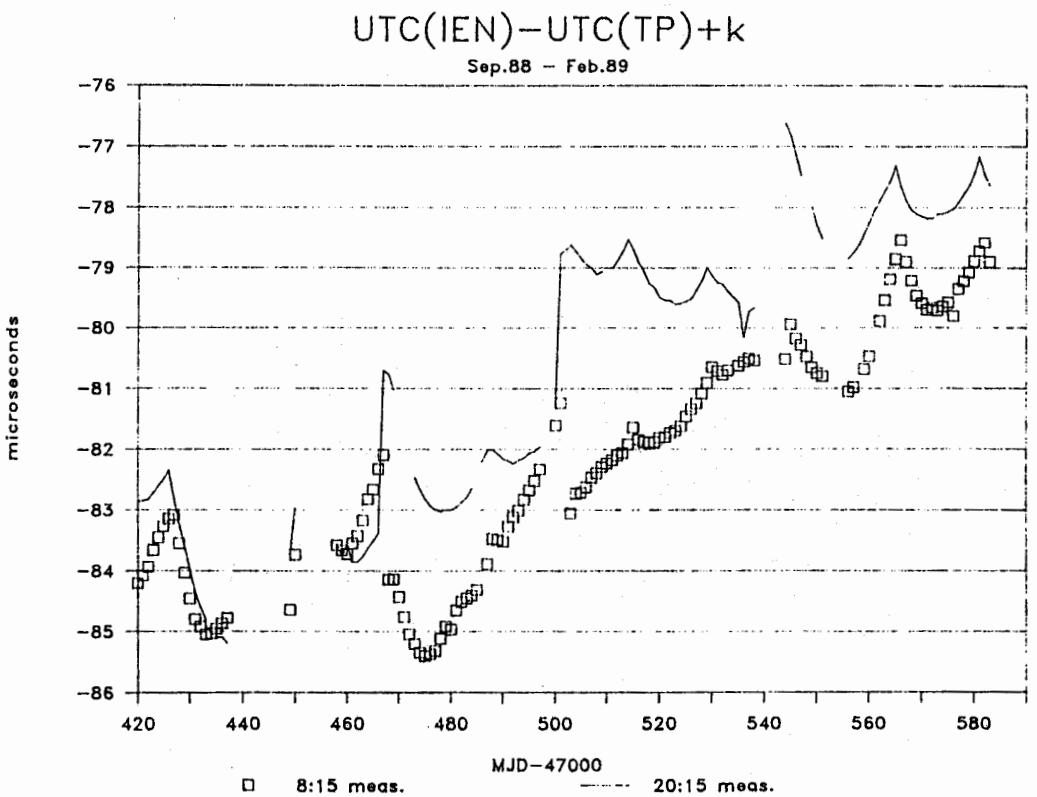
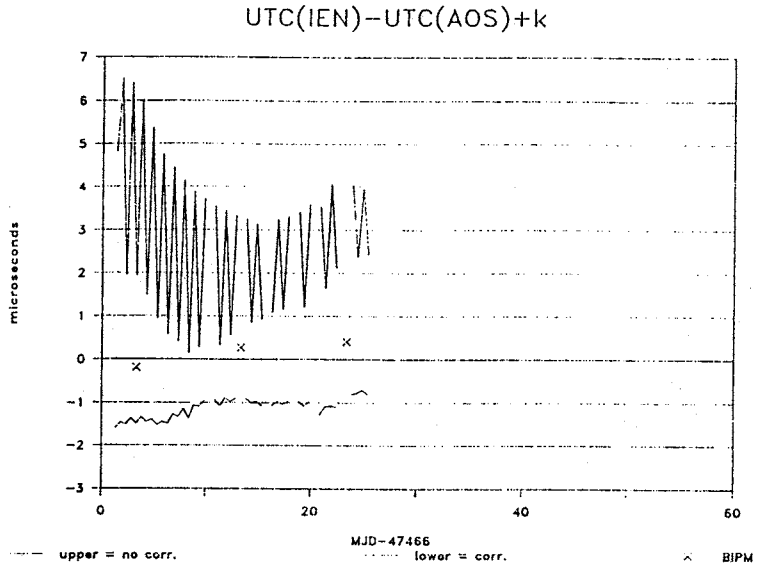
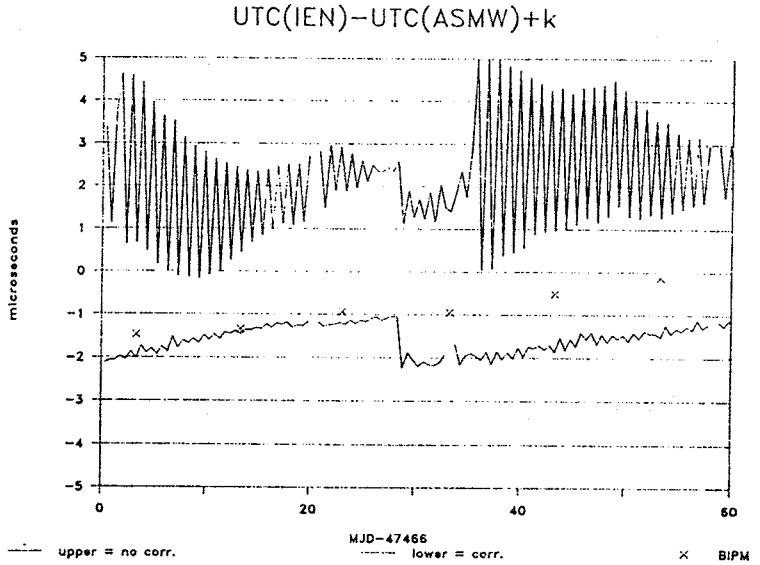


Fig. 8 - Time differences between IEN and IREE from ECS 5 satellite





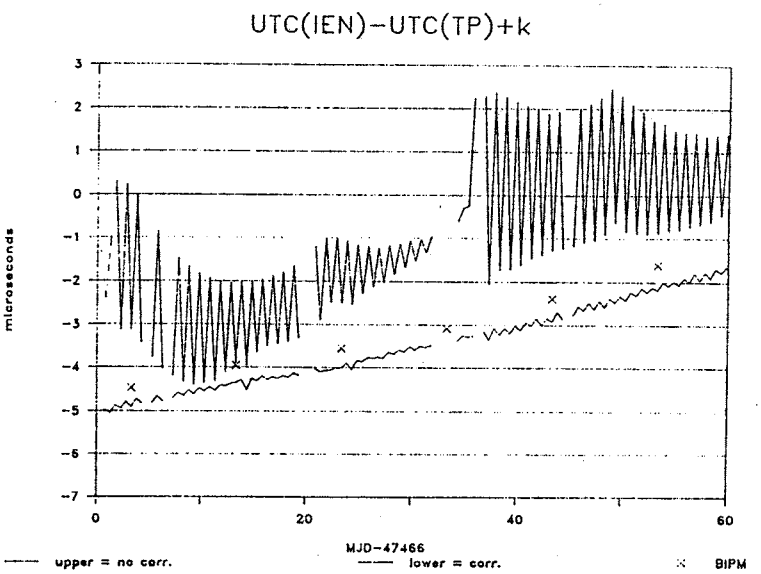


Fig. 9 - Time scales comparisons by means of ECS 5 satellite and BIPM circulars T November-December 1988

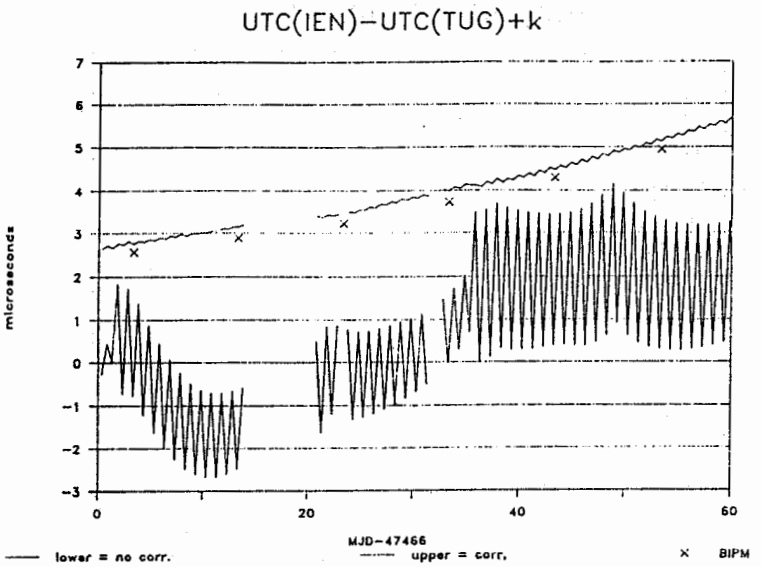


Fig. 10 - UTC(IEN) vs. UTC(TUG) from ECS 5 and BIPM circular T with TUG time scale step compensated

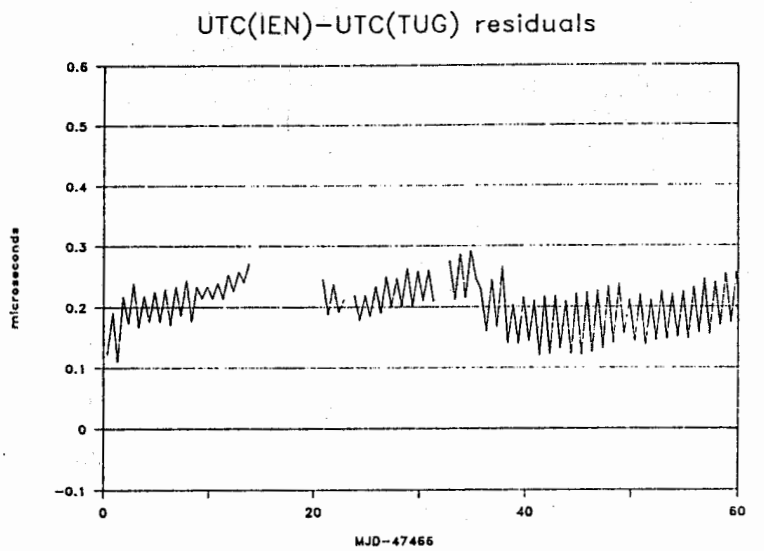


Fig. 11 - Residuals of UTC(IEN) - UTC(TUG) via GPS and ECS 5 - November-December 1988

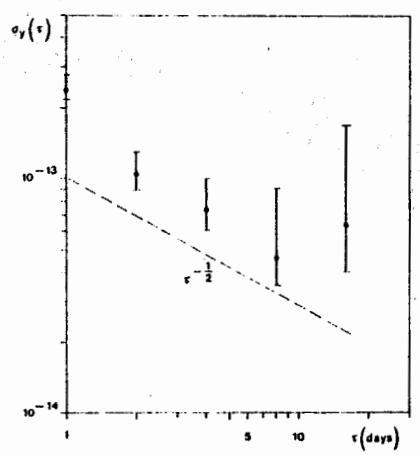


Fig. 12 - Allan Variance of the residuals obtained from GPS and ECS 5 data

FREQUENCY STABILITY AND SPECTRAL PURITY REQUIREMENTS FOR ESA'S DATA RELAY SATELLITE

Laurent-Guy BERNIER * † and Giovanni BUSCA *

* Observatoire Cantonal de Neuchâtel, Switzerland
† Formerly with Oscilloquartz SA, Neuchâtel, Switzerland

ABSTRACT

The Data Relay System (DRS) is based on 2 geostationary satellites that shall provide data-relay services between the client Low Earth Orbiting (LEO) satellites and the ground stations. In addition to the data-relay services, the DRS will provide a LEO localization service using the DRS as a ranging-relay. Thus the reference quartz oscillator on board the DRS must generate a signal with both an excellent frequency stability, for ranging purposes, and an excellent spectral purity for the generation of various carriers at microwave frequencies. This paper describes a preliminary work on several basic issues relevant to the DRS design from the point of view of frequency stability and spectral purity.

1.0 INTRODUCTION

The purpose of ESA's Data Relay System (DRS) is to provide the communication, telemetry and tracking links, necessary for mission control, between Earth terminals and Low Earth Orbit (LEO) spacecrafts using only two geostationary Data Relay Satellites (DRSS) instead of the large number of Earth tracking stations presently required. Each DRSS will carry a reference oscillator embedded into a Frequency Generation Unit (FGU) which purpose is to generate the carriers required by the different frequency trans-lations performed inside the up-link and down-link channels. The frequency stability of the reference oscillator determines the accuracy of the tracking operations performed through the DRS while its spectral purity determines the signal to noise ratio in the communication channels. This paper reports preliminary work that covers several fundamental aspects of the frequency stability and spectral purity issues involved in the DRS.

2.0 DRS OVERVIEW

The DRSS are planned to be located at 59° E and 44° W. The corresponding Earth coverage is illustrated on figure 1. A LEO spacecraft, defined as having an altitude lower than 1000 km, is in view of a single Earth station at most 10% of the time. On the other hand, the same LEO spacecraft is in view of one of the two DRSS more than 85% of the time. Another advantage of the DRS is that a complete Earth coverage is possible with all ground stations located in Europe instead of being scattered all other the world. The part of the up and down links that goes between the LEO and the DRSS is called the Inter Orbit Link (LEO) and uses either the k or the s band. The part that goes between a ground station and a DRSS is called the feeder link and uses only the k band because of the atmospheric interface. This is illustrated on figure 2.

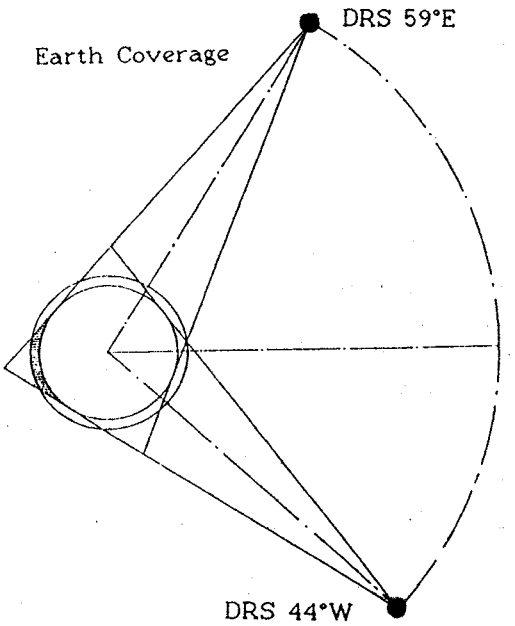


Fig. 1 Planned Earth Coverage with 2 DRSS

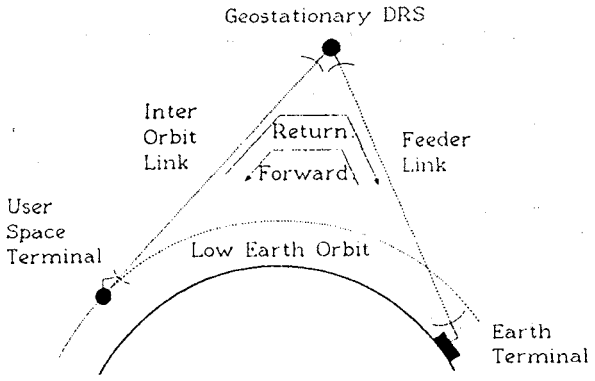


Fig. 2 Definition of Feeder and Inter-Orbit links

3.0 DRS MAIN SERVICES

The DRS will provide forward and return communication links between User Earth Terminals (UET) located on ground and User Space Terminals (UST) located on LEO or ballistic vehicles. These vehicles will include scientific and Earth observation satellites as well as sub-orbital spacecrafts such as Hermes and Ariane. Communication data will include telecommand, telemetry, payload data, voice and video. The DRS will also provide a localization service to the client satellites. Moreover TT&C Earth stations will provide services internal to the DRS such as telemetry, telecommand and ranging.

4.0 FREQUENCY STABILITY REQUIREMENTS

The DRS will use a FDM access scheme. Thus timing is not a critical issue. The frequency stability of the reference oscillators will influence mainly the precision of range and range rate measurements required for both internal and user localization services.

4.1 *Frequency stability and range measurements*

Range measurements are made by measuring the delay of a two-way transmission between the source and the target. The precision is limited by the integrated random phase fluctuations of the reference oscillator that limits its use as a clock over the time interval of the two-way transmission. The range uncertainty Δx is given by

$$\Delta x = \frac{c}{2} TIE(\tau) \qquad (1)$$

where c is the speed of light and $TIE(\tau)$ the Time Interval Error i.e. the time error accumulated during the two-way transmission interval τ. Assuming that the initial frequency offset is zero, which is true by definition since the reference oscillator is compared against itself, the TIE can be shown to be given by [1]

$$TIE^2(\tau) = 2\tau^2\sigma_y^2(\tau) \qquad (2)$$

where $\sigma_y(\tau)$ is the square root of the Allan variance. Thus if the frequency stability contribution to the range measurement uncertainty is required to be less than 5 [m] then the maximum acceptable TIE is 33 [ns] and the required frequency stability is $\sigma_y(\tau)$ = 5E-8 over the τ = 0.4 [s] two-way interval. If the DRSS reference oscillator is free running with respect to the ground station reference then the total uncertainty is obtained by adding, in the mean squares sense, the individual uncertainties. The required performance is easily achieved in a quartz oscillator.

4.2 *Frequency stability and range-rate measurements*

Range rate measurements are obtained by measuring the Doppler shift between the source and target. They reduce to a frequency difference measurement after a two-way transmission and their precision is limited by the variation of the reference frequency over the two-way transmission interval. The range rate error Δv is given by

$$\Delta v = \frac{c}{2} \Delta(\tau)\{y(t, T)\} \qquad (3)$$

where c is the speed of light and y(t) the normalized reference frequency.

$$\Delta(\tau)\{y(t)\} \equiv y(t) - y(t-\tau) \qquad (4)$$

defines the first increment operator, in this case the frequency difference over the two-way interval τ while

$$y(t,T) \equiv \frac{1}{T}\int_{t-T}^{t} y(x)\,dx \qquad (5)$$

defines the moving average operator which describes the averaging process that takes place inside the frequency counter.

Thus the range rate measurement uncertainty (3) may be interpreted as the result of the random difference between the reference frequency evaluated over the averaging interval T before and after the two-way transmission interval τ. If we choose to average the frequency over an averaging interval equal to the two-way transmission interval then we have $T \equiv \tau$ and the range rate uncertainty becomes

$$\Delta v_{rms} = c\,\sigma_y(\tau) \qquad (6)$$

since the Allan variance may be defined as

$$\sigma_y^2(\tau) \equiv \tfrac{1}{2}\,E\{\,[\Delta(\tau)\{y(t,\tau)\}\,]^2\,\}. \qquad (7)$$

Thus if the frequency stability contribution to the range-rate uncertainty is required to be less than 5 [mm/s] then the required frequency stability is $\sigma_y(\tau)$ = 1.7E-11 over τ = 0.4 [s]. This performance is achievable in a quartz oscillator.

4.3 *Cancellation of the DRSS residual Doppler shift*

If the reference oscillator on the DRSS if free running with respect to the reference oscillator on ground, then a beacon signal transmitted from the DRSS to ground is needed in order to cancel the unknown DRSS reference frequency in the range-rate receiver.

However, a single beacon signal does not cancel out the Doppler shift produced by the residual range-rate of the geostationary DRSS with respect to ground. If the maximum value of this residual range-rate is negligible with respect to the required range-rate uncertainty, then further Doppler cancellation is not necessary. If it is not negligible, then two beacon signals must be used; a direct beacon transmitted from the DRSS to ground and an indirect beacon transmitted from the DRSS to the LEO and then back to the ground through the DRSS.

Figure 3 shows the signal path between the UET, the DRSS and the UST. On the base of this diagram the return frequency f_2 can be computed as a function of the transmitted frequency f_1 as follows:

$$f_2(t) =$$

$$f_1(t-2\tau_1-2\tau_2)\times\left(\frac{x}{y}\right)\times\left(1+\frac{V_{sr}}{c}\right)^2\times\left(1+\frac{V_{er}}{c}\right)^2$$

$$+$$

$$L\times f_0(t-\tau_1-2\tau_2)\times\left(\frac{x}{y}\right)\times\left(1+\frac{V_{sr}}{c}\right)^2\times\left(1+\frac{V_{er}}{c}\right) \qquad (8)$$

$$-$$

$$M\times f_0(t-\tau_1)\times\left(1+\frac{V_{er}}{c}\right)$$

where τ_1 is the transmission delay between Earth and the DRSS while τ_2 is the delay between the DRSS and the LEO vehicle. Both delays are about the same thus the total two-way delay is approximately 4 times a single Earth-DRSS delay i.e. $\simeq$ 0.4 [s]. V_{sr} is the range-rate of the LEO vehicle with respect to the DRSS while V_{er} is the range-rate of the DRSS with respect to Earth. L and M are the multiplication factors for the frequency translations in the up and down links, x/y is the coherent repeater factor in the UST and f_0 is the frequency of the DRSS reference oscillator.

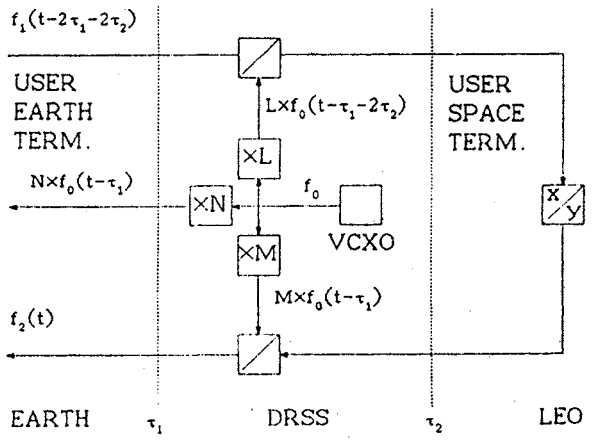


Fig. 3 DRS Signal Path

Equation (8) shows that the information about the range-rates is distributed into 3 terms. Only the first term contains the wanted Doppler shifts with respect to the up-link frequency f_1. The two other terms, however, contain Doppler shifts with respect to the unknown DRSS reference frequency f_0 and must be removed in order to resolve the range-rate of the DRSS with respect to Earth. It appears that a direct beacon is sufficient to cancel out the second term but that the indirect beacon defined above is necessary in order to cancel out the third term.

5.0 SPECTRAL PURITY REQUIREMENTS

Figure 4 illustrates the phase noise analysis of the carrier generation in the k band. Assuming the phase noise specification shown on the figure and that the carrier-to-noise ratio of an hypothetical up-link pilot is 60 dB, it appears that it is not possible to improve the spectral purity of the DRSS reference oscillator by phase locking it to an up-link pilot generated on ground. Therefore the DRSS reference quartz oscillator must have by itself a spectral purity sufficient for the local synthesis of all the carriers required in the satellite operation. The choice of the quartz oscillator nominal frequency, which is also the synthesis start frequency, is a compromise between the close-in phase noise, which is likely to be best if the nominal quartz frequency is chosen in the range of a few MHz, and the white phase noise floor which can be improved by increasing as much as possible the synthesis start frequency.

6.0 CONCLUSION

The DRS system parameters are not fully specified yet. The purpose of this paper was to explain the basic design issues relevant to the point of view of frequency stability. The main conclusion is that the frequency stability requirements for the DRSS reference generator, relevant to the range and range rate measurements, are not very difficult to realize.

On the other hand, the spectral purity requirements are likely to be much more difficult to realize and rely entirely on the DRSS reference quartz oscillator for Fourier frequencies smaller than the PLL bandwidth of the Tunable Frequency Converters (TFC) used inside the FGU. For higher Fourier frequencies, on the other hand, the spectral purity will be determined by the phase noise of the VCOs used in the TFCs.

The decision about the use of one or more beacons, finally, depends upon the range-rate resolution that is to be specified. If the residual range-rate of the DRSS with respect to ground is negligible as compared to the wanted resolution, then it makes sense to leave the reference oscillator free running and to transmit its frequency to ground using a single beacon. If, on the other hand, the residual DRSS Doppler is to be removed, then it is probably much better to lock the DRSS reference oscillator to the up-link frequency rather than to transmit two different beacons.

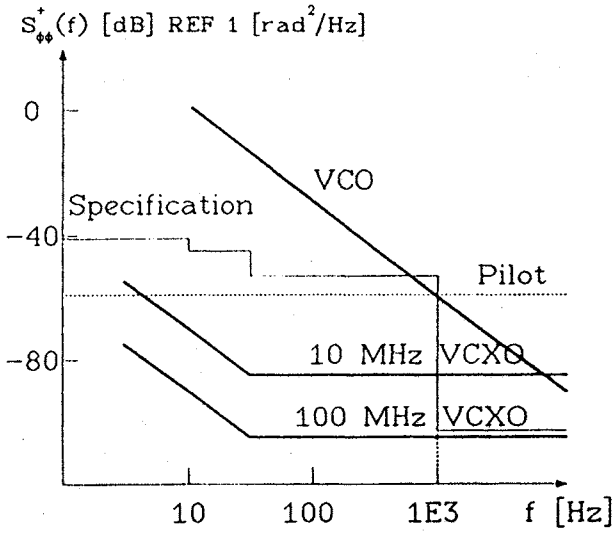


Fig. 4 Phase Noise Analysis in the k Band

REFERENCES

[1] Bernier L.G., "Linear Prediction of the Non-Stationary Clock Error Function", *Proc. 2nd European Frequency and Time Forum*, Neuchâtel, March 1988, pp. 125-137.

[2] De Agostini A., "Data-Relay Satellite (DRS) Ground-Segment Design", ESA Journal, vol. 12, 1988, pp. 249-266.

[3] "Data Relay System (DRS) System Performance Specification", ESA document SP-7-0-1 issue N° 3, June 1988.

Acknowledgements

The work reported in this paper was performed for ESA at Oscilloquartz, Switzerland, under a contract with Marconi Space Systems, U.K.

SYNCHRONISATION VIA LASER TECHNIQUES
THE LASSO EXPERIMENT

Jean Gaignebet
Groupe de Recherches de Géodésie Spatiale
Grasse - France

Sigfrido Leschiutta
Politecnico di Torino - Elettronica
Torino - Italy

The LASSO experiment is an attempt to explore the capabilities and the operational problems of a new time comparison method that is based on the bouncing of optical pulses from a spacecraft, suitably fitted with retroreflectors.

The host satellite, the ESA Meteosat-2 meteorological satellite, was succesfully launched on June 1989, and after a period devoted to the placement of the satellite to its station and the commissioning procedures, the experiment was started with the first operations of laser ranging, needed in order to obtain the precise orbit needed for the subsequent timing tests.

After a summary of the technical characteristics of the experiment, some news are given about the first ranging results and the relevant accuracies, as obtained with the Lunar Laser ranging station of CERGA.

3RD EUROPEAN FREQUENCY AND TIME FORUM - Besançon - March 21-23, 1989

RESONATEURS A ONDES ELASTIQUES DE VOLUME UTILISANT DES MATERIAUX AMORPHES

par J. Beaussier, M. Breuzet, J.C. Craveur, M. Gay et J. Uebersfeld

Office National d'Etudes et de Recherches Aérospatiales,
BP 72, 92322 Châtillon Cedex, France

BULK ACOUSTIC WAVES RESONATORS MADE OF AMORPHOUS MATERIALS

L'expérience acquise à l'ONERA dans le domaine de l'action et de la détection capacitive a permis de réaliser des résonateurs de forme parallélépipédique dans des matériaux amorphes conducteurs ou isolants. Les meilleurs résultats ont été obtenus avec le mode particulier dit "mode de Lamé" dont la description fait l'objet de la première partie de l'exposé.

La seconde partie traite des recherches effectuées pour obtenir un coefficient de surtension élevé. En l'absence de précautions particulières, les pertes par les supports constituent la source principale de dissipation d'énergie. Il s'est avéré nécessaire de définir un type de fixation qui minimise le couplage support-élément vibrant. Des méthodes de calcul par éléments finis ont été dans ce but largement utilisées.

Une troisième partie résume les résultats expérimentaux obtenus sur des résonateurs en AU 4G (Duralumin) et en silice vitreuse. Dans ce dernier cas, le coefficient de surtension atteint 4.10^6 à 200 kHz et à température ambiante. Un balayage thermique a été effectué depuis des températures cryogéniques de quelques degrés Kelvin jusqu'au delà de 100°C. Les variations de la fréquence et du coefficient de surtension ont été mesurées.

Abstract

The ONERA's experience in the field of electrostatic technique applied to capacitive sensors and actuators led to design parallelepipedic resonators made of conducting or insulating amorphous materials. The best results have been obtained with a particular mode of vibration called "mode of Lamé", which is described in the first part of this paper.

The second part deals with the improvement of the quality factor. Without any particular precautions, energy dissipation is mainly due to losses in the supports. It is necessary to define a type of resonator-mounting that reduces the coupling factor between the vibrating mass and the supports. For that, finite element calculations have been widely used.

The last part summarizes the experimental results obtained with AU 4G and SiO_2 resonators. With the latter, a quality factor of 4×10^6 at 200 kHz and room temperature has been achieved.

Finally, frequency and quality factor have been measured for temperature ranging from a few Kelvin up to 400° K.

Introduction

La plupart des résonateurs équipant les oscillateurs à fréquences radio sont en quartz piezoélectrique. Un certain nombre de limitations propres à ce matériau sont maintenant bien connues. Pour améliorer encore les performances des oscillateurs, des matériaux nouveaux sont étudiés, qu'ils soient ou non piezoélectriques. Bénéficiant de l'expérience et des moyens acquis à l'ONERA en matière de détecteurs capacitifs et de générateurs de forces électrostatiques, des résonateurs en matériau non piezoelectrique conducteurs ou isolants à action et détection capacitives, ont été mis au point à l'Office. Parmi tous les modes qui ont été expérimentés, un mode particulier dit "de Lamé" est apparu particulièrement intéressant.

Modes de vibration dans les plaques

- Le cisaillement d'épaisseur (figure 1) est le mode le plus utilisé pour les résonateurs piezoélectriques équipant les oscillateurs. Le plan médian, à équidistance des plans limitant la plaque, est un plan nodal pour le mode fondamental de cisaillement d'épaisseur. Pour exciter ce mode par méthode capacitive, il faut placer une électrode entre le plan médian et la surface de la plaque, ce qui devient délicat à faible épaisseur.

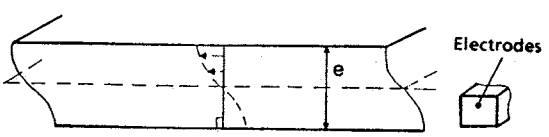


Fig. 1 - Cisaillement d'épaisseur en mode fondamental.

- Les modes de flexion (figure 2) ne permettent pas en général l'obtention d'un coefficient de surtension suffisant pour l'application "oscillateur" envisagée. En effet, la flexion crée des variations locales de volume engendrant des pertes thermo-élastiques importantes qui limitent le facteur de qualité. On montre que pour réduire ces pertes, il faut augmenter la fréquence du résonateur, donc diminuer ses dimensions. La fixation du résonateur qui nécessite un positionnement très précis sous peine d'avoir un mauvais découplage et un effondrement du coefficient de qualité pose alors un problème.

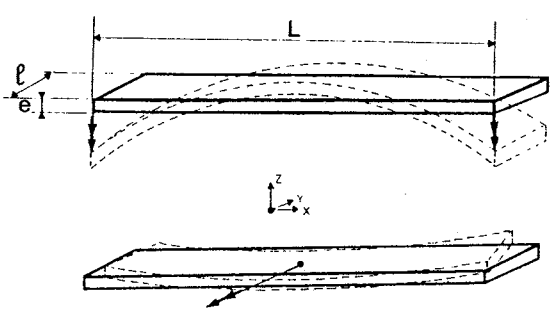


Fig. 2 - Mode de flexion (partiel 1) d'une poutre libre-libre
- suivant l'épaisseur
- suivant la largeur.

- Les modes d'extension-compression, créant des variations locales de volume, posent le même problème que les modes de flexion.

Dans les plaques rectangulaires pour lesquelles le rapport longueur/largeur n'est pas trop important (< 5), apparaissent des modes dits "de contour". Ces modes sont caractérisés par des déformations dans le plan de la plaque entraînant une déformation du contour de celle-ci. Ces modes, ne s'accompagnant pas de variation locale de volume, ont un coefficient de qualité intrinsèque élevé. Ces modes de contour ne sont en général pas calculables analytiquement et diverses méthodes sont utilisées ici pour leur détermination : des calculs par éléments finis (code Astronef de l'ONERA) et la méthode de Rayleigh-Ritz (principe variationnel) (figure 3).

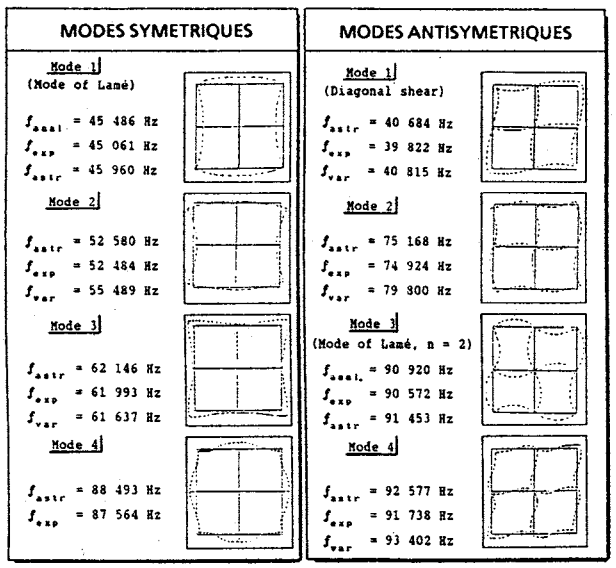


Fig. 3 - Modes de contour d'une plaque carrée.

Fréquences

f_{anal} = analytique (mode de Lamé uniquement)
f_{var} = méthode variationnelle
f_{astr} = calcul par éléments finis
f_{exp} = mesure expérimentale

Dans le cas d'un résonateur carré ou rectangulaire avec L/ℓ entier, il existe une solution analytique exacte satisfaisant l'équation de propagation et les conditions aux limites : c'est le mode inventé par le physicien Lamé en 1852 (figure 4). Il s'est avéré particulièrement bien adapté aux résonateurs non piézoélectriques pour 5 raisons essentielles :

- son coefficient de qualité intrinsèque élevé,
- sa configuration nodale permettant une fixation aisée (les 5 noeuds de vibration sont les 4 sommets de la plaque et son centre),
- sa facilité d'excitation par méthode capacitive,
- sa parfaite connaissance analytique,
- la simplicité géométrique du résonateur (plaque carrée ou rectangulaire).

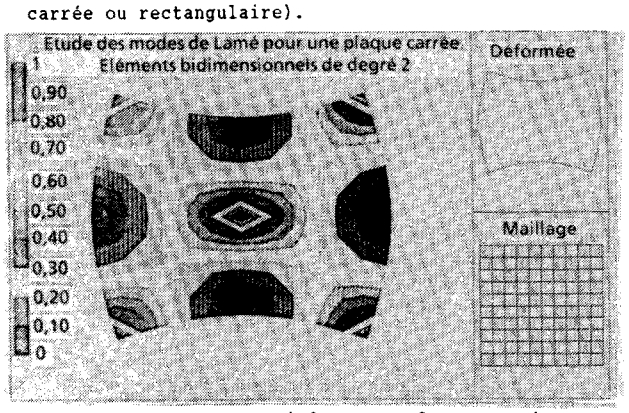


Fig. 4 - Mode de Lamé dans une plaque carrée.

Coefficient de qualité et pertes dans les résonateurs

Les pertes dans les résonateurs à ondes acoustiques ont trois origines distinctes :

- les pertes par la surface, dues par exemple à des impuretés ou à des défauts superficiels, à la métallisation indispensable pour les matériaux non conducteurs ou à la transmission d'énergie dans le gaz environnant. Le coefficient de surtension qui correspond à ces pertes sera noté, Q_s.
- les pertes dans le volume, dues à des dissipations intrinsèques au matériau : Q_i.
- les pertes par les fixations : Q_f.

Le coefficient de qualité global du résonateur est alors donné par la relation :

$$\frac{1}{Q} = \frac{1}{Q_s} + \frac{1}{Q_i} + \frac{1}{Q_f}$$

Les pertes par la surface deviennent négligeables dès que la métallisation éventuelle est limitée à une surface minimale, que le résonateur subit un traitement chimique approprié et qu'il est placé dans une enceinte sous un bon vide secondaire.

Les pertes intrinsèques peuvent être minimisées en sélectionnant un matériau de bonne qualité soumis à un traitement thermique spécifique.

La limitation principale du coefficient de qualité du résonateur est alors due à son montage. Les fixations transmettent de l'énergie vibratoire au support et cet effet est d'autant plus important que le coefficient de qualité intrinsèque est élevé : il est donc primordial d'étudier leur influence et de minimiser les pertes qu'elles induisent. Les calculs par éléments finis ont été largement utilisés pour cette partie de l'étude.

Fixations

Diverses configurations de supports ont été étudiées pour un résonateur carré en mode de Lamé, les sommets du résonateur étant utilisés comme points de fixation (figure 5).

Pour les configurations 1 et 2, la masse du support est plus importante que celle du résonateur ; c'est l'inverse pour la configuration 3.

Pour la configuration 1, il existe un mode de vibration du support proche du mode de Lamé $\left(\frac{\Delta f}{f} = 4\ \%\right)$: il y a couplage entre ces modes (figure 6). Pour la configuration 2, l'écart relatif de fréquence est passé à 6 % et le couplage est moins important que dans le cas précédent.

Pour augmenter cet écart de fréquence, le support est choisi 3 fois plus léger que le résonateur (configuration 3) : l'écart relatif de fréquence est alors de 24 % malgré un couplage aussi important que dans la configuration 1.

Il existe donc un couplage dynamique entre le mode de Lamé et les modes des pattes de fixation. Ces dernières jouent le rôle de filtre mécanique et il est donc impératif de veiller à ce que la fréquence du résonateur soit suffisamment éloignée des fréquences de la patte de fixation pour qu'il n'y ait pas de couplage.

Les déformées obtenues par le calcul par éléments finis montrent que les pattes de fixation sont sollicitées en flexion avec pour conditions aux limites un encastrement du côté support et une articulation du

côté résonateur. Pour que la fréquence de Lamé soit inférieure à la fréquence de flexion de la patte, la condition à respecter est :

$$\frac{e}{\ell_1^2} > 1{,}4117\, f_{\text{lamé}} \sqrt{\frac{\rho}{E}}$$

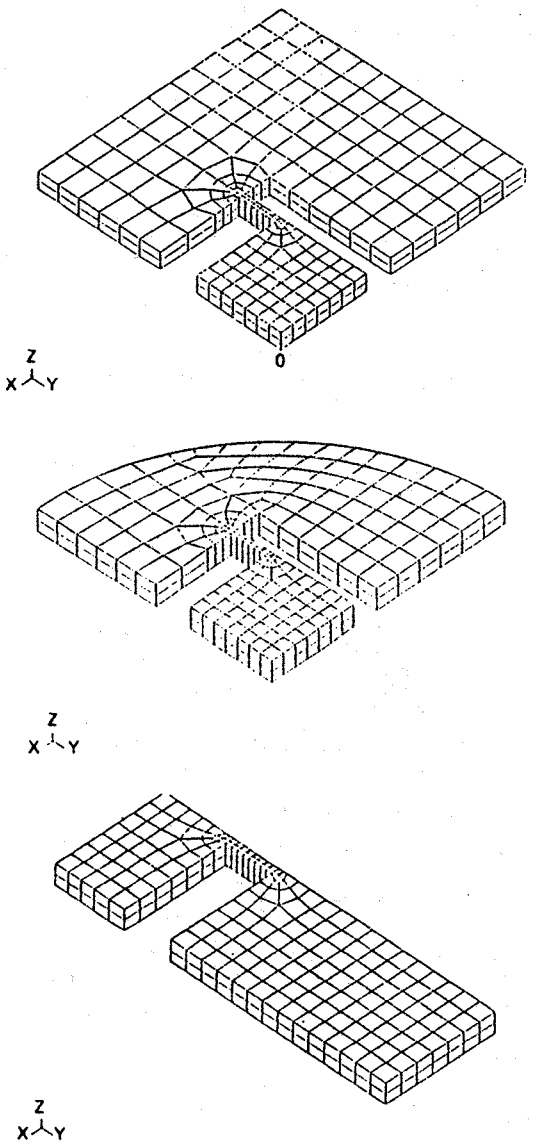
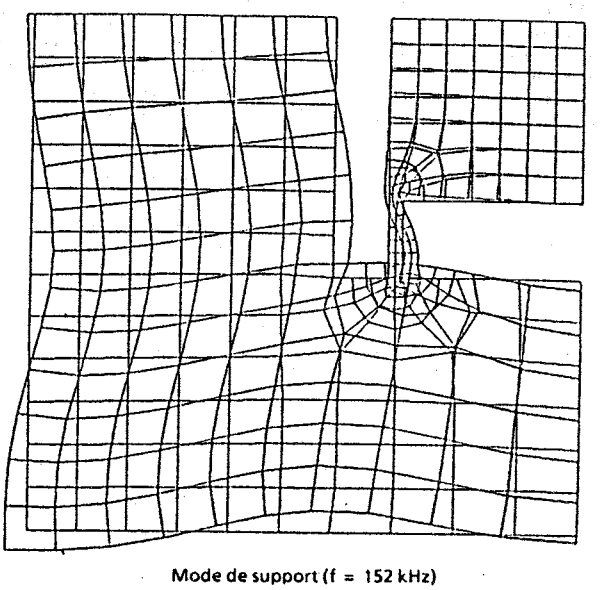


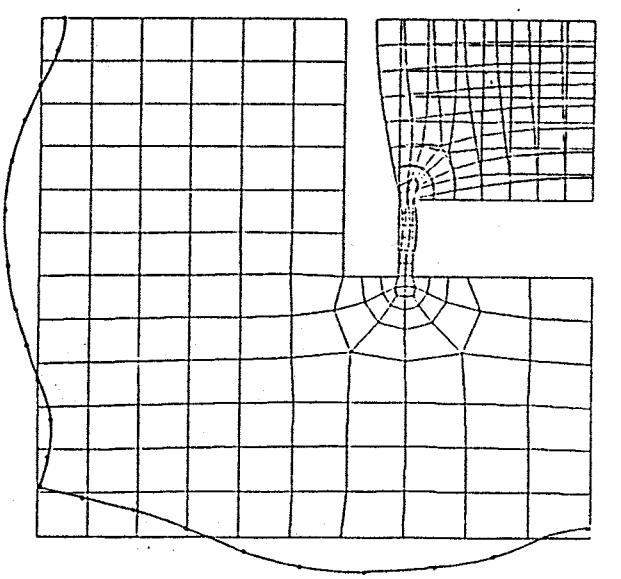
Fig. 5 - Diverses configurations de support.

Seule une partie de la structure est modélisée pour le calcul

Configuration 1
Configuration 2
Configuration 3



Mode de support (f = 152 kHz)



Mode de Lamé (f = 159 kHz)
avec déformation fortement
amplifiée du contour du support

Fig. 6 - Couplage mode de Lamé - Mode de support

Mode de support (f = 152 kHz)

Mode de Lamé (f = 159 kHz) avec déformation fortement amplifiée du contour du support

On montre d'autre part, que pour minimiser les efforts transmis au support, il faut réduire l'épaisseur e des pattes de fixation. Pour qu'il n'y ait pas de couplage dynamique, il est nécessaire de diminuer l'épaisseur des pattes (en tenant compte des limitations technologiques) et de calculer la longueur ℓ_1 des pattes par la formule ci-dessus (figure 7).

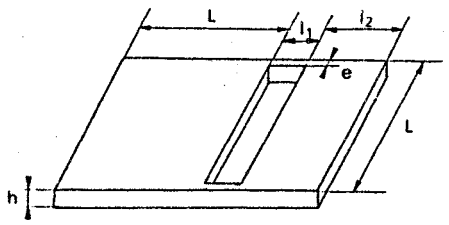


Fig. 7 - Mode de Lamé

Configuration optimisée

$$\begin{cases} \dfrac{e}{\ell_1^2} > 1{,}4117\, f_{Lamé} \sqrt{\dfrac{\rho}{E}} \\ e \text{ minimal} \end{cases}$$

Les résonateurs calculés dans ces conditions puis réalisés et essayés ont tous donné entière satisfaction en ce qui concerne leur comportement dynamique (figure 8).

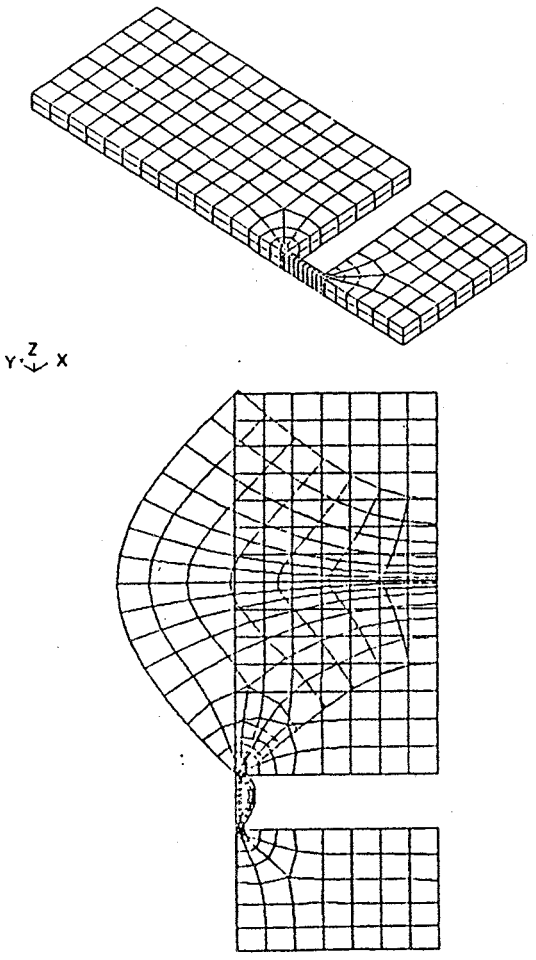


Fig. 8 - Résonateur optimisé : maillage et mode de Lamé.

Etude expérimentale

Des résonateurs vibrant sur le mode de Lamé ont été réalisés en alliage 2017 (appellation officielle de l'AU 4G connu également sous le nom de Duralumin) et en silice vitreuse SiO_2.

L'étude la plus complète a été menée sur les résonateurs en 2017. Ce matériau présente l'avantage d'être facilement usinable par des moyens classiques, et d'être métallique ce qui permet l'excitation directe par moyen capacitif sans électrodes. Un dispositif cryogénique a en outre permis l'étude des résonateurs entre la température ambiante et quelques degrés Kelvin.

Les résonateurs (2017 ou SiO_2) sont mis en vibration forcée par une force capacitive extérieure et l'effet de cette force est mesuré par un circuit détecteur séparé. On mesure la fréquence de résonance en faisant coïncider, par balayage, la fréquence de la tension d'excitation et la fréquence propre de résonateur : la tension détectée est alors maximale.

Le facteur de qualité est obtenu à partir de la décroissance de l'amplitude des vibrations après coupure de l'excitation.

Sur les figures 9 et 10, sont représentés d'une part, le schéma de disposition des électrodes, d'autre part, le montage expérimental réalisé pour cette étude.

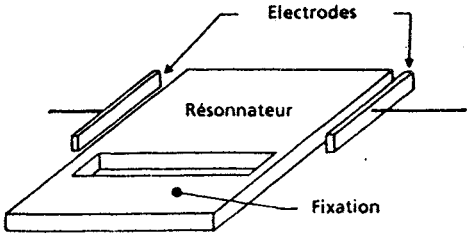


Fig. 9 - Disposition des électrodes.

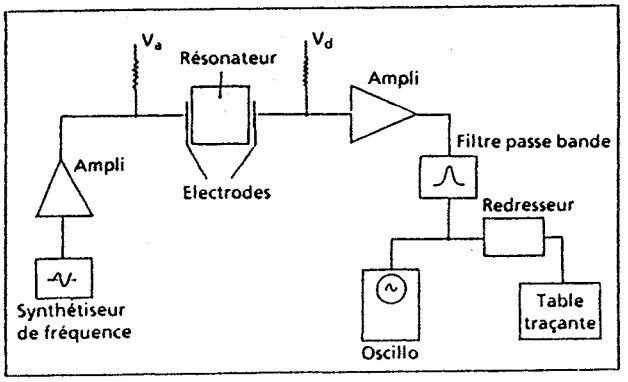


Fig. 10 - Montage expérimental pour la mesure des caractéristiques des résonateurs.

Résultats expérimentaux

Résonateurs en alliage 2017

Les modes de contour, décrits dans les paragraphes précédents, ont pu être mis en évidence dans une plaque carrée en 2017 de 49 mm de côté. Les fréquences théoriques et expérimentales sont en excellent accord.

Un résonateur, vibrant à 160 kHz sur le mode de Lamé, et présentant un coefficient de qualité de

400 000 après traitements chimique, thermique et mise sous vide secondaire, a été étudié en température. La plage de mesure s'étend de 4 K (helium pompé) à 400 K.

La variation de la fréquence en fonction de la température est représentée sur la figure 11. Entre 70 et 400 K, la pente est constante, et en-dessous de 20 K, la pente peut être considérée en première approximation comme nulle.

Les causes de la variation de la fréquence en fonction de la température sont d'une part, la variation des dimensions du résonateur sous l'effet des dilatations thermiques, d'autre part, la variation du module d'Young. On peut établir la relation :

$$\frac{\Delta f}{f} = \frac{1}{2}\left(\frac{\Delta E}{E} + \frac{\Delta L}{L}\right)$$

Le coefficient de dilatation linéaire, qui relie $\frac{\Delta L}{L}$ à la variation de température, est quasiment constant de 50 à 400 K. Le terme en $\frac{\Delta L}{L}$ est inférieur de 2 ordres de grandeur à $\frac{\Delta f}{f}$: il est donc possible d'accéder au module d'Young par intégration numérique (figure 11).

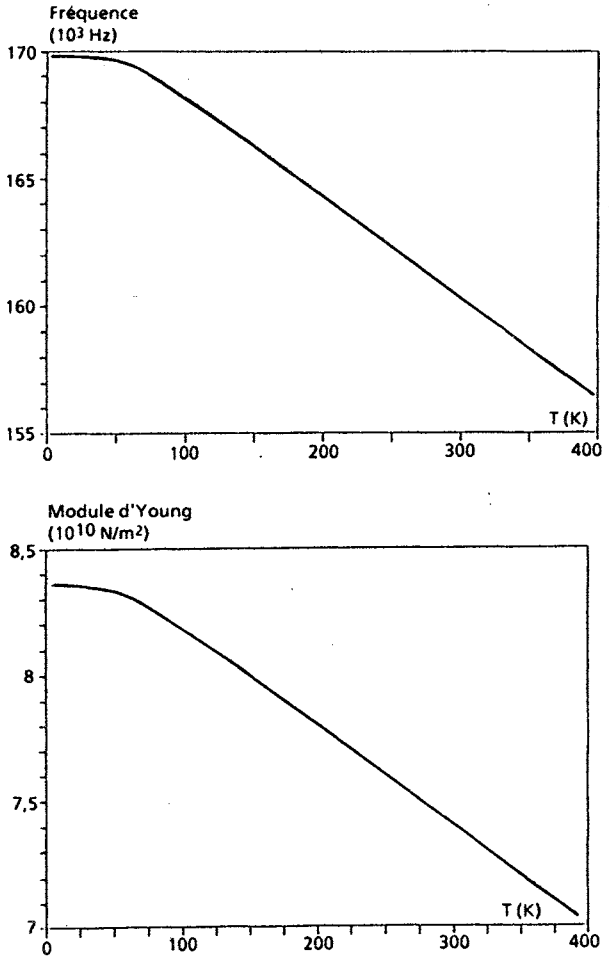


Fig. 11 - Résultats obtenus pour l'alliage 2017 (AU 4G)

Courbe fréquence-température pour un résonateur à 160 kHz. Variations du module d'Young en fonction de la température pour ce résonateur.

La figure 12 représente la variation du coefficient de qualité avec la température. La valeur maximale de 5×10^6 est atteinte pour 40 K et on observe l'apparition d'une série de pics de pertes pour des températures inférieures. Pour l'alliage 2017, ces pics sont attribués à la relaxation des impuretés. Ils sont caractéristiques de ces impuretés et de leur localisation.

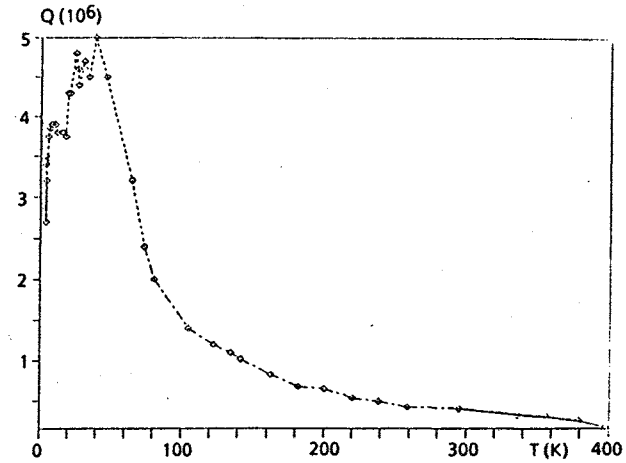


Fig. 12 - Résultats obtenus pour l'alliage 2017 (AU 4G)

Variations du coefficient de qualité en fonction de la température pour le résonateur à 160 kHz.

Résonateurs en silice vitreuse SiO^2

Un bloc parallélépipédique de 50 x 50 x 5 mm, suspendu par des fils de nylon, a été mis en résonance sur le mode de Lamé à 53,5 kHz à température ambiante et un coefficient de qualité de 10^7 a été mesuré.

Un résonateur en silice, identique du point de vue géométrique au résonateur en alliage 2017 présenté dans le paragraphe précédent, a fait l'objet d'une étude en température. Le coefficient de qualité du résonateur à la température ambiante est de 4×10^6.

Le coefficient de dilatation de la silice est encore beaucoup plus faible que celui de l'alliage 2017, permettant de relier directement fréquence et module d'Young.

La fréquence a été calculée à partir du module d'Young donné dans la littérature et, à partir de la fréquence expérimentale, le module d'Young expérimental a été déterminé par la procédure inverse : les résultats sont présentés sur la figure 13.

La figure 14 regroupe les variations thermiques expérimentales du module d'Young de l'alliage 2017 et de la silice vitreuse.

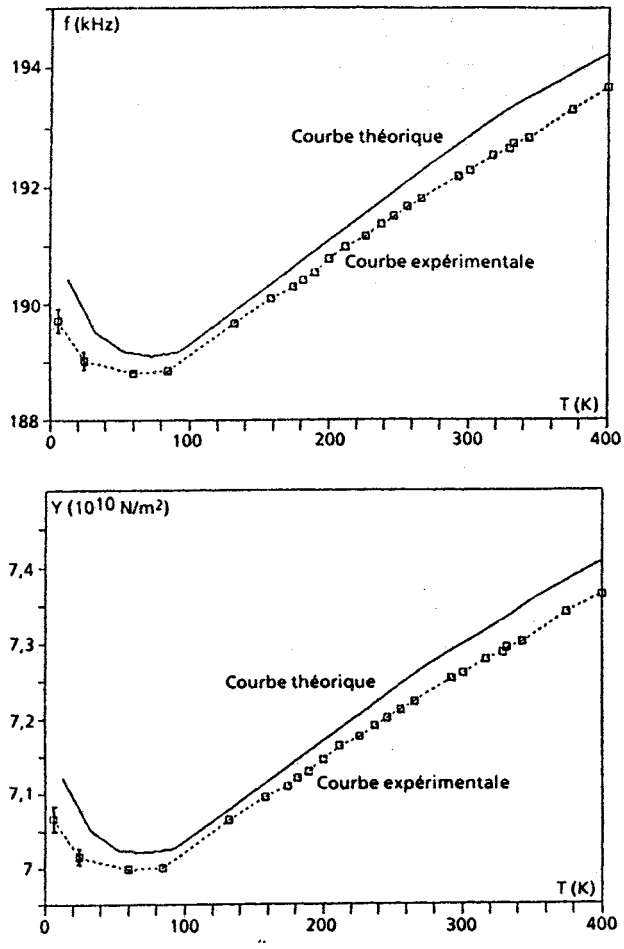


Fig. 13 - Résultats obtenus pour la silice vitreuse ($Si\ O_2$)
Courbe fréquence-température pour un résonateur à 192 kHz : comparaison entre les variations théoriques (trait plein) et expérimentales (trait pointillé).

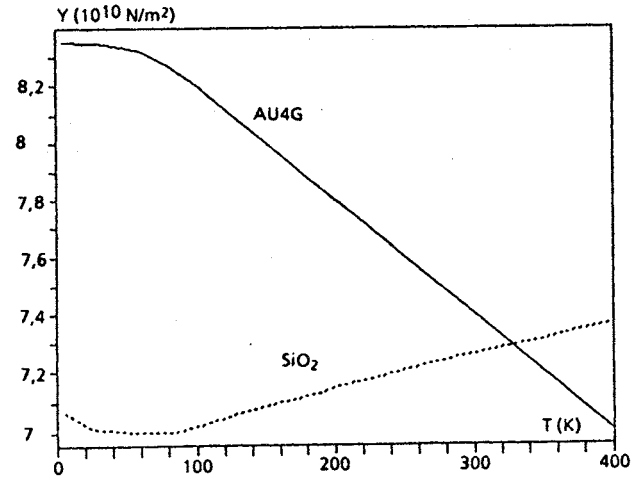


Fig. 14 - Comparaison des variations du module d'Young de l'alliage 2017 et de la silice vitreuse en fonction de la température.

Conclusion

Tout d'abord, cette étude a permis de démontrer la faisabilité de résonateurs à ondes élastiques de volume utilisant des matériaux amorphes, adaptés à l'excitation capacitive. Ces résonateurs vibrent sur un mode de contour particulier dit "mode de Lamé". Les pertes par couplage avec le support ont pu être minimisées par optimisation des fixations.

Des résonateurs en alliage 2017 ont permis de mesurer une surtension d'environ 400 000 à température ambiante, et de 5×10^6 à basse température (40 K). Pour la silice vitreuse, une forte surtension de 4×10^6 a été mesurée à température ambiante.

La linéarité de la relation fréquence-température dans une large gamme de température peut permettre l'emploi de ces résonateurs comme éléments sensibles de capteurs de température.

Par ailleurs, l'excellente concordance entre les valeurs du module d'Young déduites de l'expérience et celles de la littérature montre l'intérêt de cette méthode comme outil d'investigation des propriétés mécaniques des matériaux. En outre, cette méthode bénéficie de la résolution très élevée des mesures de fréquence. Il est possible d'utiliser d'autres matériaux, en particulier le silicium monocristallin qui peut permettre l'obtention de surtensions élevées et l'intégration de résonateurs de faible coût par les méthodes de la micro-électronique.

DETERMINATION OF THE CUT-OFF FREQUENCY OF UNELECTRODED PLATES AND EFFECTS OF LOCAL THICKNESS MODIFICATIONS IN PLANE RESONATORS

J. Détaint*, J. Schwartzel*, C. Joly*, E. Philippot**

*Centre National d'Etudes des Télécommunications Dept PAB/BAG/MCT/CMM Bagneux France
**Laboratoire de Physico-Chimie des Solides Université du Languedoc Montpellier France

SUMMARY

The cut-off frequency of the unelectroded flat plate, the electrical frequency lowering, the mass loading and the electrode geometry are the most important parameters of the plane resonators. They determine the modes and the electrical response of the resonators. The exact knowledge of the first quantity is of prime importance for the modelization and to optimize several steps of the elaboration of the resonators. An approximate determination of this quantity is usually made by an air gap frequency measurement of the unelectroded plates. To obtain a greater precision for these measurements, electrodes of size much smaller than the plate diameter are generally used, since they lead to a simpler frequency response characteristic of a trapped mode. The frequency so obtained differs from the cut-off frequency by a quantity that can be quite important and is a function of the gap, of the electrodes dimensions and of the coupling coefficient of the material. In the first part of this communication we present a method to extract the cut-off frequencey from such air gap measurements with a model of the corresponding resonators.
In a second part, we consider the interest of modifications of the cut-off frequencies of selected regions of a plane resonator. This can be achieved by local thickness modifications, and permits to obtain new degrees of freedom in the design of resonators. The interest of this possibility has already been demonstrated in one case by Lukaszek who has proposed the use of resonators with electrodes embedded in the plate to gain a better control of the unwanted response and of the impedance level of V.H.F. resonator. A theory of the corresponding resonators is presented and the possibility given by this design explored. If three regions of the resonator are allowed to have different thicknesses, new possibilities are gained, such as to eliminate a selected unwanted anharmonic mode or to obtain a faster decrease of the principal mode across the plate. A theory of such resonator with 3 regions is presented, it is based on the approximate equation of the thickness vibrations of piezoelectric plate established by H.F. Tiersten and coworkers.

I - INTRODUCTION

More and more modern Telecommunication systems make use of very high frequency piezoelectric devices for filtering or for frequency generation. This will be the case for the new Pan-European numerical radiotelephone system, now in a development phase, to be operational in the early nineties. It will make use of a very large number of V.H.F devices (Antenna filters, intermediate frequency filters etc..).In the course of a study concerning bulk wave versions of the latter filters, we have observed that significative conceptual and technical advances were needed to establish designs for these filters that permits to obtain the wanted characteristics together with the possibility of a large volume production at a cost compatible with the objectives of a very large diffusion of the mobile and portable stations.

The first point observed was that a very tight control and a very high reproductibility of electrical parameters of the piezoelectric components of the filter were required at a high frequency (near 70 MHz). Since the electrical parameters (equivalent scheme, unwanted responses) of a plane resonator are mostly dependant on three parameters (the thickness [or the cut-off frequency of the unelectrode plate], the electrodes geometry and the plate back) and since the two latter ones parameters can be accurately controled using the modern technology (masquing techniques and precise quartz thin film thickness monitoring), it was deduced that a critical aspect will be to achieve a very precise control of the plate thickness (in the order of 20 micron for the berlinite version of the filter). A method, based upon the theory of the thickness vibrations of the piezoelectric plates (1)(2)(3), permitting to obtain very accurate measurements of the cut-off frequency of very high frequency plates was investigated. This method will be described in the first part of this paper.

In order to obtain a better control of the anharmonic spectrum of the high frequency resonators (the rejection of the anharmonic modes has to be very high for this application) we have theoretically investigated, a method originally proposed par Lukaszek (4), which consists in depositing the electrodes in a previously etched region of the plate. This design allows also to obtain a much lower impedance level for the resonators. The corresponding theory and results will be given in paragraph III.

Other local modifications of the plate thickness can allow to minimize the lateral dimensions of the devices or to integrate several resonators on a single plate. To investigate this point a theoretical model of resonators having three regions with different thickness was established, it will be described in paragraph IV.

All these investigations are based upon the theory established by H.F. Tiersten and coworkers of the thickness vibration of the piezoelectric plates. They extend previous works concerning the modelization of plane resonators with electrodes of arbitrary geometry (5) and of corrugated resonators (6).

II - Accurate determination of the cut-off frequency of unelectroded plane piezoelectric plates

II.1 - Air gap measurements and the cut-off frequency

This quantity (f_{ci}) which is the antiresonance frequency of an infinite plate having a zero mass perfectly conductive electrode is, as shown by the theory, with the mass loading and the electrode geometry, a fundamental parameter of the thickness mode plane resonators. It is the parameter that include the thickness effect and most of the material properties of the resonator.

In practice, the exact knowledge of this quantity is required for the modelization of the resonator and also to make resonators at the wanted frequency with a precise mass loading. Errors on these quantities can have detrimental effects on the equivalent scheme and on the parasitic response of the resonator.

The frequency directly measured from air gap measurements differs from f_{ci} by a quantity which is, in relative value, of the order of $4\ k^2/\pi^2$ (2,6 ‰ for AT quartz and 4 ‰ for AT Berlinite). This error which is very dependent of several experimental parameters (gap, electrode dimension) is an important fraction of the mass loading used for high frequency resonators (about 0,25 to 1 %) so that this frequency cannot be used directly when a precise control of the mass loading is required.

To obtain from air gap measurements a quite exact value of f_{ci} it is necessary to use a theoretical model of the corresponding experimental set up. Two type of experimental set up are of common use. In the first one the plate is situated between two plates having the same dimensions at the plate *(figure 1)* and the quantity measured is either the anti resonance frequency at zero gap or the resonance frequency with a large gap. f_{ci} can be determined by a theory of the vibration of unelectroded square or circular plates made by H.F. Tiersten (7). However, with this experimental set up, the electrical response is very often complicated by plates modes coupled at the edge and is perturbated (antiresonance measurements) by the parasitic capacitance.

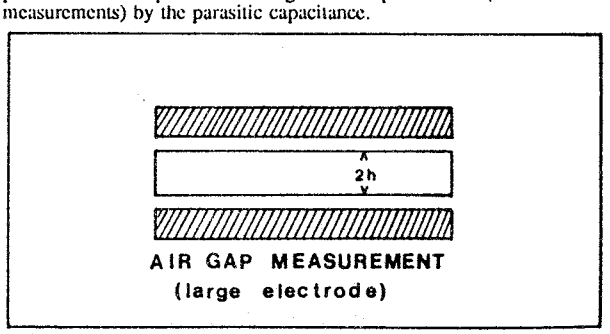


Figure 1 : Air gap measurements with a large electrode

The second experimental set up is represented in figure 2, it uses an electrode with a diameter much lower than the lateral dimensions of the plate. A small gap is generally used and transmission measurements of the resonance frequency most generally employed.

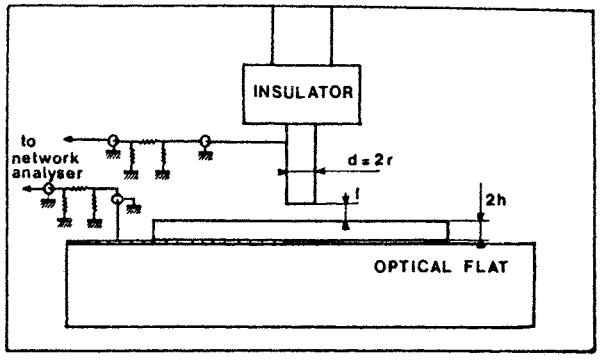


Figure 2 : Air gap measurements with a small electrode

With this set up the electrical frequency lowering in the region under the electrodes is generally sufficient to induce the trapping of the vibration mode at the center of the plate so that "clean" frequency responses with low losses are obtained. This permits to make precise frequency measurements. A typical exemple of the obtained response is given in figure 3. It concerns a 20 MHz AT quartz plate, the electrode diameter was 2,000 mm and the gap varied from 10 µm to 130 µm with 10 µm increments. On figure 3 it is possible to observe that the antiresonance frequency varies with the gap. This can be explained by the equivalent scheme of figure 4. The variation of the Cp capacitance of the gap produces the variations of the antiresonance frequency.

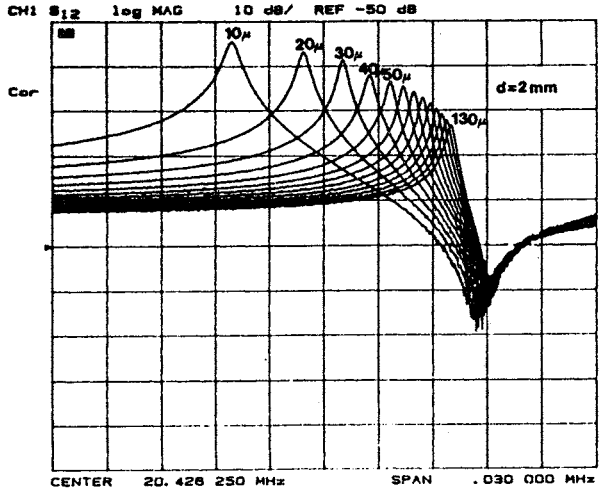


Figure 3 : Example of the frequency responses obtained with a variable gap

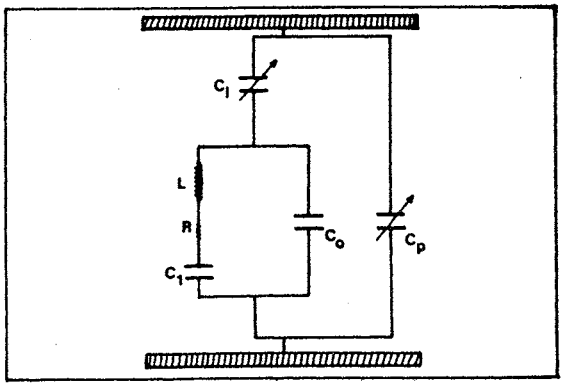


Figure 4 : Equivalent scheme

If a zero gap condition is realized, without applying any stresses on the plate, then, the resonance frequency is that of a resonator having perfectly conductive electrodes with a zero mass. Thin can be accurately simulated by the mathematical extrapolation of the resonances frequencies measured with different gaps.

II.2 - Modelization of a resonator where the energy trapping results only of the electrical lowering

As a basis of this model, we use the scalar equation governing the lateral dependance of the anharmonics in the vicinity of one overtone (number n) of a given thickness mode. This equation was first established for monoclinic plates (2) then for plates of an arbitrary orientation (3). In this case the equation has, in a coordinate system with x_2 normal to the thickness and x_1 x_3 choosen to eliminate the mixed derivative (3) the following form :

$$M_n \frac{\partial^2 \hat{u}_1^n}{\partial x_1^2} + P_n \frac{\partial^2 \hat{u}_1^n}{\partial x_3^2} - \frac{n^2\pi^2 C^*}{4h^2}\hat{u}_1^n + \rho\omega^2 \hat{u}_1^n = \rho\omega^2(-1)^{(n-1)/2}\frac{e_{26}4V_0 e^{j\omega t}}{C^{(1)}n^2\pi^2} \quad [1]$$

In this equation, $\hat{u}_2(x_1,x_3,t)$ is the lateral dependance of $\hat{u}_1^n(x_1,x_2,x_3,t) = \hat{u}_1^n \sin(n\pi x_2/2h)$.

$\hat{u}_1$ arises from a transformation of the usual component of the displacement u_1 and of ϕ the potential , made to replace the inhomogeneous boundary conditions $\phi = \pm V_0 e^{j\omega t}/2 at x_2 = \pm h$ by the homogeneous conditions $\hat{\phi} = 0$. The inhomogeneous term of the equation results from this transformation.

$$u_1^n = \hat{u}_1^n - \frac{e_{26}V_0 e^{j\omega t}}{c^{(1)}2h}\cdot x_2;$$

$$\phi = \hat{\phi} + \frac{x_2 V_0 e^{j\omega t}}{2h};$$

$$u_1 = \sum_n u_1^n$$

c* is either $\bar{c}^{(1)}$ for the unelectroded part of the resonator or $\hat{c}^{(1)}$ for the electroded part. $\bar{c}^{(1)}$ is the stiffened elastic constant relative to the corresponding one dimensional mode (Eigen-value of the Christoffel Matrix).

$$\hat{c}^{(1)} = \bar{c}^{(1)}(1 - 8k^{(1)2}/n^2\pi^2)$$

$\hat{c}^{(1)}$ is a constant that includes the electrical effect of the metallization. $4k^{(1)2}/n^2\pi^2$ is approximatively the relative frequency lowering due to the electrical effect of the metallization $c^{(1)} = \bar{c}^{(1)}(1 - k^{(1)})^2$ is a pseudo ordinary elastic constant. $k^{(1)}$ is the coupling coefficient of the corresponding one dimensional mode. M , P are intricate functions of the material constants and of the plate orientation that can be obtained from the method described in reference (3) or in reference (8).

- Boundary and continuity conditions : The traction free conditions on the major free surfaces, normal to the thickness are automatically verified as a consequence of the method used to establish the dispersion relation. On the surfaces limiting the electroded and the unelectroded regions of the resonator we have to specify the continuity of u_1 and of its normal derivative.

In this paper we suppose that the amplitude of the vibration mode near the edges of the plate are negligible so that no boundary condition at the plate edge are taken in account. It was previously shown that, in this case, they induce a negligible modification of the resonances frequencies [1][2].

Eigen mode analysis : The eigen modes of the resonator can be obtained following the simultaneous solution of the equations for the electroded and unelectroded region taking into account the continuity conditions at the electrode edge. At V=0 the homogeneous form of the approximate equation for the two regions reduces to :

$$M'_n u_{1,11} + P'_n u_{1,33} + 4\pi^2\rho(f^2 - f^{*2}) = 0 \quad [2]$$

with :

$$f^* = f_{ce1} = \frac{n}{4h}\sqrt{\frac{\hat{c}^{(1)}}{\rho}}$$

for the electroded region.

or:

$$f^* = f_{el} = \frac{n}{4h}\sqrt{\frac{\bar{c}^{(1)}}{\rho}}$$

for the unelectroded region.

The coordinate transformations :

$$X_1 = \frac{C^{*1/2}}{M'_n}\cdot\gamma x_1 \qquad X_3 = \frac{C^{*1/2}}{P'_n}\cdot\gamma x_3$$

$$X_1 = r\cos t \qquad X_3 = r\sin t \qquad [3]$$

with: $\gamma^4 = \frac{n^2\pi^2}{2h_0^2}$

$c^* = \bar{c}^{(1)}$ (electroded region)

$c^* = \bar{c}^{(1)}$ (unelectroded region)

leads to the equation :

$$u_{1,rr} + \frac{1}{r}u_{1,r} + \frac{1}{r^2}u_{1,tt} + Au_1 = 0 \qquad [4]$$

$$Where: A = \frac{n\pi}{4}\left(\frac{f^2 - f^{*2}}{f^{*2}}\right)$$

that can be separated as $u_1(r,t) = R(r).T(t)$ in :

$$r^2R'' + rR' + R(r^2A - \upsilon^2) = 0 \qquad [5]$$
$$T'' + \upsilon^2 T = 0$$

For the electroded region the symetrical solution bounded at r = 0 can be expressed as (6) :

$$u_1 = \sum_{m=0}^{\infty} A_m J_m(r\sqrt{A})\cos mt \qquad [6]$$

with m = υ even integer

For the unelectroded region the symetrical solution bounded at r =∞ is (since A is then negative)

$$u_1 = \sum_{m=0}^{\infty} B_m K_m(r\sqrt{A'})\cos mt \qquad [7]$$

with: m = υ even integer and A' = -A

Jm and Km are respectively the Bessel function of first kind and the modified Bessel function of second kind, of order m.

At all points of the electrode edge we have to specify the continuity of u_1 and of its normal derivative. As previously discussed (6)(11)(12)(13), this can be approximated by the expression of these conditions at only a discrete number q_1 of points (figure 4). It the solutions for the two regions are :

$$u_1^I = \left(\sum_{m=0} A_m J_n(r\sqrt{A})\cos mt \sin(n\pi x_2/2h)\right.$$

$$u_1^{II} = \left(\sum_{m=0} B_m K_m(r\sqrt{A'})\cos mt \sin(n\pi x_2/2h)\right. \qquad [8]$$

at the point $P(x_1^P x_2^P x_3^P)$ we have :

$$u_1^I(x_1^P x_2^P x_3^P) = u_1^{II}(x_1^P x_2^P x_3^P)$$

$$n_1\frac{\partial u_1^I(P)}{\partial x_1} + n_3\frac{\partial u_1^I(P)}{\partial x_3} = n_1\frac{\partial u''_1(P)}{\partial x_1} + n_3\frac{\partial u_1^{II}(P)}{\partial x_3} \qquad [9]$$

If we choose to express these conditions at q points at the electrode edge, we obtain 2q linear relations between the coefficients A_0, A_2, A_{2q-2},..., B_0, B_2, B_{2q-2}, to obtain a solution we have to truncate the series to q terms so that the 2q relations constitute an homogeneous linear system in the 2q coefficients. The determinant of this system has to vanish to have a non trivial (0) solution in the coefficients. This condition constitutes a frequency equation that is solved numerically for the eigen frequencies. For each eigen frequency the coefficients of the eigen modes are found as the solution of the homogeneous linear system. The forced modes can then be obtained as described in reference [6].

II.3 - Application of the model to determine f_{ci}

As previously indicated, the airgap frequencies are measured as a function of the gap and then numerically extrapolated to a zero gap. The zero gap frequency is considered as resonance frequency of a resonator having a zero mass electrode of diameter d = 2r. By an iterative process an apparent (or corrected) thickness of the resonator is numerically searched, using the described model, so that the computed resonance frequency corresponds to the experimental one. The f_{ce1} and f_{ci} frequency are then determined.

Since the resonators are designed using similar models making use of the f_{ce1} and f_{ci} frequencies (determination of the electrode radius and of the mass loading to have a spurious free response and a given value of inductance), this global process eliminates a large part of the modelization approximations and of experimental uncertitudes.

II.4 - Experimental and computed results

This method was applied to quartz and berlinite plates with thickness ranging from about 200 to 20 microns that were obtained by conventional polishing techniques.

The experimental set-up is represented on *figure 5*. An HP 8753 network analyser is used for the electrical measurements on *figure 6 and 7* are represented the relative variations of the resonance frequencies measured as a function of the gap for a 51 MHz and a 71 MHs resonator ; the corresponding zero gap frequencies are given on the figures.

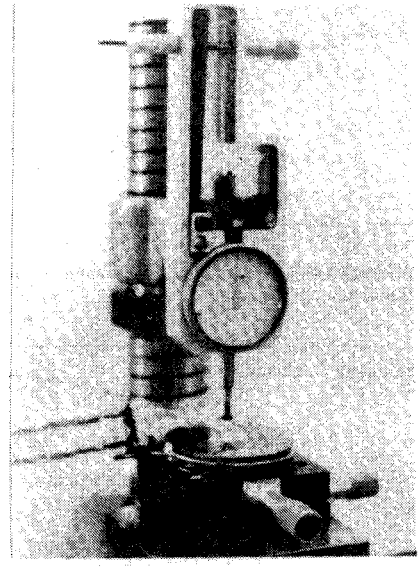


Figure 5 : Photography of the experimental set up

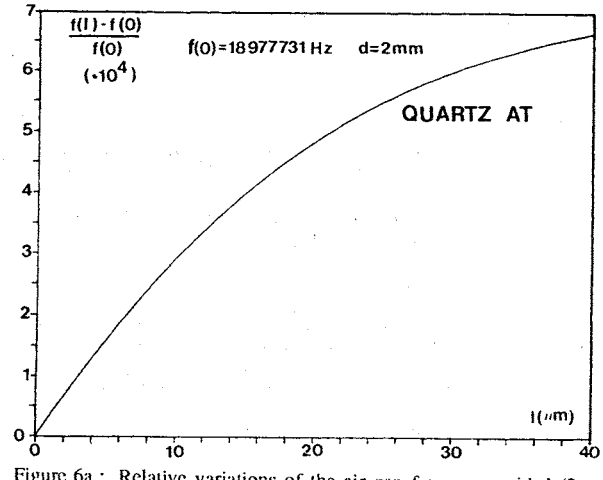


Figure 6a : Relative variations of the air gap frequency with l (2 mm probe ; 2 h ## 87,5 μ)

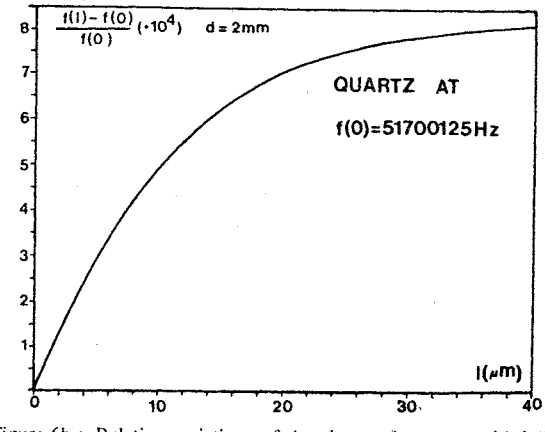


Figure 6b : Relative variations of the air gap frequency with l (2 mm probe, 2 h = 32 μ)

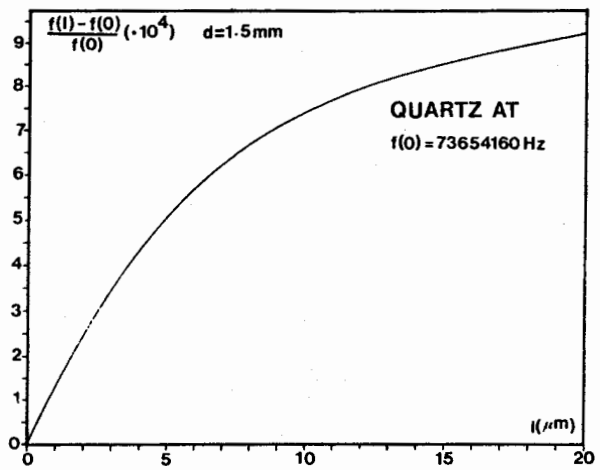


Figure 7a : Relative variations of the air gap frequency with l (1,5 mm probe, 2 h = 22,5μm)

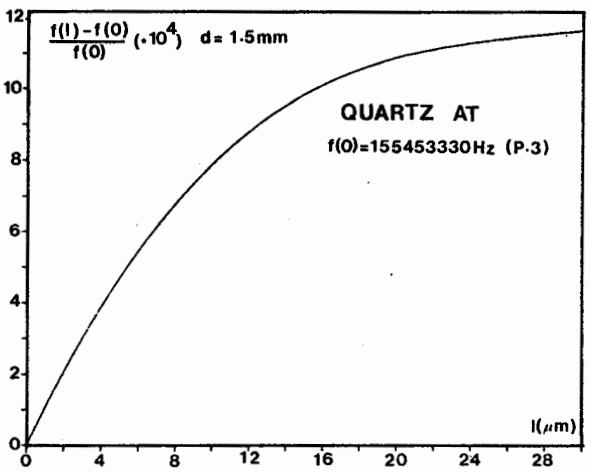


Figure 7b : Relative variation of the air gap frequency with l (1,5 mm probe, 2 h = 22,5 μm)

The frequency response of a 71 MHz resonator measured with a probe having a diameter of 2.000 mm and a gap of a few micron is given on *figure 8*. It can be observed in this case that with this large electrode the trapping is not sufficient to have no unwanted modes. In *figure 9* is given the frequency response of the same plate with a probe having a diameter of 1,50 mm ; we have now a "clean" response. This illustrate the importance of the choice of the probe diameter that must be such that the vibration mode have a negligible amplitude at the edge of the plate.

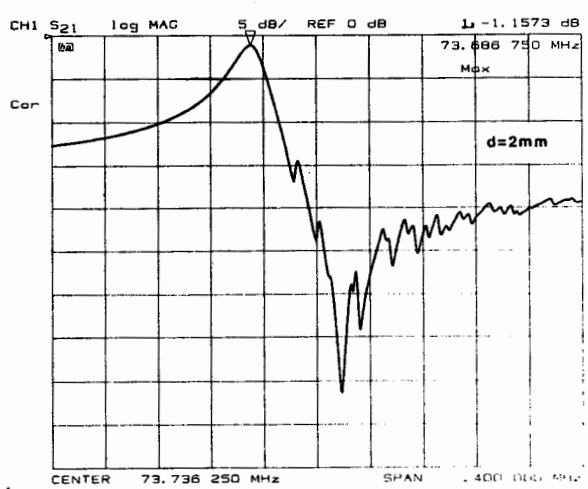


Figure 8 : Response of a 73 MHz resonator with a 2 mm probe

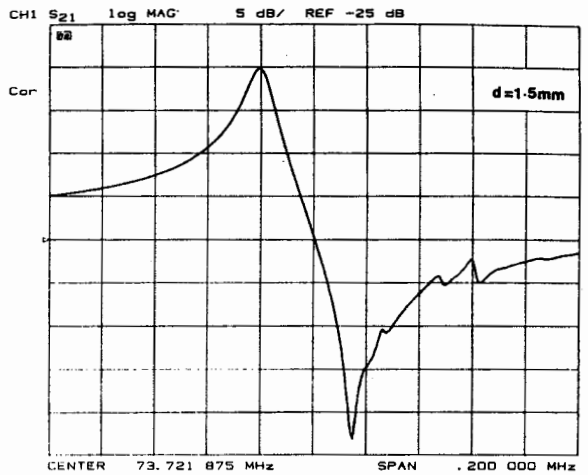


Figure 9 : Response of the same resonator with a 1,5 mm probe

With this experimental set up it was possible to make measurements at much higher frequencies, the 3rd overtone of a 71 MHz plate is displayed in *figure 10*. Due to the very small value of the reactance of the static capacitance at elevated frequency reflection measurements (S_{11}) are to be prefered above say 200 MHz. The S_{11} frequency response of the fifth overtone of a 71 MHz plate is given on *figure 11*.

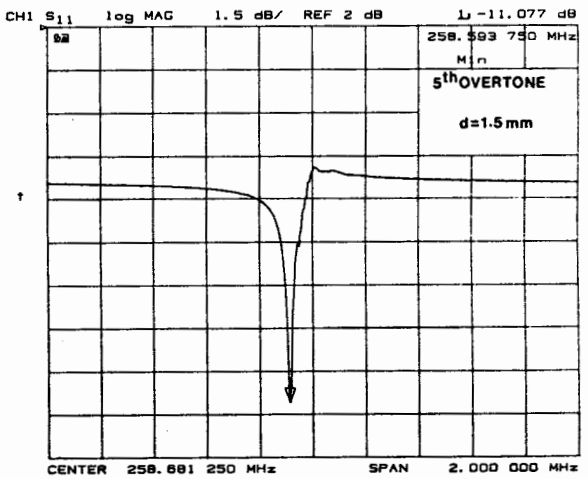


Figure 10 : Air gap response of the 5rd harmonic of a 51 MHz plate (S_{11} measurement)

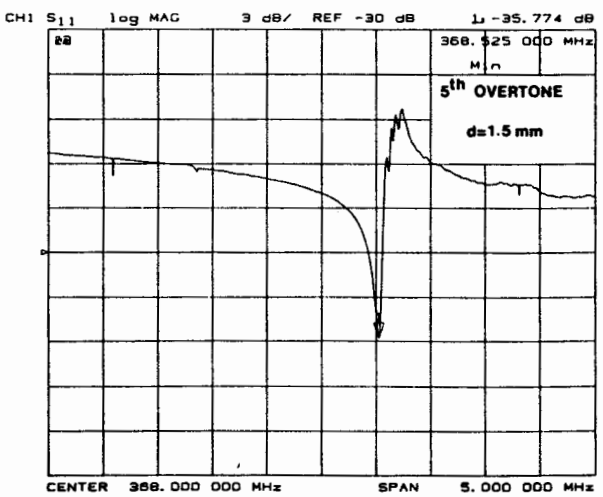


Figure 11 : Air gap response of the 5th harmonic of a 71 MHz plate (S_{11} measurement)

Typical exemples of application of this method are displayed *in table 1 and 2* were are given the thickness of the plate measured by a conventional method, the zerogap frequencies and the extracted values of f_{c1}, f_{ce1} and of the computed "apparent" thickness (relative to the values of the material constant used in the computation).

TABLE 1 : AIR GAP MEASUREMENTS

RESULTS OBTAINED FROM MEASUREMENTS WITH r = 1 mm AT QUARTZ EXEPT SAMPLE N° 5 (Y+3° BERLINITE)

$2h_{measur.}$	$f_{airgap=0}$	f_{ce1}	f_{c1}	$2h_{calc.}$
32.0	51569793.	51662672.	51825912.	32.0579
80.9	20466933.	20424233.	20488768.	81.0899
81.1	20414451.	20371758.	20436127.	81.2988
87.5	18977731.	18935085.	18994914.	87.4672
187.2	9244542.	9195947.	9252228.	186.68

TABLE 2 : AIRGAP MEASUREMENTS

RESULTS OBTAINED FROM MEASUREMENTS WITH r = 0,75 mm

$2h_{measur.}$	$f_{airgap=0}$	f_{ce1}	f_{c1}	$2h_{calc.}$
31,9 F	51799083.	51750620.	51914138	32.0034
31,9 P3	155453330	155439444	155493787	32.0546
22,5 F	73654160	73614188	73846788	22,4983

To avoid the calculation process made to extract the values of f_{ce1} and f_{c1} in the cases where only a moderate precision is required, abaci giving in reduced coordinates the dependance of the zero gap resonance frequency to the electrode diameter were calculated. From the reduced frequency $\Omega^2 = (f_r^2 - f_{ce1}^2)/f_{ce1}^2$ for a given 2r/2h ratio, the f_{ce1} frequency can be calculated from the measured value f_r ; the f_{c1} frequency can then be determined knowing the value of $k^{(1)2}$. They are given on figure 12 (AT cut of quartz) and on figure 13 (AT cut of Berlinite).

To illustrate the dependance upon the electrode dimension of the resonance air gap frequency in a simpler manner we have plotted on figure 14 the zero gap frequencies for a plate of AT quartz, 0.1 mm thick as a function of the electrode diameter.

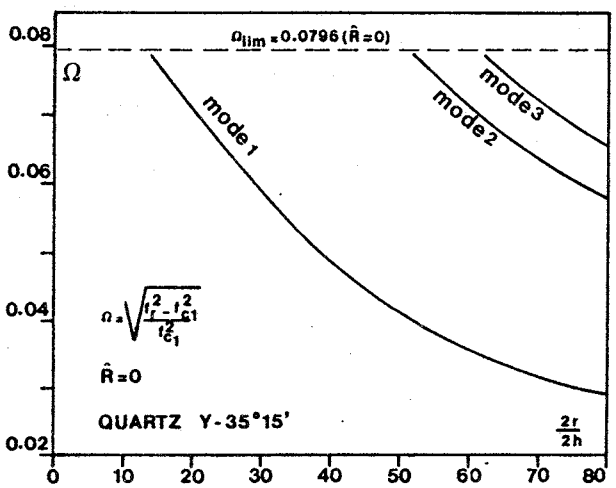


Figure 12 : Correction diagram for air gap measurements (Y-35°15' Quartz)

III - Resonators with electrodes embedded in the plate

This type of resonator was proposed by Lukaszek (4) to control the unwanted responses and the impedance level of VHF resonators for filter applications. As shown in figure 15 the electrodes are deposited in an etched region of the plate. These resonators are of the energy trapping type if the cut-off frequency of the electroded central region (f_{ce1}) is lower than f_{c1}. These frequencies differ by a quantity that depends on the electrode thickness, the depth of the grooved region and on the electrical frequency lowering. If we define a new apparent mass loading $\hat{R}'$ by :

$$f_{ce1} = f_{c1}\left(1 - \frac{4k^2}{n^2\pi^2} - \hat{R}'\right)$$

$$\hat{R}' \# \frac{\rho_m(2h' + 2h'') - (2h_1 + 2h_2)\rho_q}{\rho_q . 2h}$$

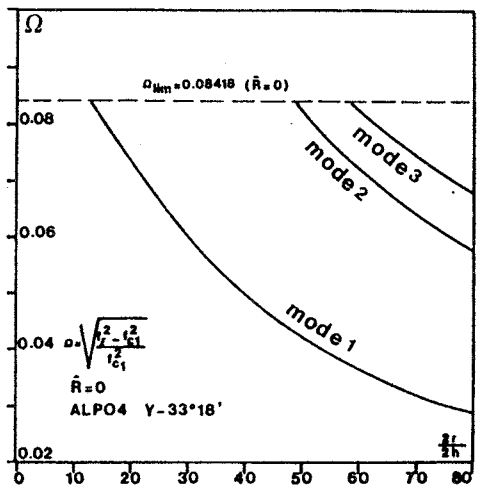


Figure 13 : Correction diagram for airgap measurements (Y-33°18' Berlinite)

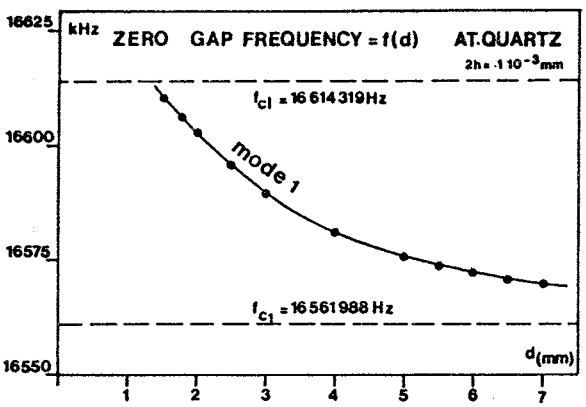


Figure 14 : Resonance frequency (l=0) of an AT Quartz (2h = 0,100 mm)

where :

h, h', h'', h_1, h_2 are defined on figure 15. ρ_m and ρ_q are the densities of electrodes and quartz.

The theory of paragraph II can be applied to these resonators if $\left(\frac{4k^2}{n^2\pi^2} + \hat{R}\right)$ is > 0 to ensure that the trapping of the energy occurs with this new value of f_{ce1}).

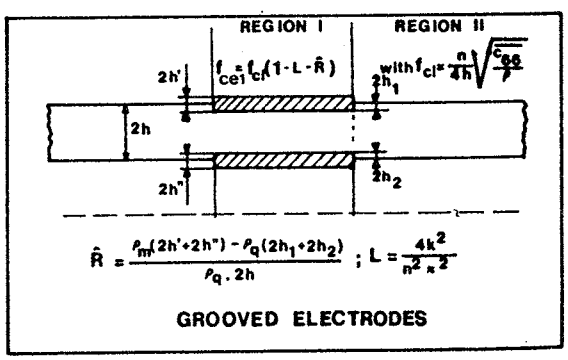


Figure 15 : Resonator with embedded electrodes

Several calculations were made for the case of 72 MHz resonators using quartz and berlinite (the thickness is about 23 µm for AT quartz and 20 µm for AT Berlinite). We have assumed that the interesting values of $\hat{R}'$ were between - 1.10^{-3} and 8.10^{-3}. For a total thickness of aluminium electrodes of about 2 x 1000 Å, this corresponds to an etching ranging from about 2 x 1100 Å to nearly 2 x 200 Å. The results are given in the form of a diagram giving for AT Quartz, the reduced resonance frequency Ω as a function of 2r/2h. On this diagram we can observe (figure 16) that, for a given value of $\hat{R}'$, the trapped modes can exist between $\Omega = 0$ and a limit Ω_{lim} which is a fonction of $\hat{R}'$.

$\Omega_{lim} \# 8k^2/n^2\pi^2 + 2\hat{R}'$

The principal effect of reducing $\hat{R}'$ is to reduce the number of modes that can exist, so that it is possible to obtain resonators with no anharmonics for much larger values of 2r/2h than with conventional values of the mass loading.

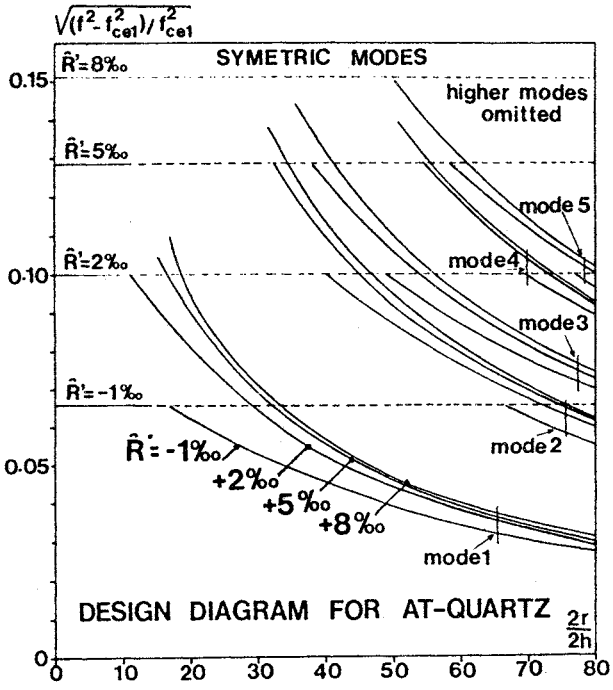


Figure 16 : Diagram for the design of resonators with embedded electrodes

The effect on the computed values of the inductance (6) is displayed on Figure 17 and in table 3. For this computation a thickness of 23,015 µ was used and the values of the electrode diameter were choosen to obtain round values of 2r/2h. Two important facts can be observed. First, as expected, the inductance decreases when 2r is increased, but, not, in following a linear low. The second fact is relative to the influence of $\hat{R}'$, for the low values of 2r/2h, the inductances decreases strongly when $\hat{R}'$ is increased, whereas, for the higher values of 2r/2h, the inductance is quasi independant of $\hat{R}'$. The combination of figure 16 and 17 permits to find a solution to the problem of designing filter resonators with no anharmonics and a low value of the inductance. However, in practice, the possibility of using small values of $\hat{R}'$ is limited by an eventual lack of parallelism of the plates ; so that this parameter must be also taken in account.

Another important parameter to consider, since it determines the required lateral dimensions of the plate, is the lateral extension of the mode. This parameter is closely related to the value of the inductance since this quantity is proportional to the integral of u on the surface of the resonator divided by the square of the integral of u_1 on the electoded surface. For a given - 2 r/2h ratio, the value of L indicates roughly the fraction of the energy which is not confined under the electrodes, but it is always more accurate to compute the vibration mode u (x_1, x_3).

Examples of modes computed for AT Berlinite resonators near 70 MHz are given an Figure 18. On figure 18a we have the mode of a conventional resonator with $\hat{R}' = 1$ % design to thave nearly the minimum inductance together with no anharmonics (this has conducted to use an electrode diameter of 0,40 mm). On figure 18b the value of $\hat{R}'$ was choosen to be 0,25 %, so that it is now possible to have an electrode diameter of 0,70 mm without any anharmonic mode. On Figure 18c a value of $\hat{R}' = 0$ is choosen, it is now possible to use an electrode diameter of 0,95 mm.

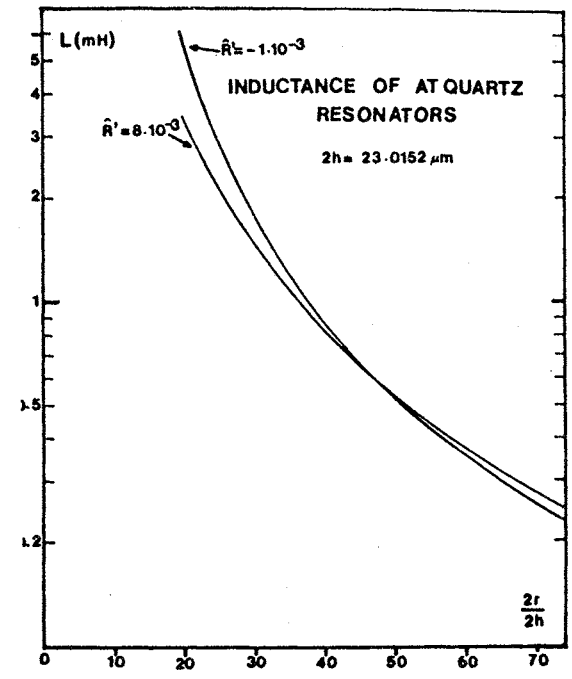


Figure 17 : Inductances of 23 µm AT Quartz plate as a function of $\hat{R}'$ and 2r/2h

TABLE 3 : INDUCTANCE

COMPUTED VALUES OF L for 72 MHz AT QUARTZ
2 h = 23,01524 µm - Values in mH

2r/2h \ $\hat{R}'$	- 1.10^{-3}	2.10^{-3}	5.10^{-3}	8.10^{-3}
20	5.507*	3.891*	3.492*	3.3349*
30	1,761*	1,471*	1.430*	1.430*
40	0,8598*	0,7985	0,8025	0,8135
50	0,5206*	0,5110	0,5205	0,5304
60	0,3539*	0,3582	0,3671	0,3748
80	0,1988	0,2065	0,2124	-

(*) monomode response

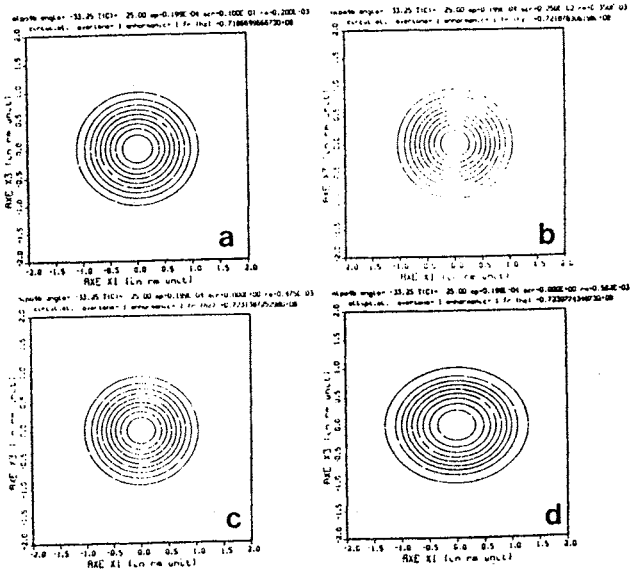


Figure 18 : Computed vibration mode of 72 MHz berlinite resonators

With the same value of $\hat{R}$' it is still possible to have a lower inductance (a greater electrode surface) using an other geometry. This is demonstrated in Figure 18d where an elliptical electrode permits to have an electroded surface equal to that given by a 1 mm electrode without any anharmonics (6).

A careful examination of the value of C_1/C_1 in the different cases would have also revealed that the use of embedded electrodes permits also to have more favourable values of this ratio, together with a response free of spurious modes. This indicates that this type of resonators is also very interesting for VCXO applications.

IV - Theory of resonators having 3 regions with different thickness

As represented on *figure 19* we consider a resonator that have 3 regions with different thickness, the central region noted 1 being electroded.

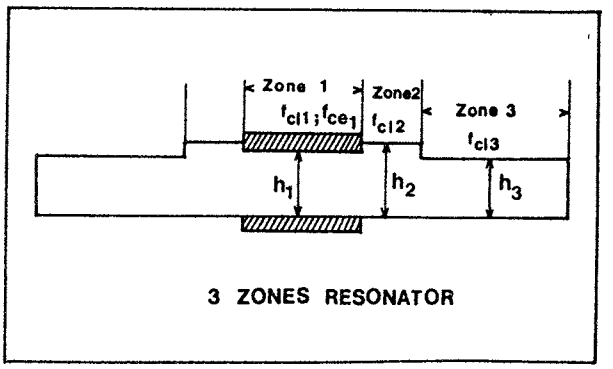


Figure 19 : Resonator with 3 regions of different thickness

In each region the lateral dependance of u_1 is governed by the approximate equation and at each boundary separating 2 regions we have to express the continuity of the displacement and of its normal derivative we also suppose that the lateral dimension of region III are such that the trapped mode considered has a negligible amplitude at the edge of the plate. We suppose that the difference of thickness between two regions is small so that no boundary conditions are to be expressed on the small step existing there.

Eigen mode Analysis : For each region we have to solve the homogeneous Helmoltz equation [2]. In each region the quantity A* which carry the frequency dependance has a different expression

$$A^* = \frac{n\pi}{4}\left(\frac{f^2 - f^{*2}}{f^{*2}}\right)$$

f being either f_{ce1}, f_{cl2}, f_{cl3}.

These cut-off frequencies being defined in regions 1, 2 and 3 by the same formula as f_{cl1} and f_{cl} in paragraph 2, in function of the thickness, mass loading, the electrical lowering, and the appropriate constant C*. To have trapped modes we have also to suppose that f_{cl3} has the highest value among (f_{cl1}, f_{cl2}, f_{cl3}). The same method of resolution by separation given in graph II is used then for the three Helmoltz equation.

The solution for region 1 is now either

$$u_1^I = \sum_m A_m J_m(r\sqrt{A})\cos mt \; if A > 0$$

or

$$u_1^I = \sum_m A_m I_m(r\sqrt{|A|}\cos mt \; if A < 0$$

The solution for region 2 is :

$$u_1^{II} = \sum B_{m_m} \mathcal{B}_m \; (r\sqrt{|A|})\cos mt + \sum_m C_m \mathcal{B}'_m \; (r\sqrt{|A|}\cos mt$$

where $\mathcal{B}$ and $\mathcal{B}'$ are either J_m and Y_m or I_m and K_m depending on the sign of A. J_m, Y_m, I_m, K_m being the Bessel functions or modified Bessel functions of order m.

The solution for region 3 is :

$$u_1^{III} = \sum_m D_m K_m(r\sqrt{|A|})\cos mt$$

The continuity conditions relative to the continuity for u_1 and of its normal derivative at the boundary of regions 1 and 2 and of regions 2 and 3 are, as in paragraph 2, expressed at a discrete number of points (p). Again this conduct to truncate the series to p terms so that we have 4 p linear relations between the coefficients A_i, B_i, C_i and D_i in the form :

$$\alpha A = \beta B + \gamma C$$

$$\alpha' A = \beta' B + \gamma' C$$

$$\beta'' B + \gamma'' C = \delta D$$

$$\beta''' B + \gamma''' C = \delta' D$$

where : A, B, C, D are vectors containing each the p coefficients of the tuncated series ; $\alpha, \beta, \gamma, \ldots, \delta'$ are $\rho \times \rho$ matrix containing the values of the # functions at the points of the boundaries.

From these relations it is possible to obtain the homogeneous system

$$\begin{pmatrix} \alpha & -\varepsilon \\ \alpha' & -\varepsilon' \end{pmatrix}\begin{pmatrix} A \\ D \end{pmatrix} = 0 \quad with \quad \begin{pmatrix} \varepsilon \\ \varepsilon' \end{pmatrix} = \begin{pmatrix} \beta & \gamma \\ \beta' & \gamma \end{pmatrix}\begin{pmatrix} \beta'' & \gamma'' \\ \beta''' & \gamma''' \end{pmatrix}^{-1}\begin{pmatrix} \delta \\ \delta' \end{pmatrix}$$

The determinant of this system must vanish to have a solution in the A and D coefficients. This condition constitute a frequency equation which is solved numerically to obtain the eigen frequencies. For each of the eigen frequencies,the corresponding eigen mode is obtained as the solution of the homogeneous linear system which gives the A and D coefficients. The B and C coefficients are obtained by back substitution in one of the following relations:

$$\begin{pmatrix} B \\ C \end{pmatrix} = \begin{pmatrix} \beta'' & \gamma'' \\ \beta''' & \gamma''' \end{pmatrix}^{-1}\begin{pmatrix} \delta \\ \delta' \end{pmatrix} D$$

$$\begin{pmatrix} B \\ C \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta' & \gamma' \end{pmatrix}^{-1}\begin{pmatrix} \alpha \\ \alpha' \end{pmatrix} A$$

The forced modes are then obtained as a linear combination of the eigen mode in a manner described in reference (6).

VI - CONCLUSION

A method is proposed to obtain precise measurements of the cut-off frequencies of unelectroded plates. These quantities can then be used in a model, similar to those which has permitted the extraction of the cut-off frequency, to determine the other parameters of the resonator (electrode geometry and mass loading). This permits to obtain precise values of the equivalent scheme and, if necessary, a response free from anharmonic modes, even for V.H.F resonators. A modelization of the resonators with embedded electrodes proposed, by Lukaszek has been made, and, the properties of this type of resonators computed. It appears that they are an efficient solution to obtain filter or VCXO resonators with more favourable values of the equivalent scheme.

A modelization of resonators having 3 regions with different thicknesses was made and is now being used to find designs that permit to have the advantages of the resonators with grooved electrodes together with a much reduced sensitivity ot the defects of paralellism of the plates and the possibility of reducing the lateral dimensions of the resonators. We expect that the design found using the model of the grooved resonator or of the 3 regions resonators will permit to have much better solutions for the intermédiate frequency filter of the new numerical radiotéléphone system.

Aknowledgments : The authors whish to acknowledge M. H. Carru for many indications given during the course of this work and R. Lefèvre for kindly lending the airgap apparatus of figure 5. They are indebted to MM. Chenebault, Douis and Mrs Daniel for the elaboration of the many samples used during this study, they thank particularly Mrs M.P. Louis and P. Durand for the preparation of this manuscript.

REFERENCES

[1] H.F. Tiersten
J. Acoust. Soc. Am vol 59 p 879 (1976)

[2] H.F. Tiersten, R.C. Smythe
J. Acoust. Soc. Am vol. 65 n° 6 p 1455 (1979)

[3] D.S. Stevens, H.F. Tiersten
J. Acoust. Soc. Am vol. 79 n° 6 p 1811 (1986)

[4] T.J. Lukaszek
IEEE trans. Sonics vol. S.U. 16 n° 4 p 238 (1971)

[5] J. Détaint, A. Zarka, B. Capelle, Y. Toudic, J. Schwartzel, E. Philippot, J.C. Jumas, A. Goiffon, J.C. Doukhan
Proc. 40th Annual Frequency Control Symp. p 101 (1986)

[6] J. Détaint, J. Schwartzel, C. Joly, E. Philippot
Proc. 41st Annual Frequency Control Symp. p 314 (1987)

[7] H.F. Tiersten, R.C. Smythe
Proc. 41st Annual Frequency Control Symp. p 230 (1981)

[8] R.C. Peach
Proc. 42nd Annual Frequency Control Symp. p 38 (1988)

[9] Dulmet, H.F. Tiersten
Proc. 1986 Ultrasoncis Symp. p 355 (1986)

[10] J. Détaint, H. Carru, J. Schwartzel, B. Capelle, A. Zarka
Proc. 42nd Annual Frequency Control Symposium p 19 (1988)

[11] K. Nakamura, M. Shimizu
Proc. 1976 Ultrasonics Symposium p 606 (1976)

[12] H. Seikimoto
IEEE trans. Sonics Vol. SU 31 n° 6 p 664 (1984)

[13] H. Carru, R. Lefèvre
Proc. 42nd Annual Freq. Control Symp p 101 (1988)

A 524 kHz THERMALLY COMPENSATED QUARTZ RESONATOR FOR WRIST-WATCHES

C. Bourgeois

Centre Suisse d'Electronique et de Microtechnique SA
Maladière 71, 2007 Neuchâtel, Switzerland

ABSTRACT

This paper describes a quartz resonator for wrist-watch applications which oscillates at 524 kHz and is thermally compensated up to the second coefficient.

This compensation is achieved by the mechanical coupling of two low-frequency contour modes of a hollowed-out structure, namely a flexure-type and an extensional mode.

Such resonators have an inactive area where stresses and movements are evanescent, so that they can be glued with epoxy onto a tiny pedestal. They could be mass-produced by using photolithography and chemical etching.

The coupling is controlled by the shape of the resonator rather than by the angles of cut. The structure has been devised so as to minimize the sensitivity to under-cutting.

These resonators are also characterised by a good shock resistance and a quality factor typically in the range of 0.5×10^6.

INTRODUCTION

In piezoelectric resonators, electromechanical coupling makes it possible to excite elastic modes. Most of them are characterized by isolated classes of resonance modes such as thickness shear, surface shear, flexural and extensional modes. One mode is usually desired and all the others are unwanted.

In some cases two modes are deliberately elastic-coupled in order to achieve a particular property, most often the frequency-temperature dependence of the resonator.

In the GT or ZT resonators [1,3,2,5] two extensional contour modes of a rectangular plate are coupled together through an elastic cross-constant (Poisson's ratio). The resulting frequency has a very small change over a wide temperature range. In the tuning-fork the flexural mode can be coupled with a torsional mode of the arms [4].

This paper concerns a thermally compensated resonator for wrist-watch applications which can be clamped and which oscillates at a low frequency in order to minimize oscillator power consumption. This compensation is achieved by the coupling of an extensional mass-loaded mode with a flexure-type mode as shown in Fig. 1. They belong to the same family of dilatation contour modes [6,7].

The structure is such that the elastic part of the coupling can be balanced with an inertial one for any crystallographic orientation of the plate. The coupling factor vanishes, but not its thermal derivatives. Thus its influence is limited only to the temperature derivatives of the frequency. It is suggested that this resonator be designated FXT, to stress the fact that it uses a FleXure-eXTensional coupling.

FREQUENCY - TEMPERATURE DEPENDENCE

The frequency of a resonator at a given temperature can be expressed as

$$F(T) = F_o \left(1 + \alpha(T-T_o) + \beta(T-T_o)^2 + \gamma(T-T_o)^3\right)$$

where
- T = temperature
- T_o = reference temperature
- F_o = frequency at T_o
- α = first-order temperature coefficient
- β = second-order temperature coefficient
- γ = third-order temperature coefficient

COUPLING FACTOR

Elastic cross-constant (Poisson's ratio or mutual constants) or inertial forces allow energy exchanges between modes. The modes are coupled. The frequency of one mode as well as its thermal properties depend on the proximity of the other.

In a first approximation the coupling effect on one mode can be expressed as

$$F = F_o \left[1 + K^2/4 \cdot \Delta_f\right]$$
$$\alpha = \alpha_o + KK'/2 \cdot \Delta_f$$
$$\beta = \beta_o + \left[KK'' + K'^2\right]/2 \cdot \Delta_f$$
$$\gamma = \gamma_o + \left[KK''' + K'K''\right]/2 \cdot \Delta_f$$

where
- F = frequency of the mode with coupling
- F_o = frequency of the mode without coupling
- F_c = frequency of the compensation mode
- $\Delta_f = (F_o - F_c)/F_o$
- $\alpha_o\ \beta_o\ \gamma_o$ = frequency-temperature coefficient without coupling
- $K = K_e - K_i$ = coupling factor
- K_e = elastic coupling
- K_i = coupling with inertial forces
- $K'\ K''\ K'''$ = thermal derivatives of coupling factor K.

K is negative if $K_i > K_e$.

The elastic coupling K_e depends mainly on the cut angles of the resonator. The inertial coupling K_i basically depends on the shape and the boundary conditions. Thus the thermal derivatives of K_i depend on the thermal expansion and are much smaller then those of K_e, which depend on the elastic constantes.

K will be chosen as small as possible in order to minimize the sensitivity of the first thermal coefficient α (i.e. the ZT cut). On the other hand, the second thermal coefficient β of an uncoupled mode is generally negative if $\alpha = 0$ [8]. That means that the frequency of the compensated mode must be greater then that of the compensation mode.

RESONATOR EXTENT

The resonator extent can be defined as

$$E = F*L*2/V_l$$

where F = frequency
L = maximum dimension
V_l = longitudinal wave velocity (~ 5500 m/s)

In a GT or ZT rectangular plate E = 1.

FLEXURE-EXTENSIONAL COUPLING

In the FXT structure (Fig 1) the flexure-extensional coupling allows low-frequency modes and minimizes the consumption of the oscillator.

The coupling factor K can be adjusted by the distance between the external arms, or by the width ratio of the external arms to the central arm. K can be negative if these two parameters decrease.

These two parameters can at the same time minimize the coupling factor K and the sensitivity due to undercutting during chemical etching.

K is little sensitive to the ratio of arms length to total length, as well as the frequency of the extensional mode, if this ratio is about 1/3. This parameter is used to adjust the frequency of the flexure-type mode in order to optimize the thermal dependence of the extensional mode.

The cut of the FXT resonator is such that the first temperature coefficient α is close to zero without coupling in order to minimize its sensitivity to the dimensions.

The ZT cut allows a piezoelectric coupling with a perpendicular field (Fig. 3).The lateral faces are tilted through chemical etching. Thus the contour modes are strongly coupled with unwanted out of plane modes.

The +2° X cut is well adapted to chemical etching. The piezoelectric coupling is obtained with an in-plane field (Fig. 4). The two rectangular plates are shifted by 10 µm in the X direction in order to compensate the effect of the Poisson's ratio σ_{12} on the movement of the clamping zone.

The FXT structure has been analyzed with a contour mode computer program[7]. The unwanted out-of-plane modes have been analysed with a finite-element program[9].

EXPERIMENTAL RESULTS

Cut		+2° X
Frequency		524 kHz
Length		3.7 mm
Width		2.75 mm
Thickness		0.1 to 0.2 mm
Extent	<	0.7
Quality factor	>	450'000
β	<	$10*10^{-9}/°C^2$

CONCLUSIONS

The thermal properties of the described resonator are at least 4 times better then those of the flexural tuning-fork.

The structure with 3 arms minimizes the sensitivity of the chemical etching process.

REFERENCES

1. MASON, W.P.: A New Quartz-Crystal Plate, Designated the GT. Proc. IRE 1940, 220-223.
2. HERMANN, J. & BOURGEOIS, C.: A New Quartz Crystal Cut for Contour Mode Resonators. Proc. 33th Annual Frequency Control Symp. 1979, 255-262.
3. KAWASHIMA, H. et al.: New Frequency Characteristics of Miniaturized GT-Cut Quartz Resonators. Proc. 34th Annual Frequency Control Symp. 1980, 131-135.
4. KOGURE, S. et al.: New Type Twin Mode Resonator. Proc. 34th Annual Frequency Control Symp. 1980, 160-166.
5. HERMANN, J.: A Novel Miniature ZT-Cut Resonator. Proc. 39th Annual Frequency Control Symp. 1985, 375-380.
6. HOLLAND, R. & EERNISSE, E.P.: Contour Extensional Resonant Properties Rectangular Piezoelectric Plates. IEEE Trans. Sonics and ultrasonics, SU-15, 1968, 97-105.
7. BOURGEOIS, C.: Decoupled Families of Contour modes of Planar Thin Plates. Proc. 34th Annual Frequency Control Symp. 1980, 419-425.
8. GERBER, E.A. & BALLATO, A.: Precision Frequency Control. Academic Press, Inc. 1985.
9. TRÜMPY, K.: Calcul de Perturbation avec Élements Finis Aplliqué aux Résonateurs à Quartz. 1st European Time and Frequency Forum. 1987, 17-124.

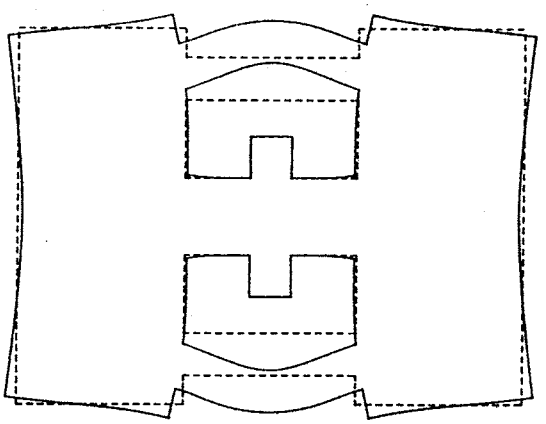


Figure 1a: Flexure-type mode

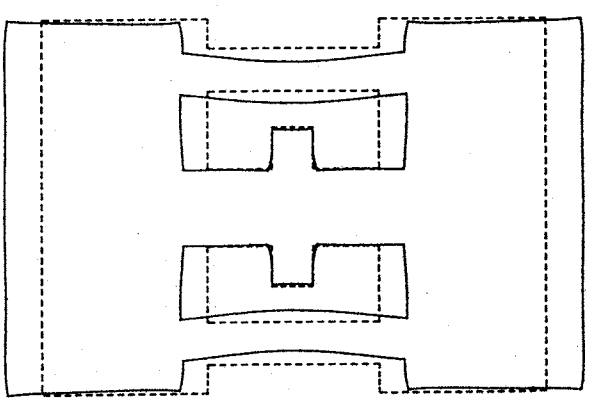


Figure 1b: Extensional-type mode

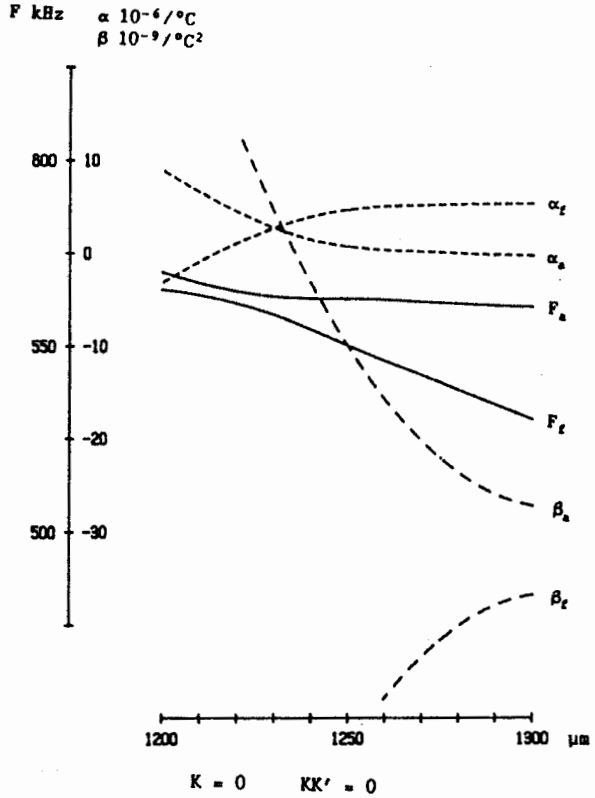
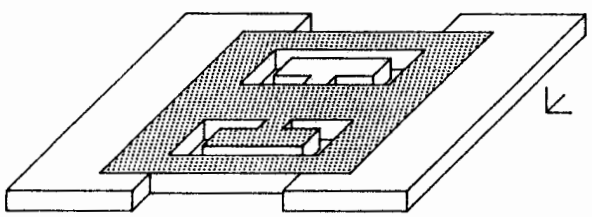


Figure 2: Arms length dependance



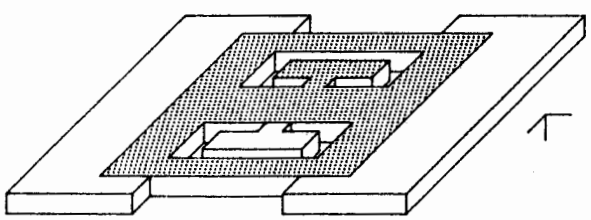


Figure 3: Electrode configuration with perpendicular field

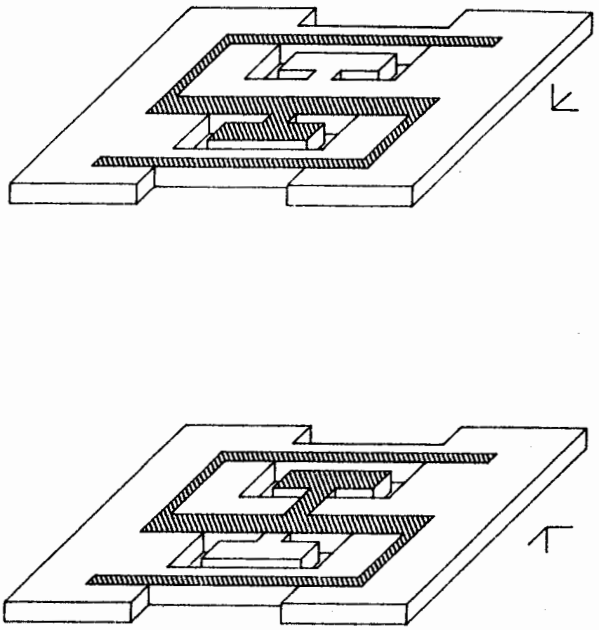


Figure 4: Electrode configuration with in-plane field

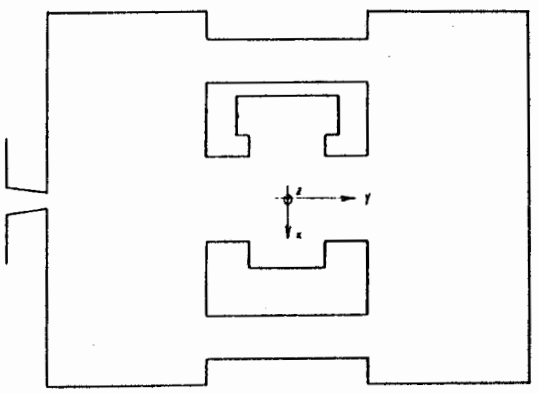


Figure 5: Outline of the +2°X cut

INFLUENCE DES CONTRAINTES MECANIQUES INTERNES SUR L'HYSTERESIS THERMIQUE ET LE VIEILLISSEMENT DES RESONATEURS A QUARTZ

J. Beaussier

Office National d'Etudes et de Recherches Aérospatiales
BP 72, 92322 Châtillon Cedex, France

RESUME

La première partie de l'exposé regroupe des résultats expérimentaux obtenus à l'ONERA sur des résonateurs de structure différente. Ces résultats concernent en particulier la sensibilité et l'hystérésis thermique ainsi que le coefficient de surtension et le vieillissement. Le but de cette présentation est d'obtenir des relations qui, pour un même résonateur ou une même catégorie de résonateurs, apporteraient quelques précisions sur les phénomènes physiques mis en cause dans les dérives typiquement non réversibles que sont l'hystérésis thermique et le vieillissement. Les difficultés rencontrées pour ce travail montrent la complexité et la multiplicité de ces phénomènes.

La seconde partie tente de justifier cette complexité par le rôle des contraintes mécaniques internes au quartz, tel qu'il résulte des connaissances acquises actuellement dans ce domaine. Ces contraintes peuvent être soit d'origine extérieure au quartz, soit d'origine interne. Dans le premier cas, elles sont créées par des forces de surface transmises par les électrodes et les supports. Dans le second cas, elles sont créées par des forces de volume. Quelle que soit leur nature, les contraintes internes évoluent en fonction du temps selon un processus classique de relaxation qui contribue au vieillissement. Elles évoluent en fonction de la température selon un processus plus complexe lié principalement aux dilatations et aux variations des coefficients élastiques. Ces deux effets modifient l'état de contrainte interne du quartz et provoquent des dérives irréversibles dues à des phénomènes d'élasticité non linéaire.

1 - INTRODUCTION

La présente communication traite de l'état actuel des études effectuées à l'ONERA sur l'hystérésis thermique de résonateurs à quartz destinés à des oscillateurs de hautes performances. Elle constitue la suite d'un premier exposé fait sur le même sujet au précédent Forum 88 [1].

La mesure de l'hystérésis thermique, qui doit être dissociée d'autres causes de dérives, nécessite l'utilisation de programmes complexes d'acquisition et de traitement de données. Un certain nombre de résultats ainsi obtenus en 1988 concernent des résonateurs de coupe SC, soit BVA miniatures du LCEP, soit QAS de la CEPE. A partir d'une interprétation de ces résultats, une modélisation des phénomènes d'hystérésis thermique a été effectuée.

Le comportement thermique des contraintes mécaniques internes au quartz apparaît comme la cause privilégiée d'hystérésis thermique. Une triple origine caractérise ces contraintes. Elles peuvent en effet être soit liées aux défauts du matériau, soit dues aux électrodes lorsque celles-ci sont adhérentes, soit transmises par les fixations. Une étude physique de ces contraintes permet de définir les mécanismes susceptibles de provoquer l'hystérésis sur la fréquence, en particulier pour les premières et les troisièmes. L'importance respective de ces mécanismes peut être évaluée pour les divers types de résonateurs étudiés. Un certain nombre de possibilités existent pour réduire le phénomène. Elles concernent la sélection du quartz, des précautions d'assemblage et des traitements thermiques appropriés.

2 - RESULTATS EXPERIMENTAUX

2.1 - Mode opératoire

Le mode opératoire et les difficultés liées à la mesure de l'hystérésis thermique ayant déjà été exposés en [1], seuls quelques points essentiels sont rappelés :

- le balayage thermique est effectué symétriquement autour du point d'inversion avec des amplitudes $2\Delta T$ variant de ± 2 à ± 10°C ;
- ce balayage est suffisamment lent (0,5°C/heure) pour éliminer l'effet des sensibilités thermiques dynamiques ; dans ces conditions, une correction des dérives de vieillissement est nécessaire ;
- la connaissance précise de la température est impérative et implique, en particulier, la prise en compte des constantes de temps thermiques boîtier-quartz ;
- lorsque toutes ces précautions sont prises, la courbe d'hystérésis obtenue par plusieurs balayages se présente sous la forme d'un cycle limite et l'amplitude relative maximale de l'hystérésis, notée $\delta F_1/F_1$, est obtenue et mesurée au point d'inversion T_1 (Fig. 1).

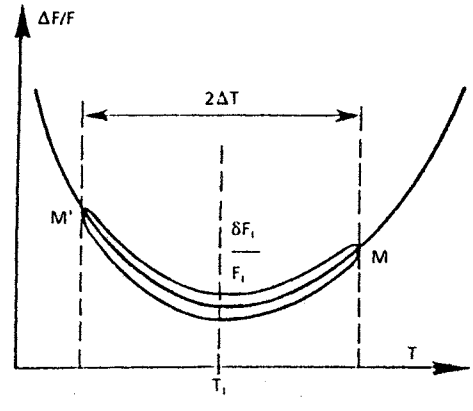


Figure 1 - Cycle limite d'hystérésis.

Dans ces conditions, des logiciels complexes d'acquisition et de traitement de données ont été mis au point pour la mesure précise de la sensibilité et de l'hystérésis thermique. Des approches polynomiales de degré 2 ou 3 ont été utilisés et ces logiciels

effectuent les opérations suivantes.

- Mesure de ΔT et de $\delta F_1/F_1$ - La valeur maximale $\delta F_1/F_1$ apparaît proportionnelle au carré de l'amplitude ΔT du cycle. Le logiciel définit par conséquent un coefficient K_H tel que $\delta F_1/F_1 = K_H \Delta T^2$ et calcule les valeurs de K_H.

- Identification des coefficients a, b, c du polynome du troisième ordre représentatif de la sensibilité statique et tel que:

$$\Delta F/F = a(T-T_1) + b(T-T_1)^2 + c(T-T_1)^3.$$

Cette identification est faite, par conséquent à $T = T_1$ pour des cycles thermiques d'amplitude $T - T_1 = \Delta T$. Les coefficients c, de l'ordre de quelques $10^{-11}/°C^3$ ne varient pas et les coefficients linéaires sont négligeables au voisinage du point d'inversion. Dans ces conditions, seul le coefficient quadratique b est différent dans les cycles thermiques aller-retour et l'écart δb, considéré comme représentatif de l'hystérésis thermique, est aussi calculé par le logiciel. On rappelle que b est défini par l'expression :

$$b = \frac{1}{2F}\left(\frac{d^2F}{dT^2}\right)_{T = T_1}.$$

2.2 - Résultats obtenus

Les mesures d'hystérésis thermique effectuées à l'ONERA en 1988 et utilisées dans la présente communication concernent huit résonateurs dits "BVA miniatures", en coupe SC, réalisés par le LCEP et un résonateur QAS réalisé en même coupe par la CEPE. Le tableau I regroupe, pour ces neuf résonateurs, les résultats d'hystérésis thermique complétés par trois caractéristiques supplémentaires : constante de temps boitier quartz τ, coefficient de sensibilité dynamique $\tilde{\alpha}$ et coefficient de surtension, calculés simultanément. Les résonateurs BVA ont été classés par ordre de décroissance du coefficient K_H qui varie de $3{,}4.10^{-10}/°C^2$ à $0{,}104.10^{-10}/°C^2$, ces deux valeurs constituant les extrêmes pour l'ensemble des résonateurs évalués en 1988. Pour la majorité des résonateurs, K_H est compris entre 0,35 et $0{,}75.10^{-10}$ et ces valeurs montrent une amélioration d'un facteur 2 environ par rapport aux mesures effectuées en 1987 sur le même type de résonateur. Le résonateur QAS type RPI de la CEPE bien que de conception très différente présente une valeur de $K_H = 0{,}64$, par conséquent très cohérente avec les BVA. Il faut aussi souligner que la plus faible valeur de K_H pour le BVA n° 8 correspond aussi aux valeurs optimales de Q, α, τ ainsi qu'à la valeur la plus élevée de T_1 et la plus faible de b et que ce résonateur est fabriqué avec un cristal de quartz différent de celui utilisé pour les autres résonateurs. D'une manière générale, le coefficient de surtension croît lorsque K_H décroît. On notera aussi les valeurs dispersées et peu significatives du coefficient $\tilde{\alpha}$.

Un point très intéressant de ce tableau concerne les écarts constatés entre K_H et δb. Ces écarts ont été calculés en valeur relative et changent de signe en dessous de $K_H = 0{,}7$. La différence $K_H - \delta b$ est susceptible de renseigner sur l'hystérésis thermique. Elle fait l'objet des paragraphes suivants.

2.3 - Première interprétation graphique des résultats

Avec l'indice A pour le cycle thermique aller et R pour le cycle retour et en introduisant les variations $\delta F_1/F_1$ de la fréquence d'inversion et δb du coefficient b autour de la température T_1, une expression générale de l'hystérésis thermique entre T_1 et $T_1 + \Delta T$ est donnée par :

$$\left(\frac{\delta F}{F}\right)_H = \left(\frac{\delta F_1}{F_1}\right)_A - \left(\frac{\delta F_1}{F_1}\right)_R + [(\delta b)_A - (\delta b)_R]\,(\Delta T)^2$$

Des considérations de symétrie permettent de simplifier cette expression générale dont l'utilité est principalement expérimentale. La symétrie par rapport à l'axe des fréquences est la conséquence du fait que seul le terme quadratique est pris en compte dans la caractéristique fréquence-température et cette propriété permet de ne considérer que la moitié ΔT de l'excursion thermique totale. Une autre symétrie peut être introduite par rapport à la caractéristique nominale fréquence-température. Elle consiste à écrire que :

$$\left(\frac{\delta F_1}{F_1}\right)_A = -\left(\frac{\delta F_1}{F_1}\right)_R$$

et

$$(\delta b)_A = -(\delta b)_R$$

de sorte que l'hystérésis totale est :

$$\left(\frac{\delta F}{F}\right)_H = 2\left(\frac{\delta F_1}{F_1} - \delta b\ (\Delta T)^2\right).$$

Il n'y a pas discontinuité due à l'hystérésis aux points d'inversion $T_1 \pm \Delta T$ du cycle thermique si : $\delta F_1/F_1 - \delta b\,(\Delta T)^2 = 0$.

N°	T_1 (°C)	b $10^{-9}\cdot°C^{-2}$	K_H $10^{-10}\cdot°C^{-2}$	δv $10^{-10}\cdot°C^{-2}$	$K_H - \delta b$ K_H	τ (s)	$\tilde{\alpha}$ $10^{-6}\,s\cdot°C^{-1}$	Q 10^6
1	78,7	- 3,8	3,4	1,8	+ 0,47	480	0,25	1,134
2	78,8	- 3,8	1,88	1,2	+ 0,36	560	2,7	1,057
3	71,38	- 5,2	0,7	0,39	+ 0,44	350	1,2	1,083
4	86,3	- 2,4	0,68	1	- 0,47	440	3,1	1,051
5	76,8	- 4,2	0,55	0,9	- 0,63	470	4	1,07
6	70,8	- 5,4	0,35	0,45	- 0,28	490	3,3	1,20
7	70,2	- 5,4	0,35	0,47	- 0,34	300	2,2	1,12
8	84,4	- 2,2	0,104			123	0,71	1,23
QAS	62,6	- 7,1	0,64	0,66	+ 0,03	45	0,3	1,06

Tableau I
Résultats d'évaluation de huit résonateurs BVA miniatures SC et un résonateur QAS (CEPE).

Autrement dit, les deux caractéristiques thermiques aller-retour se coupent aux points d'inversion M et M' du cycle (Fig. 1). C'est dans ce cas seulement qu'il est possible de définir un coefficient K_H tel que :

$$\delta F_i / F_i = \delta b \, (\Delta T)^2 = K_H (\Delta T)^2 \quad \text{avec } K_H = \delta b \, .$$

Mais, le tableau I montre que la condition $K_H = \delta b$ n'est en général pas satisfaite et que les variations relatives $(K_H - \delta b) / K_H$ se situent entre + 0,47 et - 0,63 pour les résonateurs considérés.

Il est facile de montrer que cette différence implique une valeur non nulle de l'hystérésis aux points d'inversion du cycle thermique, c'est-à-dire une discontinuité des deux caractéristiques fréquence-température en ces points. Soit en effet $(\delta F/F)_M$, cette discontinuité au point M à la température $T_M = T_i + \Delta T$, le calcul de K_H est tel que :

$$K_H = \delta b + \frac{1}{(\Delta T)^2} \left(\frac{\delta F}{F} \right)_M .$$

Selon le signe de $(\delta F/F)_M$ par rapport à celui de δb, K_H peut être supérieur ou inférieur à δb, en concordance avec les résultats expérimentaux. Les figures 2a et 2b schématisent ces deux cas.

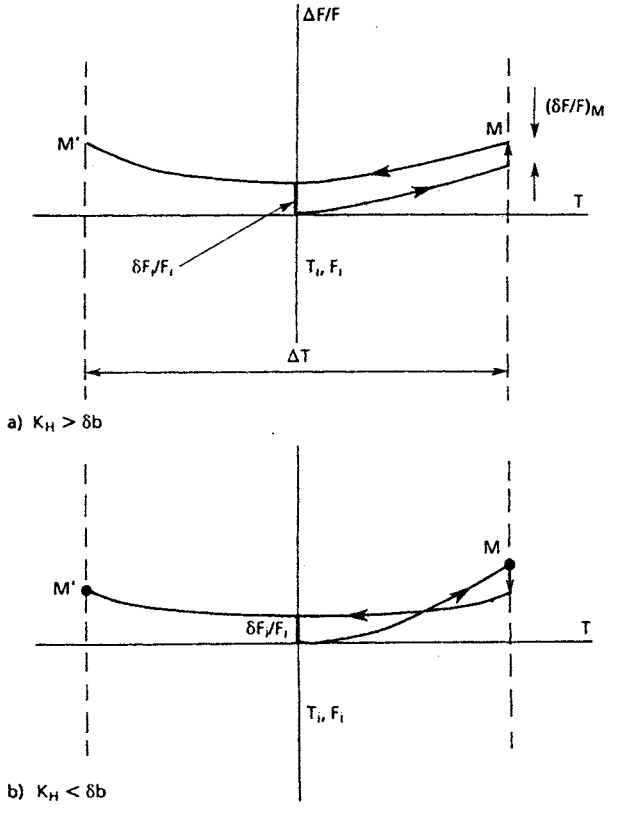


Figure 2
Discontinuités aux points d'inversion du cycle thermique.

Pour simplifier le diagramme et conformément à la relation précédente, le cycle "aller" est supposé confondu avec la caractéristique nominale :

$$\frac{\Delta F}{F} = b \, \Delta T^2$$

et l'hystérésis thermique se situe sur le cycle retour.

La suite de l'exposé a pour but de donner une réalité physique à cette discontinuité $(\delta F/F)_M$. Il apparaît que cette réalité peut se trouver dans l'action sur la fréquence des contraintes mécaniques internes au quartz.

3 - CONTRAINTES MECANIQUES INTERNES AU QUARTZ

3.1 - Définition des contraintes et effets sur la fréquence

Dans des conditions dites de référence pour la température T_0, trois types de contraintes initiales sont à considérer (Fig. 3).

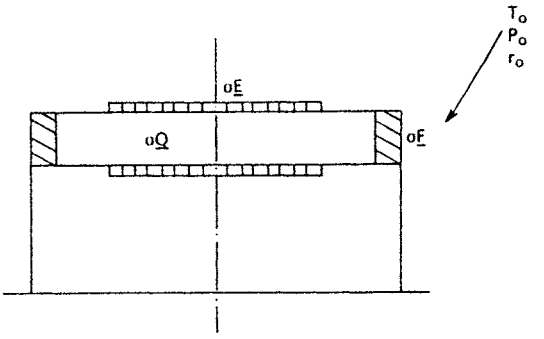


Figure 3 - Contraintes mécaniques internes.

a) Les contraintes d'origine du quartz dues à la synthèse et notamment aux défauts créés par la croissance ainsi que par les opérations ultérieures de coupe, usinage, surfaçage du quartz. Les dislocations jouent un rôle essentiel dans ces contraintes, notées σ_Q, et qui sont par conséquent créées dans le volume du quartz.

b) Les contraintes dues aux électrodes n'existent que dans les résonateurs à électrodes adhérentes. Ce sont des contraintes créées en surface. Elles sont notées σ_E.

c) Les contraintes transmises par les fixations du cristal résultent soit d'un effet ressort au montage du résonateur, soit des déformations dues à la mise sous vide du boîtier, soit de l'orientation de la pesanteur par rapport aux axes de référence du résonateur. Ces contraintes sont créées en surface et leur distribution dépend beaucoup du type de fixation. Elles sont notées σ_F.

Pour chacune des six composantes usuelles des contraintes en modèle tridimensionnel ou des trois composantes du modèle bidimensionnel, la contrainte initiale sera la somme :

$$\sigma_0 = \sigma_Q + \sigma_E + \sigma_F$$

et la fréquence de référence F_0 dépend de cet état initial de contrainte.

En dehors de cet état initial, il y a évolution des contraintes internes et cette évolution est une des causes majeures des dérives de fréquence des résonateurs. La présente communication montre de quelles manières ces contraintes et leurs évolutions interviennent dans les dérives irréversibles que sont l'hystérésis thermique et le vieillissement. Trois mécanismes d'évolution des contraintes initiales sont à considérer :

- les contraintes d'origine du quartz ainsi que celles dues à la métallisation et à la fixation du cristal sont normalement affectées par un processus de relaxation temporelle dû au fait que tout système physique tend à minimiser son énergie potentielle ; ce processus, partiellement responsable des dérives de

vieillissement, est activé thermiquement ;

- les variations de contraintes dues à la pression extérieure P_0 et à l'orientation θ du résonateur par rapport à la pesanteur ou, d'une manière générale, par rapport aux accélérations appliquées, provoquent des variations de fréquence par les effets bien connus des coefficients élastiques non linéaires ;

- les variations de température entraînent principalement des dilatations et des variations des coefficients élastiques. Il en résulte des déformations du cristal et une évolution des contraintes internes qui réagissent sur la fréquence par les mêmes effets que précédemment.

Si $\Delta\sigma_0$ désigne les variations des contraintes initiales dues aux diverses sollicitations qui viennent d'être évoquées, on rappelle que les variations résultantes de fréquence sont linéaires en fonction de $\Delta\sigma_0$: $\frac{\Delta F}{F} = \Sigma K_i \Delta \sigma_0$, i varie de 1 à 6 dans le cas général et il existe par conséquent six coefficients K_i qui dépendent beaucoup de la coupe du quartz. En cisaillement d'épaisseur il suffit, pour appliquer cette relation, de considérer les seules contraintes au centre du quartz.

L'hystérésis thermique, mesurée dans les conditions définies au paragraphe précédent, représente la part d'irréversibilité de ces dérives. Elle n'est par conséquent pas prise en compte par cette relation linéaire. On montre dans les paragraphes suivants comment l'hystérésis peut s'introduire à partir de l'évolution thermique des contraintes internes.

3.2 - Contraintes mécaniques dues aux fixations

La figure 4 schématise la disposition en support couronne avec ponts de liaison qui particularise le quartz utilisé à la fois pour les résonateurs BVA à électrodes non adhérentes et pour les résonateurs QAS à électrodes adhérentes. Les QAS sont réalisés par la CEPE et les BVA par le LCPE de Besançon qui est à l'origine de cette disposition.

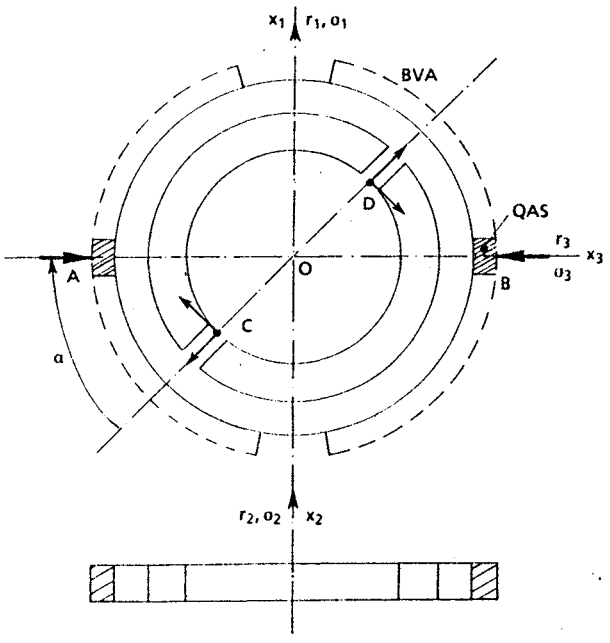


Figure 4 - Forces appliquées par les fixations dans les résonateurs BVA et QAS

Pour les résonateurs QAS, la fixation du quartz est assurée en deux points diamétraux par deux lamelles support. Tant pour l'état initial du quartz que pour les variations de pression ou d'accélération, les seules forces à considérer en A et B sont celles de compression ou d'extension diamétrale [3]. Celles-ci provoquent, en chaque point C et D d'encastrement des ponts, deux composantes dont la première est elle-même diamétrale, la seconde une force de cisaillement dirigée selon la tangente. Une première phase d'optimisation consiste à améliorer la conception des lamelles de suspension par un certain nombre d'artifices destinés à atténuer les forces diamétrales en A et B (parallélogramme, Muguet, etc...). Une seconde phase consiste à positionner les ponts C et D en dex points qui représentent le meilleur compromis entre le zéro de sensibilité de la fréquence aux forces diamétrales (Ratajski) et le zéro de sensibilité aux forces de cisaillement, ces deux points étant en général distincts. Dans ces conditions, la sensibilité accélérométrique est surtout due à la composante axiale Γ_2.

Pour les résonateurs BVA, l'optimisation des fixations se présente de manière différente [4, 5]. La distance quartz électrodes doit être maintenue rigoureusement constante et la fixation des électrodes ne peut être dissociée de celle du cristal. La solution considérée comme optimale est la fixation par deux demi-coques en quartz complétée par des ressorts et des brides assurant un encastrement périphérique rigide. Dans ces conditions, on montre que, si comme précédemment l'effet de la composante axiale Γ_2 de l'accélération appliquée est prépondérant, celui-ci doit être attribué à un fléchissement de la pastille de quartz selon une surface sphérique avec une amplitude maximale et un axe de symétrie au centre.

Dans les deux cas, les contraintes prépondérantes à considérer pour les variations de fréquence sont les contraintes σ_1 et σ_3 dans le plan du disque. On écrit par conséquent :

$$\frac{\Delta F}{F} = K_1 \Delta\sigma_1 + K_3 \Delta\sigma_3 .$$

Pour une coupe SC, les coefficients K_1 et K_2 sont de l'ordre de 10^{-12} N^{-1} m^2.

Dans les résonateurs QAS, les contraintes σ_1 et σ_3 sont constantes dans l'épaisseur de la plaque de quartz. Par contre, dans les résonateurs BVA et compte tenu du fléchissement de la plaque, ces mêmes contraintes sont nulles au centre du disque et maximales et de signe opposé à la périphérie.

Dans tous les cas, les relations entre les contraintes σ_1 et σ_3 et les forces appliquées sont linéaires et s'expriment en fonction de seuls paramètres géométriques. Une meilleure connaissance des contraintes est actuellement obtenue par l'utilisation de méthodes d'éléments finis.

3.3 - Contraintes mécaniques dues aux électrodes

Quelle que soit la procédure de métallisation utilisée, celle-ci laisse subsister des contraintes mécaniques dans le plan de la surface de séparation métal substrat. Une partie de ces contraintes, due à la différence des coefficients de dilatation électrodes-quartz est bien définie et prévisible. Mais les contraintes réelles sont très supérieures à celles qui résultent de cet effet. Elles doivent être attribuées à des phénomènes de nucléation ou d'interfaçage de gains pendant la croissance des électrodes et les cycles thermiques associés. La contrainte résultante est la superposition de celles créées par chaque électrode. Elle est aussi dirigée

dans le plan de la plaque. Mais cette contrainte n'est à prendre en considération que pour les résonateurs QAS réalisés en coupe AT, la coupe SC ayant été calculée pour minimiser cet effet. Elles ne seront pas considérées dans la suite de l'exposé.

3.4 - Défauts linéaires et contraintes internes associées

Les dislocations sont des anomalies du réseau cristallin du quartz dont les dimensions sont de l'ordre des distances interatomiques. Elles sont créées soit dans l'élaboration par synthèse hydrothermale, soit dans les déformations mécaniques ultérieures que le matériau peut subir. Elles constituent le défaut linéaire le plus répandu et le plus difficile à maîtriser. Une particularité essentielle des dislocations est leur mobilité sous action mécanique ou thermique. Ce déplacement s'effectue le long de plans appelés plans de glissement avec une résistance de type frottement sec.

Dans tous les cas, les dislocations sont associées à des contraintes internes. Un cristal sans dislocation serait parfaitement élastique et à contrainte interne nulle. Cette dualité dislocation/contrainte est conforme à la notion de compatibilité dans les milieux continus. Dans un tel milieu, il ne peut, en effet, y avoir déformation sans contrainte et ce même milieu ne peut supporter sans perdre sa compacité n'importe quel champ de déformation. Les déformations totales doivent en effet dériver d'un champ de déplacement et obéir aux équations de compatibilité dites de Saint-Venant. Pour qu'il en soit ainsi, le milieu crée des déformations élastiques additionnelles.

Il n'est évidemment pas possible de raisonner sur une dislocation isolée. Il faut aussi tenir compte des interactions élastiques à distance entre dislocations et des interactions entre dislocations et défauts ponctuels. Ce problème global s'aborde par l'énergie emmagasinée dans le cristal par la présence de dislocation et cette énergie fait que le cristal se trouve écarté de son état d'équilibre d'énergie minimale. Les mécanismes qui lui permettraient d'y revenir par restauration ou recristallisation le conduiraient à passer momentanément par des états encore plus écartés de l'équilibre et qui constituent des barrières de potentiel. En l'absence de toute activation mécanique ou thermique, le cristal reste alors dans un état stable bien que hors d'équilibre.

Il n'en est pas de même après un cycle thermique fermé de quelques degrés à quelques dizaines de degrés qui provoque une évolution de cet état stable. Sous l'effet d'une élévation de température et des dilatations qui en résultent, il s'instaure un nouvel état d'énergie et de contrainte dans des conditions susceptibles de provoquer un glissement des dislocations. Un certain nombre d'obstacles à ce glissement, dus par exemple aux interactions, peuvent en effet être franchis par activation thermique. Le déplacement s'effectue à volume constant, sans transport de matière, avec une amplitude au moins égale à la période du réseau cristallin, dans des conditions dissipatives assimilables à un frottement sec. Après retour de la température à sa valeur initiale, la non réversibilité des phénomènes énergétiques liés au franchissement d'obstacles provoque l'hystérésis intrinsèque du quartz mesurable si la résolution du dispositif chargé de la mesure est suffisante. Le paragraphe suivant montre comment ces phénomènes peuvent s'intégrer dans une étude plus globale des effets thermiques.

4 - EFFETS THERMIQUES STATIQUES DANS LES RESONATEURS

Une élévation lente de température appliquée au quartz provoque deux sortes de phénomènes.

a) Une dilatation thermique statique dont la similitude avec la dilatation mécanique est mise en évidence par la relation générale :

$$d\epsilon = \frac{1}{E} d\sigma + \alpha_L \, dT$$

E étant le module d'Young et α_L le coefficient de dilatation du quartz dans la direction ϵ. On remarque que le produit $E\alpha_L$ représente la contrainte mécanique qui produirait un effet identique à une variation thermique de 1°C. Avec $E = 7{,}8.10^{10} N/m^2$ et $\alpha_L = 13{,}7.10^{-6}/°C$, $\sigma_E \simeq 10^6 N/m^2$. Lorsque les actions mécaniques et thermiques sont séparées, cette relation s'intègre pour les dilations en $\epsilon = \Delta L/L = \alpha_L (T_M - T_i) = \alpha_L \Delta T$.

b) Une variation des coefficients élastiques - En adoptant pour ces coefficients la notation tensorielle C_{ij} telle que :

$$\sigma_i = C_{ij} \epsilon_j \text{ et } C_{ij} = \frac{d\sigma_i}{d\epsilon_j},$$

la variation des coefficients élastiques avec la température sera limitée pour la présente application à un développement linéaire :

$$C_{ij}(T) = C_0 + \frac{dC_{ij}}{dT} \Delta T \text{ avec } \Delta T = T_M - T_i.$$

On définit aussi un coefficient de variation relative :

$$\alpha_{C_{ij}} = \frac{1}{C_{ij}} \frac{dC_{ij}}{dT}.$$

Cette caractérisation concerne en réalité deux types d'effets thermiques difficiles à dissocier. Les coefficients C_{ij} sont en effet la somme d'un coefficient fondamental, lié au seul matériau sans défaut et parfaitement élastique, et d'un terme dit non linéaire qui introduit une variation du coefficient élastique avec les déformations statiques $\bar{\epsilon}_k$.

De manière très approchée : $C_{ij} \simeq C^*_{ij} + C_{ijk} \bar{\epsilon}_k$ ou, avec les coefficients de souplesse s : $C_{ij} \simeq C^*_{ij} + C_{ijk} s_{jk} \alpha_k$ α_k étant le coefficient de dilatation linéaire dans la direction k. Le coefficient de variation relative $\alpha_{C_{ij}}$ s'obtiendrait en divisant les deux nombres de cette relation par C_{ij}.

Il est facile de montrer que les contraintes qui permettent d'obtenir des dérivées thermiques de coefficient élastiques constantes et surtout indépendantes de la température sont celles qui varient linéairement en fonction de T :

$$\sigma(T) = \sigma_0 + \frac{d\sigma}{dT} \Delta T \qquad \text{avec } \frac{d\sigma}{dT} = c^{te}.$$

Toutes les contraintes thermiques dues aux dilatations rentrent en général dans ce modèle.

La variation de fréquence en fonction de la température résulte des variations des trois paramètres e, ρ, C_{ij} qui interviennent au premier degré dans l'expression générale de la fréquence :

$$F \simeq \frac{n}{2e} \sqrt{\frac{C_{ij}}{\rho}}$$

Les variations de C_{ij} incluant à la fois les effets thermiques intrinsèques dus à C^{*}_{ij} et ceux liés aux contraintes ou déformations agissant par la présence du terme non linéaire C_{ijk}. Pour une excursion thermique de part et d'autre du point d'inversion ne dépassant pas $2\,\Delta T = 10°C$ et dans un système d'axes centré en ce point, la sensibilité thermique de la fréquence se modélise par la relation quadratique déjà utilisée au § 2.3 :

$$\frac{\Delta F}{F} = b\,\Delta T^2 .$$

Le coefficient *b*, indépendant de la température, est fonction des coefficients de dilatation et des coefficients thermiques des coefficients élastiques. En particulier $\alpha_{c_{66}}$ intervient dans *b* par le terme $\frac{1}{4}\alpha^2_{c_{66}}$.

Lorsque la température provoque des variations de contrainte $\delta\sigma_i$ et, par conséquent, des variations de coefficient élastique qui n'entrent pas dans celles incluses dans *b*, ces variations doivent être traitées comme des contraintes parasites usuelles. Dans ces conditions, une expression plus générale de la sensibilité thermique statique est donnée par :

$$\boxed{\frac{\Delta F}{F} = b\,\Delta T^2 + \Sigma\, K_i\, \delta\sigma_i}$$

5 - HYSTERESIS THERMIQUE

L'hystérésis se définit comme une différence aller-retour dans la variation de fréquence provoquée par un cycle thermique fermé du type $T_i \nearrow T_M \searrow T_0$. On écrira :

$$\left(\frac{\delta F}{F}\right)_H = \left(\frac{\Delta F}{F}\right)_A - \left(\frac{\Delta F}{F}\right)_R = \frac{\delta F_i}{F_i} \text{ pour } T = T_i .$$

Les indices $_A$ et $_R$ désignent les variations de fréquence en cycle aller et retour.

A partir des études précédentes [1] et de celles des variations thermiques des contraintes exposées en § 3, deux causes majeures d'hystérésis doivent être dissociées et introduites dans cette relation.

La première, interne au quartz, est due au coefficient élastique dont les caractéristiques thermiques pour un cycle fermé (Fig. 5) montrent :

- *une discontinuité* $C'_1 - C'_0$ au point de retour du cycle,
- *une variation de pente* $\frac{dC}{dT}$ entre les trajets aller et retour.

Dans ces conditions, l'hystérésis totale à $T = T_i$ sur le coefficient élastique est :

$$(\delta C_{ij})_H = C'_1 - C'_0 + \left[\left(\frac{dC}{dT}\right)_R - \left(\frac{dC}{dT}\right)_A\right]\Delta T$$

C_{ij} représentant la tangente au diagramme contrainte-déformation pour une température T, une discontinuité sur C_{ij} correspond à un point de rebroussement du diagramme contrainte-déformation.

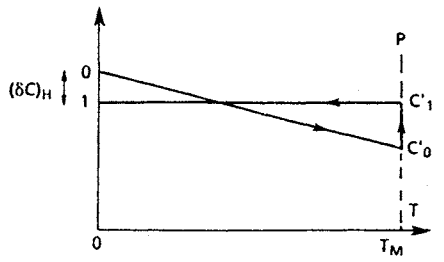


Figure 5 - Hystérésis du coefficient élastique.

La seconde, extérieure au quartz, est aussi une discontinuité au point de retour du cycle mais qui intervient cette fois sur les contraintes d'origine extérieure transmise par les supports. On note $\delta\sigma_F$ cette discontinuité (Fig. 6). Elle peut être attribuée à des phénomènes de plasticité dans les assemblages quartz-boîtier qui utilisent des collages ou des contacts avec frottements secs. Par contre, la pente de la caractéristique linéaire, soit $d\sigma/dT$ est supposée inchangée dans les trajets aller et retour.

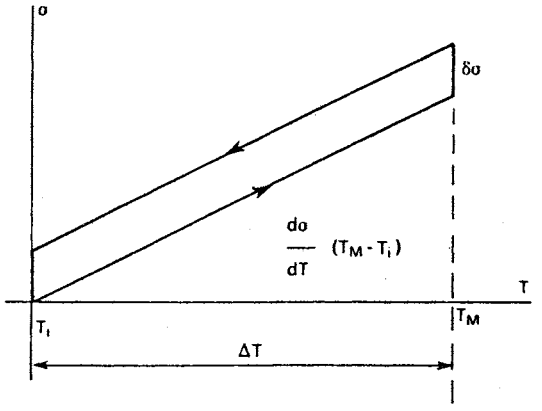


Figure 6
Hystérésis des contraintes d'origine extérieure.

L'hystérésis thermique sur la fréquence se déduit directement de l'hystérésis sur les coefficients élastiques et sur les contraintes extérieures au quartz, ces paramètres étant introduits dans l'expression générale de la sensibilité thermique statique établie au paragraphe précédent. La discontinuité $C'_1 - C'_0$ serait en effet équivalente à une variation de contrainte $\delta\sigma_Q$ telle que :

$$\delta\sigma_Q = \frac{1}{C_{ijk}\, s_K}\left(C'_1 - C'_0\right)$$

Dans la direction i, la variation totale de contrainte interne est :

$$\delta\sigma_i = \delta\sigma_Q + \delta\sigma_F .$$

Les variations $\delta\alpha = \frac{1}{C_{ij}}\left[\left(\frac{d\,C_{ij}}{dT}\right)_R - \left(\frac{d\,C_{ij}}{dT}\right)_A\right]$ provoque sur le coefficient thermique *b* une variation $\delta b = \frac{1}{2}\,\alpha\,\delta_\alpha$.

Dans ces conditions, la valeur de l'hystérésis thermique à la température $T = T_i$ après retour d'un cycle thermique ΔT est (Fig. 7) :

$$\boxed{\frac{\delta F_i}{F_i} = \Sigma\, K_i\, \delta\sigma_i + \delta b\,(T - T_i)^2}.$$

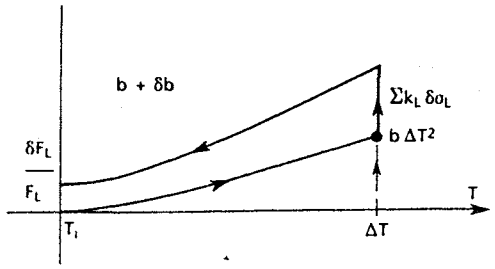


Figure 7 - Hystérésis thermique sur la fréquence.

La discontinuité $\Sigma K_i \delta\sigma_i = \Sigma K_i (\delta\sigma_{iQ} + \delta\sigma_{iF})$ représente par conséquent l'hystérésis à $T_M = T_i + \Delta T$. Cette discontinuité peut être nulle soit si $\delta\sigma_{iQ}$ et $\delta\sigma_{iF}$ sont simultanément nuls, soit si ces deux paramètres sont égaux et de signe contraire. Dans ces deux cas seulement :

$$\frac{\delta F_i}{F_i} = \delta b\,(T_M - T_i)^2 = K_H\,\Delta T^2 .$$

Les caractéristiques thermiques aller-retour se coupent alors à $T = T_i + \Delta T$.

S'il en est autrement, il existe une différence $\delta b - K_H$ telle que :

$$\boxed{\delta b - K_H = -\frac{1}{\Delta T^2}\,\Sigma K_i\,(\delta\sigma_{iQ} + \delta\sigma_{iF})} .$$

Il n'est pas possible de préjuger a priori du signe de la sommation selon les règles du calcul tensoriel. Pour les contraintes σ_F il a été montré (§ 3.2) que les valeurs i = 1 et i = 3 sont à considérer. Pour les contraintes σ_Q cette détermination est plus complexe et une information globale sur le signe de $\Sigma K_i \delta\sigma_i$ ne peut être obtenue que par les résultats expérimentaux. Il en est de même d'une relation entre δb et $\delta\sigma_{iQ}$, les mêmes phénomènes physiques liés aux défauts du quartz étant à l'origine de ces deux types d'écarts. Les études sur l'hystérésis thermique se poursuivent actuellement dans cette direction. Bien que des progrès significatifs aient été faits dans la connaissance du phénomène, il est encore difficile d'obtenir des résultats fiables sur l'importance relative des deux types d'hystérésis en fonction du résonateur ou du matériau. Les écarts importants entre K_H et δb, constatés sur les résonateurs BVA, pourraient être attribués à une origine extérieure compte tenu de la complexité d'assemblage de ces résonateurs. Mais l'hystérésis globale sur les BVA est du même ordre que celle constatée sur les autres types de résonateurs.

REFERENCES

[1] - J. Beaussier, "Hystérésis thermique dans les résonateurs et les oscillateurs à quartz", 2ème Forum Européen Temps-Fréquence, Neuchâtel 1988, T.P. ONERA n° 1988-24.

[2] - F. Deyzac, "Oscillateurs à quartz thermostaté compensé", 1er Forum Européen Temps-Fréquence, Besançon 1987, T.P. ONERA n° 1987-29.

[3] - J.P. Aubry, B. Bonnotte, J.C. Craveur, F. Deyzac, D. Janiaud, "Optimisation de l'orientation des résonateurs autosuspendus", 2ème Forum Européen Temps-Fréquence, Neuchâtel 1988, T.P. ONERA n° 1988-37.

[4] - R. Delaite et J.P. Valentin, "Compensation des effets mécaniques dans un cristal de coupe SC", Revue de Physique Appliquée, Février 1987.

[5] - R. Delaite, "Précision de réalisation pour résonateurs à quartz HQ", Revue de Physique Appliquée, Janvier 1989.

FREQUENCY REFERENCE FOR BOTH SHORT AND LONG TERM STABILITY MEASUREMENTS

J. Groslambert

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

In this paper it will be shown that a frequency reference stable at both short and long term can be built around a cesium standard and a quartz oscillator. The state of the art in frequency stability research show that improved quartz oscillator characteristics in the short term domain, associated with good results in the long term domain of a cesium standard, can give a unique reference for short and long term stability measurements. This apparatus uses a digital servoloop, with a large time constant, to lock a quartz oscillator to the output frequency of a cesium standard. A servoloop with a large time constant has a very small locking range. This difficulty is removed by using a sequentially variable time constant - i.e. at the start of the operation the apparatus uses a small time constant (a few seconds) and, when the locking is effective, this time constant is automatically doubled, and so on, up to a selected constant. This servoloop is a phase lock loop which maintains the signal coming from the oscillator in quadrature phase with respect to the mean phase of the signal coming from the cesium standard over a time τ. The time constant τ of the servoloop must be choosen to conserve the good short-term characteristics of the quartz oscillator and to transfer the long term stability of the cesium. With an oscillator exhibiting a flicker noise floor of 5×10^{-13}, this time constant depends on the quality of the cesium used. It ranges from about 200 seconds at best to 1 600 seconds at worst. Stability better than 5×10^{-13} can be obtained in the 1 to 1 000 second range, depending on the flicker noise of the quartz that is used, while keeping the stability of the cesium for larger time intervals.

Introduction

The characterisation of the frequency stability of an oscillator is very important because the ultimate resolution of many instruments depends on the frequency stability of the reference oscillators.

A cesium beam standard exhibits poor relative frequency stability over an integration time in the 1 - 1 000 seconds range compared to good Xtal oscillators, but the latter have a long term aging characteristic inferior to that of the cesium beam standard.

Short term measurements must be made with a good Xtal oscillator as reference, while long term measurements require the cesium as reference.

The purpose of this paper is to realize an unique reference which exhibits both the short term stability of the quartz oscillator and the long term stability of the cesium.

State of the art

At the present time there is no available reference which exhibits both the best short term stability over an integration time in the 10^{-3} - 1 000 s and the best long term (over 1 000 s) stability.

The state of the art for commercial units is reviewed on Fig. 1, which gives the short term reference characteristics obtained with the best 5 MHz Xtal oscillator available on the market and the long term reference characteristics for cesium standards of different manufacturers (CS1, CS2, CS3).

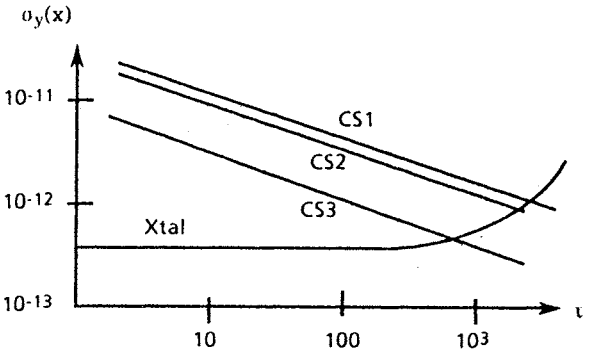


Fig. 1 : State of the art

As part of the DORIS* project, the Centre National d'Etudes Spatiales (CNES) needed a reference signal, used as master clock for orbitography measurements, with two requirements, first : a relative stability better than 10^{-12} over a 10 s integration time ; second : the long term stability of the cesium to date the events. Not even the best of the cesium standards satisfies the first requirement.

However, a Xtal oscillator phase locked to the output signal of a cesium reference, of even the worst quality can satisfy the two requirements.

The better reference is obtained with a time constant of the servoloop corresponding to the intersection of the two curves in Fig. 1 giving the stabilities of both the quartz oscillator and the cesiums. This time constant is in the range of 300 s for the better cesiums and 2 000 s for the worst.

A corresponding time constant of 2 000 s is realizable, but is not usable. A numeric locking system solves the problem. A difficulty arises because the locking frequency range of the servoloop is inversely proportional to its time constant. To give an example, for a time constant of 1 000 s the relative frequency locking range is in order of few 10^{-11}, which makes locking impossible.

This difficulty is avoided by using a sequentially variable time constant i.e. the apparatus starts with an initial time constant of few seconds to increase the capture range and when the locking is effective this time constant is automatically doubled and so on, up to a selected value.

---

* DORIS : Détermination d'Orbite et Radiopositionnement Intégrés par Satellite

The servoloop

The servoloop contains a particular phase comparator which measures the integrated phase shift between the 5 MHz cesium signal and the Xtal oscillator signal to be locked, during a time constant T. After each measurement the contents of the integrator is transfered via a temporary memory to a DAC, the output of which is applied to the electronic tuning of the oscillator.

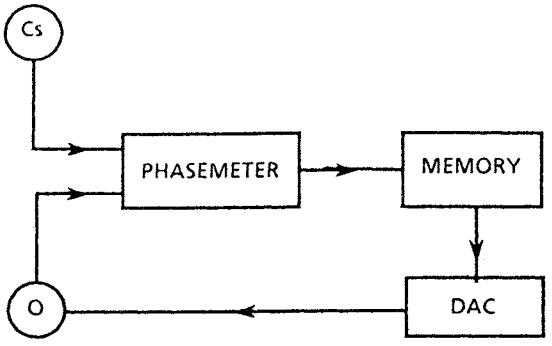


Fig. 2 : Block diagram of the servoloop

The required resolution of the digital phasemeter is determined by the need for a measurement stability of 10^{-12} on an integration time of 10 s.

$$\text{Phasemeter resolution} = 10^{-12} \times 5 \times 10^{6} \times 2\pi \times 10 = 3 \times 10^{-5} \text{ radian}$$

The clock frequency needs to detect a variation of 3×10^{-5} radian. That is

$$F_{clock} = 5 \times 10^{6} \times 2\pi / 3 \times 10^{-5} \simeq 10^{12} \, Hz$$

This clock frequency is not realizable in practice and to reduce the clock frequency an heterodyne phasemeter is constructed with a local oscillator. The two signals coming from the cesium and the oscillator are down converted by mixing with this local oscillator (Fig. 3).

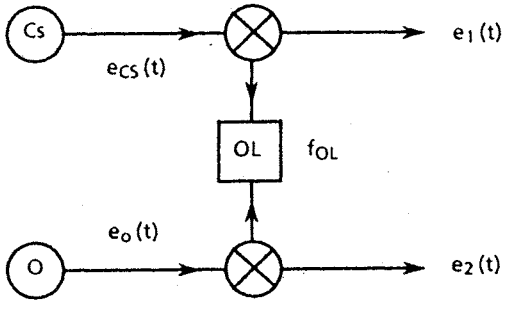


Fig. 3 : Principle of the heterodyne phasemeter

In this case the total phase difference between the two HF signals ($e_{CS}(t)$, $e_o(t)$) is transfered to the LF signals ($e_1(t)$, $e_2(t)$).

The maximal frequency of the beat ($f_{CS} - f_{OL}$) of the $e_1(t)$ and $e_2(t)$ LF signals is set by the resolution obtained with a 5 MHz as clock frequency :

$$f_{beat\ max} = 3 \times 10^{-5} \times 5 \times 10^{6} / 2\pi = 23 \, Hz$$

Therefore the local oscillator must be have a frequency given by :

$$5 \, MHz - 23 \, Hz < F_{LO} < 5 \, MHz + 23 \, Hz$$

This local oscillator frequency is obtained by a synthesis which satisfies the relation :

$$f_{LO} = f_{CS} + \frac{f_{CS}}{998} - \frac{f_{CS}}{1\,000}$$

$$f_{OL} = f_{CS}(1 + 2 \times 10^{-6}) = f_{CS} + 10 \, Hz \quad \text{if} \quad f_{CS} = 5 \, MHz$$

The schematic synthesis of the local oscillator is shown in Fig. 4.

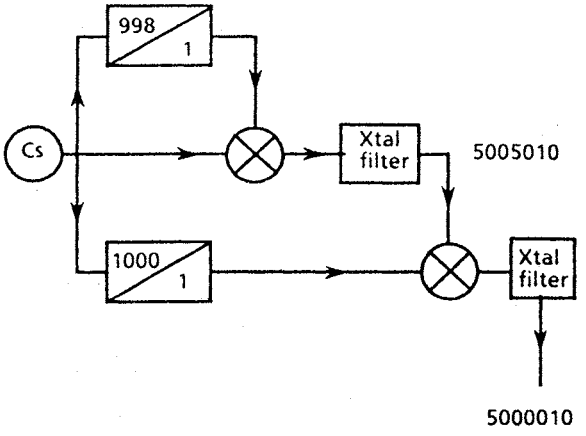


Fig. 4 : Synthesis of the local oscillator

The phasemeter principle is based on time duration measurements over N periods of the signals $e_1(t)$ and $e_2(t)$.

In each path, B_1 and B_2 in Fig. 5, a flip-flop is activated at the start of the count and remains in the one state during the N periods of the signal counted.

An integrator is constructed around a up-down counter, where the counts are conditioned by the state of the B_1 and B_2 flip-flops.

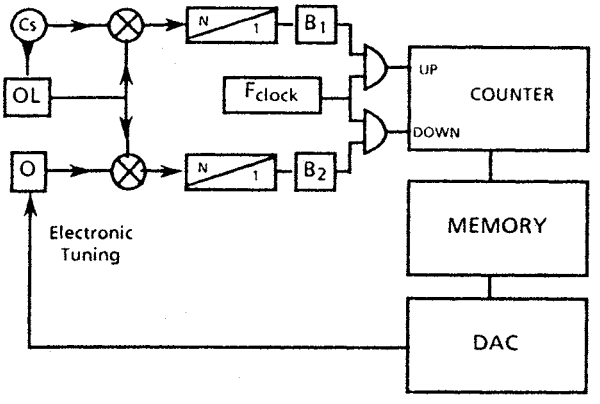


Fig. 5a : Schematic of the servoloop

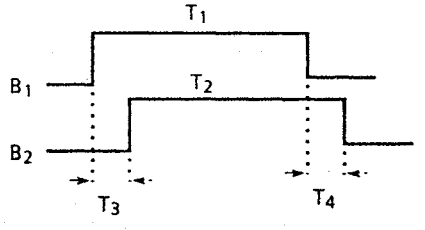


Fig. 5b : Time diagram of the state of B_1 and B_2 flip-flops

In reality the up down counter is up during time T_3 and down during time T_4. When $T_3 = T_4$ the contents of the integrator do not change and the Xtal oscillator has the same mean frequency than the cesium during a time τ.

This servoloop is a frequency lock loop. In fact the conditions $T_1 = T_2$ imposed $T_3 = T_4$ but does not imposed $T_3 = T_4 =$ cste, a phase shift can be allowed.

To obtain a phase lock loop, a phase difference corresponding to the time T_3 between the two flip-flops B_1 and B_2 signals must be imposed : the quadrature has been chosen. For this operation two different clocks at frequencies f_1 and f_2 are needed for the down count and the up count. The reference phase is introduced for each measurement by prepositionning the up down counter with extra counts of frequencies $f_1/4$ and $f_2/4$ respectively during one period T_0 of the $e_1(t)$ signal, which is equivalent to up and down count of frequencies f_1 and f_2 during a quarter of the T_0 corresponding to the quadrature.

The lock is effective when the contents of the integrator do not change. Therefore :

$$T_0 \frac{f_1}{4} - T_0 \frac{f_2}{4} + T_4 f_2 - T_3 f_1 = 0$$

Only one stable solution exists :

$$T_3 = T_4 = \frac{T_0}{4}$$

When the lock is effective a logic command automatically doubles the value of the integration time of the phasemeter by changing the N ratio of the divider to 2N, and so on, up to a selected power of 2 chosen by the user.

The transfert function of the servoloop

This servoloop is a sampling servoloop with a digital phasemeter, a sample-and-hold memory, and a digital-to-analog converter (Fig. 6).

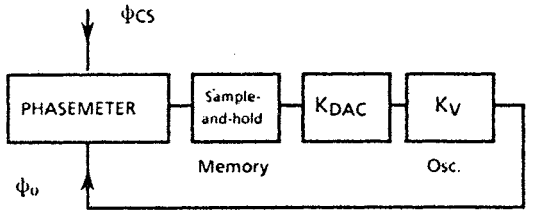


Fig. 6 : Block diagram of the digital servoloop

The theory of the sampled system uses the Z transform [1,2]. Let $\phi_{CS}(s)$ be the frequency fluctuation of the cesium, $\phi_0(s)$ the frequency fluctuation of the oscillator, where $s = j\omega$: the error signal takes the form

$$E(s) = \phi_{CS}(s) - \phi_0(s)$$

The transfert function of the sampling phasemeter which is an integrator of the error signal is given by the Z transform by

$$P(Z) = K_i \frac{Z}{Z - 1}$$

where $Z = e^{sT}$, T the sampling period and K_i the integrator coefficient.

Here, $e^{sT} = (1 + sT/2) / 1 - sT/2)$, in the Padé's approximation [2]. This $P(Z)$ function can be expressed in s terms as :

$$P(s) = K_i \frac{1 + sT/2}{sT}$$

The sample-and-hold memory is equivalent to a circuit with a transfert function : [2]

$$M(s) = \frac{1}{1 + sT/2}$$

The open loop gain is

$$G(s) = P(s).M(s).K_D K_V$$

where K_D is the converting coefficient of the DAC and K_V the sensitivity of the electronic tuning of the oscillator.

Finally

$$G(s) = K_i K_D K_V / sT = G_0 / sT$$

In a digital servoloop the "continous gain" G_0 cannot have a large value, in analogic servoloop. In fact $G_0 = 0.4$.

For the closed loop the frequency fluctuations of the locking oscillator can be expressed in terms of the frequency fluctuations of the cesium signal $\phi_{CS}(f)$ and the frequency fluctuations of the free oscillator $\phi_{o\,free}$ by the relation :

$$\Phi_{o\,lock}(f) = \left(\frac{G(f)}{1 + G(f)}\right) \Phi_{CS}(f) + \left(\frac{1}{1 + G(f)}\right) \Phi_{o\,free}(f)$$

In spectral density terms :

$$S_{\delta v\,lock}(f) = \left|\frac{G}{1+G}\right|^2 S_{\delta v_{CS}}(f) + \left|\frac{G}{1+G}\right|^2 S_{\delta v_{o\,free}}(f)$$

Let $H_{CS}(f)$ be the transfert function for the cesium frequency fluctuations and $H_{o\,free}(f)$ the transfert function for the free oscillator frequency fluctuations. These two quantities have the limiting values in Table I for $2\pi fT > 1$ and $2\pi fT < 1$.

	$2\pi fT/G_0 > 1$	$2\pi fT/G_0 < 1$
$H_S(f)$	$\left\|\frac{1}{1 + 2\pi fT/G_0}\right\|^2 \to 0$	$\to 1$
$H_{o\,free}(f)$	$\to 1$	$\to 0$

Tableau I

The transfert function of the servoloop transforms the white frequency noise of the cesium to a random walk frequency noise for frequencies higher than the cutoff frequency corresponding to the time constant T (Fig. 7).

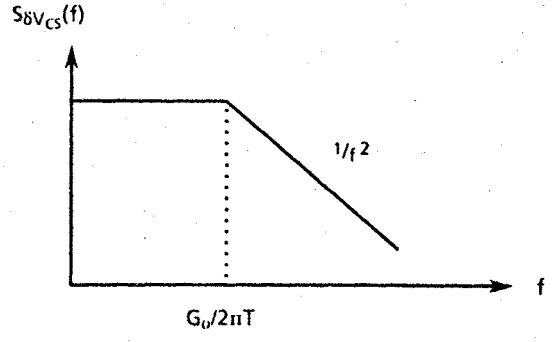


Fig. 7 : Filtering of the white frequency noise of the cesium by the servoloop

In time domain, the slope of the Allan variance exhibits a $\tau^{1/2}$ behavior for integration times smaller than T (Fig. 8).

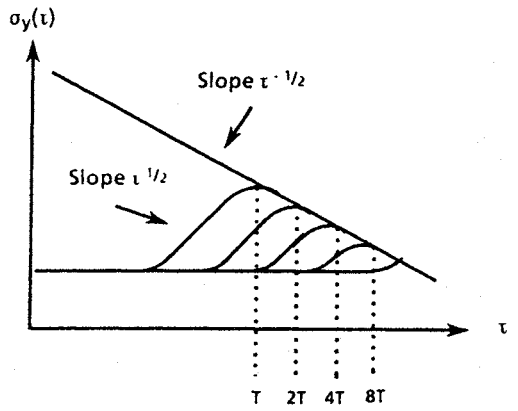


Fig. 8 : Allan variance of the locked oscillator for different time constant T

This system has been tested at the Centre National d'Etudes Spatiales versus a hydrogen maser. The measured frequency stability (Fig. 9) shows that the signal issued from this apparatus can be used as both a short and a long term reference.

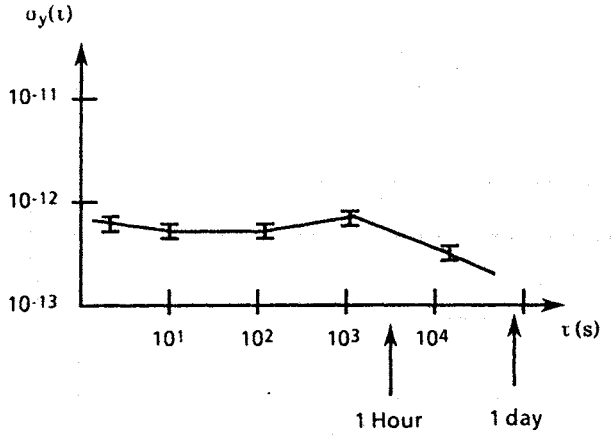


Fig. 9 : Allan variance of the locked oscillator versus a hydrogen maser

Conclusion

This system solves the problem of the unique reference in stability measurements.

This unit used as a complementary apparatus to the cesium beam standard cleans up the poor performance of this latter in the short terms domain.

When making aging measurements of a lot of a large number of oscillators with this system it is not necessary to use a large measurement time on the counter to have a good resolution. Consequantly, a larger number of oscillators can be tested in the same time, with a better resolution than when using only the cesium as reference.

References

[1] J.C. Radix, "Introduction au filtrage numérique", Eyrolles.

[2] Hansruedi Bühler, "Réglages échantillonnés", vol. 1, Traitement par la transformation en Z. Presses Polytechniques romandes.

COMPARATEUR DE PHASE ULTRA FAIBLE BRUIT POUR LES FUTURS ETALONS DE FREQUENCE

R. BARILLET

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

RESUME

Les étalons de fréquence de nouvelle génération seront basés sur des dispositifs atomiques refroidis, soit directement par cryogénie, soit encore par irradiation laser. La mesure de la stabilité de fréquence de tels étalons nécessitera des bancs de mesure améliorés, qui n'existent pas actuellement.

Nous présentons les limites ultimes des systèmes à deux mélangeurs fonctionnant à température ambiante et à 5 MHz, ainsi que les résultats expérimentaux correspondants.

Nous montrons la possibilité d'améliorer leur plancher de mesure, sans refroidissement mais avec diverses précautions. Des stabilités relatives de fréquence voisines de 5. 10^{-16} (τ = 1 s), 8. 10^{-18} (τ = 100 s) et 10^{-18} (au-delà de 1000 s) seront alors mesurables à 100 MHz.

ABSTRACT

Future frequency standards will be based on atomic systems that will be cooled, either by cryogeny or by laser-cooling. The measuremnt of their frequency stability will need improved measurement systems, which do not exist presently.

We point out the ultimate limits of the "dual mixer" systems operating at ambiant temperature and 5 MHz. Experimental results are presented.

We indicate the possibility of improving their measurement floor without cooling but with several cautions. In such conditions, it will be possible to measure frequency stabilities of the order of 5. 10^{-16} (τ = 1 s), 8. 10^{-18} (τ = 100 s) and 10^{-18} or better (τ > 1000 s) at 100 MHz.

INTRODUCTION

Le système à 2 mélangeurs avec oscillateur décalé[1] est un banc de mesure très commode lorsqu'on souhaite mesurer la stabilité de fréquence d'horloges quasi-synchrones. Toutefois, les performances usuelles de ce type de banc de mesure se situent au mieux à quelques $10^{-14}\ \tau^{-1}$ en court terme pour une bande passante de l'ordre de 10 Hz, lorsqu'on effectue la comparaison entre 2 signaux à la fréquence "Standard" 5 MHz.

L'apparition de quelques masers cryogéniques depuis 1986 et l'efficacité prouvée du refroidissement d'atomes et d'ions par laser devraient bientôt conduire à des besoins de l'ordre de 10^{-18} en plancher de mesure de la stabilité relative de fréquence. Nous présentons ci-après le niveau de performances que l'on peut mesurer à température ambiante et à 5 MHz, puis les conditions complémentaires qui permettront d'atteindre des planchers de stabilité de l'ordre de 10^{-18}. Simultanément, ce dispositif pourra être utilisé pour mesurer avec une précision exceptionnelle l'écart de phase entre deux étalons de fréquence.

I. POSSIBILITES ULTIMES A 5 MHz ET A TEMPERATURE AMBIANTE

I.1. Limites liées aux amplificateurs d'isolement

Nous supposons que les amplificateurs d'isolement possèdent la densité spectrale de phase S_φ donnée par la formule (1), ce qui correspond à un amplificateur peu bruyant, dans le domaine flicker en particulier

$$S_\varphi(f) \simeq 10^{-18,0} + \frac{1}{f}\, 10^{-15,0}\ \mathrm{rad}^2.\mathrm{Hz}^{-1} \qquad (1)$$

Ces composantes de bruit vont induire une limitation absolue de la stabilité de fréquence à court terme. Les termes "blanc de phase" et "flicker de phase" donneront les stabilités respectives $\sigma_{y_2}(\tau)$ et $\sigma_{y_1}(\tau)$ ci-dessous :

$$\sigma^2_{y_2}(\tau) = \frac{3h_2 f_c}{8\pi\tau^2} \qquad (2)$$

$$\sigma^2_{y_1}(\tau) = \frac{3h_1 \ln(2\pi f_c \tau)}{4\pi^2\tau^2} \qquad (3)$$

avec les notations habituelles de la métrologie des fréquences [2]. Dans le cas où f_c = 10 Hz, et avec les valeurs de h_2 et h_1 correspondant à la densité spectrale (1) à 5 MHz, on obtient $\sigma_{y_2}(\tau) = 2{,}2.\ 10^{-16}\tau^{-1}$ et $\sigma_{y_1}(\tau) \simeq 3{,}5.\ 10^{-15}\ \tau^{-1}$ en court terme. On constate que la composante flicker de phase est lourdement pénalisante par rapport au bruit blanc de phase. Ceci est lié à la faible bande passante considérée (10 Hz), dans laquelle la composante flicker est très nettement prépondérante.

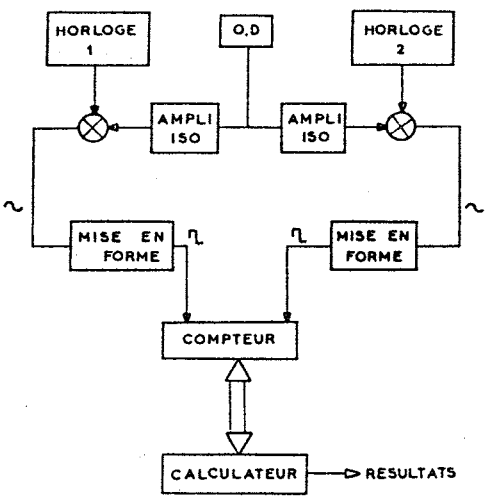


Fig. 1. Schéma de principe du système de mesure à deux mélangeurs.

I.2. Limites liées au banc de mesure

Ce banc représenté sur la figure 1, comporte plusieurs éléments susceptibles d'introduire des bruits additifs, donc de limiter le plancher de bruit du système.

I.2.1. Effet du bruit de l'oscillateur décalé (O.D.)

On montre [1][3] que la contribution de bruit de l'O.D. est notablement rejectée sous réserve que le temps δt, séparant les fronts (montants par exemple) des 2 battements, soit très petit par rapport à la période T_B de ces battements. La vérification expérimen-

tale confirme cette réjection de façon très nette. Toutefois, dans une configuration réelle de mesure, δt varie et surtout dérive lentement en raison de l'écart de fréquence (aussi faible soit il) entre les horloges sous test. Dans ces conditions, δt va évoluer progressivement de 0 à T_B, puis revenir à zéro et ainsi de suite. Pour des mesures sur des temps courts ($\tau \sim$ quelques secondes), l'ajustement initial de δt au voisinage de zéro va évoluer suffisamment peu pour que le bruit de l'O.D. ne soit pas réintroduit. Par contre, pour des mesures sur des temps relativement longs ($\tau \gtrsim 10\,000$ s), l'obtention de 100 échantillons de 10 000 s par exemple nécessitera une durée d'acquisition de 11,5 jours ; à la fin de la mesure, δt sera généralement passé une (ou plusieurs) fois par des valeurs de l'ordre de T_B. En d'autres termes, le plancher de mesure sera très fluctuant (2 ordres de grandeur par exemple) selon l'instant, et globalement il se ramènera au voisinage de sa plus mauvaise configuration. Pour remédier à cet inconvénient, on peut procéder de 2 façons :

- ou bien l'oscillateur décalé reste libre, et il faut alors ajuster très finement l'écart de fréquence entre les deux horloges sous test, de façon que δt soit toujours très petit par rapport à T_B aux instants de prélèvement de l'information (δt pouvant prendre des valeurs voisines de T_B à d'autres instants). Ceci est faisable, mais n'est pas toujours le plus commode, il faut effectuer séparément les caractérisations pour des temps courts et pour des temps longs. De plus, les fluctuations lentes de fréquence de l'O.D. par rapport aux horloges sous test risquent également de perturber la stabilité mesurable.
- ou bien on asservit l'O.D. sur l'une des horloges sous test, avec la constante de temps d'asservissement optimale. Dans ce cas, l'effet du bruit de l'O.D. sera, au pire, voisin du bruit d'une horloge sous test, et le plancher de mesure du système sera meilleur que ou voisin de la stabilité des horloges à tester lorsque δt variera de 0 à T_B, et les risques liés aux fluctuations de l'O.D. libre seront supprimés. L'étude détaillée de cet aspect (asservissement) n'est pas abordée dans le présent article. Elle fera l'objet d'un travail ultérieur.

I.2.2. Effet du bruit de l'électronique de traitement du signal de battement en basse fréquence

L'électronique de traitement du signal de battement a pour fonction d'accroître la pente du signal de battement pour rendre négligeable l'erreur de trigger du compteur, tout en délimitant une bande de fréquence f_c. Ceci doit être accompagné d'une dégradation minimale du rapport signal sur bruit "initial". Pour nous, cela signifie en fait que la perturbation de phase équivalente au bruit du banc de mesure, et rapportée à la fréquence ν_0 des horloges, c'est à dire avant le mélangeur, doit être la plus faible possible.

Enfin, il ne faut pas oublier que le compteur, qui effectue la mesure, a une bande passante très élevée ($\sim$ 100 MHz), et que la mesure effectuée prendra donc en compte la totalité du bruit superposé au signal utile dans 100 MHz de bande, et pas seulement dans la bande f_c que nous avons choisie.

La dégradation minimale vis à vis de la phase des signaux d'horloges est réalisée lorsque :

- le mélangeur utilisé est le plus performant en bruit, en particulier pour le bruit flicker de phase, dont l'effet est prépondérant (4)
- le mélangeur est attaqué par des signaux suffisamment puissants et chargé de façon à présenter le plus grand gain possible (4)
- l'amplificateur opérationnel qui suit le mélangeur est très performant en bruit blanc et en bruit flicker. Nous avons choisi d'utiliser des LT 1028, qui font partie des meilleurs amplificateurs du point de vue du bruit basse fréquence : bruit blanc à 0,9 nV/$\sqrt{\text{Hz}}$ et bruit flicker à 2 nV/$\sqrt{\text{Hz}}$ à f = 1 Hz. Dans cette configuration de bruit et avec les conditions de gain de mélangeur mentionnées en II.1. (K = 0,3 V/rad.), la perturbation de phase, sur le signal 5 MHz, équivalente au bruit d'amplitude de l'amplificateur LT 1028 s'établit à $3.\ 10^{-9}$ rad./$\sqrt{\text{Hz}}$ en bruit blanc de phase et $6{,}6\ 10^{-9}$ rad./$\sqrt{\text{Hz}}$ à f = 1 Hz en bruit flicker de phase. Pratiquement, la contribution "flicker" est plus faible que celle des mélangeurs, et la limite théorique associée au bruit blanc des LT 1028 s'établit, d'après la formule (2), à $\sigma_y(\tau) = 6{,}6\ 10^{-16}\ \tau^{-1}$ pour f_c = 10 Hz. Parallèlement, le bruit flicker de phase, pour des mélangeurs HP 10514, est de l'ordre de - 150 dBc/Hz à f = 1 Hz dans nos conditions d'utilisation (cf. II.2. 5e configuration). La limitation correspondante de la stabilité est la même que celle mentionnée précédemment pour les amplificateurs d'isolement, c'est à dire $3{,}5.\ 10^{-15}\ \tau^{-1}$ en court terme.

En ce qui concerne l'effet de la grande bande passante du compteur, on l'oublie fréquemment à cause de raisonnements du type : "le bruit ramené à l'entrée est...". Ici, il faut s'intéresser à la répartition spectrale du bruit existant à la sortie de l'électronique de traitement (entrée du compteur) lorsque le signal est au voisinage de zéro. Considérons les figures 2A, 2B et 2C qui représentent cette répartition spectrale dans le cas où le filtrage f_c est situé respectivement à la sortie du mélangeur, après un préamplificateur ou après la mise en forme. Sur la figure 2A, on constate que le rapport S/B global provient essentiellement du bruit large bande de la mise en forme, ce qui peut conduire à des résultats apparemment aberrants lorsqu'on veut améliorer les performances (cf. II.1). Sur la figure 2B, on constate que le rapport S/B global provient pour partie du bruit des signaux d'horloge et de l'étage d'entrée, pour partie du bruit large bande de la mise en forme. Dans cette configuration nous constaterons effectivement une amélioration très nette des performances. Sur la figure 2C, on obtient un rapport S/B plus favorable encore (en principe), mais la mesure est alors notablement dégradée par la faible pente du signal de sortie (augmentation de l'erreur de trigger du compteur).

La configuration "idéale" correpond à une répartition spectrale identique à celle de la figure 2C, associée à une forte pente lors du passage à zéro. Cela est théoriquement possible au moyen d'un filtrage associé à la mise en forme (et non situé après la mise en forme).

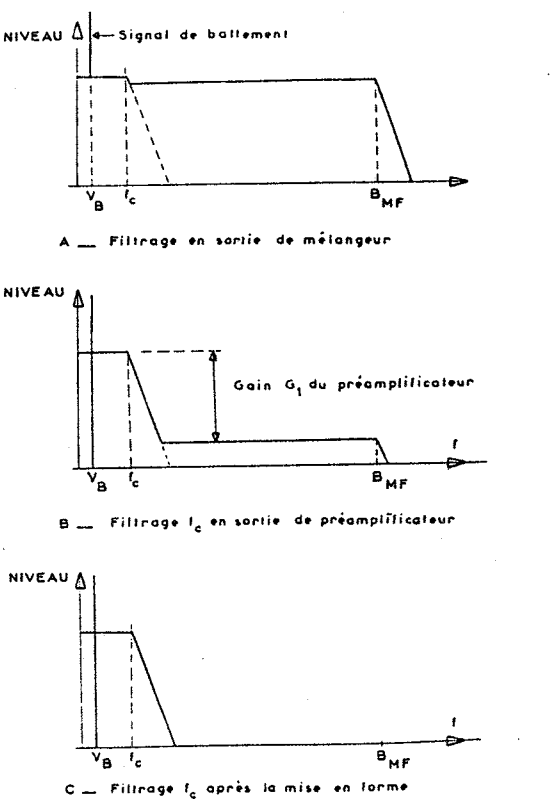


Fig. 2. Répartition spectrale du bruit existant au voisinage de zéro sur le signal d'attaque du compteur.

I.2.3. Estimation simple de la dégradation apportée par le bruit large bande existant à l'entrée du compteur

La figure 3 représente de façon assez générale la chaine électronique de traitement du signal délivré par un mélangeur.

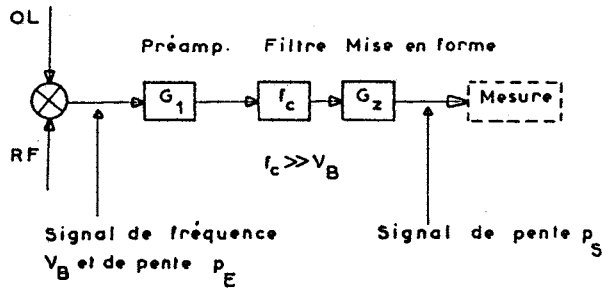


Fig. 3. Traitement du signal BF délivré par le mélangeur.

On simplifie le raisonnement en considérant uniquement les bruit BF blancs de tension, et en supposant que l'erreur de trigger du compteur est négligeable. Soient b_M^2 la densité spectrale de bruit blanc, en sortie de mélangeur, due à la superposition des effets résiduels des signaux (OL, RF) et des bruits du mélangeur et de l'amplificateur de tête, b_f^2 la densité spectrale de bruit du filtre, G_1 et G_2 les gains respectifs du préamplificateur et du dispositif de mise en forme, f_c la bande passante de mesure, b_{MF}^2 la densité spectrale de bruit propre de la mise en forme à son entrée et B_{MF} la bande passante de la mise en forme. La figure 4 montre schématiquement le signal d'attaque du compteur, de pente $p_S = G_1 G_2 p_E$ où p_E désigne la pente du signal de battement.

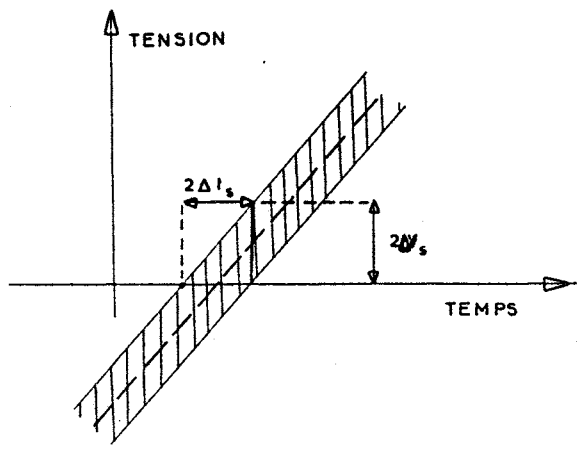


Fig. 4. Signal d'attaque du compteur au voisinage de "zéro".

La zone hachurée correspond au bruit quadratique moyen. La pente est reliée aux fluctuations (RMS) de tension de sortie et d'instant de passage à zéro par :

$$p_s = \frac{\Delta V_s}{\Delta t_s} \qquad (5)$$

ce qui permet d'expliciter Δt_s, lequel contient les fluctuations des instants de passage à zéro, donc la limitation du système de mesure

$$\Delta t_s = \frac{V_s}{p_s} = \frac{\sqrt{(G_1 G_2)^2 b_M^2 f_c + G_2^2 b_{MF}^2 B_{MF} + G_2^2 b_F^2 f_c}}{G_1 G_2 p_E} \qquad (6)$$

soit encore :

$$\Delta t_s = \frac{\sqrt{b_M^2 f_c}}{p_E} \sqrt{1 + \left(\frac{1}{G_1}\right)^2 \frac{b_{MF}^2 B_{MF} + b_F^2 f_c}{b_M^2 f_c}} \qquad (7)$$

Le bruit "normal" à prendre en compte correspond à $\sqrt{b_M^2 f_c}$ associé à la pente p_E. On constate que s'il n'y a pas de préamplificateur avant le filtre ($G_1 = 1$), la limitation de performance passe, relativement, de 1 à $\sqrt{1 + \frac{b_{MF}^2 B_{MF} + b_F^2 f_c}{b_M^2 f_c}}$

Dans l'hypothèse où f_c = 10 Hz, b_f = 13nV/$\sqrt{\text{Hz}}$, b_M ≠ 2 nV/$\sqrt{\text{Hz}}$, b_{MF} = 3,5 nV/$\sqrt{\text{Hz}}$ et B_{MF} = 2 kHz, on constate une dégradation théorique de 29 dB (!) par rapport à l'effet du bruit blanc "normal" b_M. La limite théorique de la stabilité devient alors $\sigma_y(\tau) = 3{,}8\ 10^{-14}\ \tau^{-1}$ au lieu de $1{,}4.\ 10^{-15}\ \tau^{-1}$.

Si l'on utilise une préamplification de gain G_1 = 30 avant le filtre, le terme correctif chute de $\sqrt{750}$ à $\sqrt{2}$, soit 3 dB, ce qui est tout à fait acceptable. Au niveau expérimental, l'amélioration est très nette, mais moins spectaculaire (cf. II.2) en raison d'autres limitations (bruit flicker, résiduel 50 Hz...).

La configuration "idéale" d'un filtrage judicieusement associé à la mise en forme permet de conserver seulement la bande f_c pour toutes les contributions de bruit ($B_{MF} = f_c$), ramenant alors le terme correctif très sensiblement à 1. Cette amélioration passera totalement inaperçue car, pour une faible bande f_c, le bruit blanc n'est pas le processus limitatif. Nous ne l'avons donc pas expérimentée.

II. RESULTATS EXPERIMENTAUX A 5 MHz

Les limites théoriques mentionnées précédemment laissent entrevoir des planchers de mesure de l'ordre de quelques $10^{-15}\ \tau^{-1}$ en court terme à 5 MHz, avec une bande passante de l'ordre de 10 Hz. Ceci est sensiblement meilleur que les planchers de stabilité mentionnés dans la littérature, à 5 MHz ou même plus haut en fréquence (1)(5)(6). Nous avons tenté de vérifier dans quelle mesure l'expérience procurait des planchers de mesure du même ordre, par utilisation d'un système à 2 mélangeurs à 5 MHz. Afin de comprendre les mécanismes limitatifs et de vérifier les idées développées dans la 1ère partie de cette communication, nous nous sommes placés dans des circonstances défavorables (pas de préamplification avant filtrage) puis favorables (préamplification avant filtrage). Les résultats sont les suivants :

II.1. Banc de mesure sans préamplificateur avant le filtre (f_c)

Le dispositif de test du plancher de mesure est représenté sur la figure 5. Pour ce test, les deux horloges sont remplacées par un même oscillateur Q_1, et l'oscillateur Q_2 (généralement appelé oscillateur commun ou décalé, dans la littérature) est décalé de 1 Hz pour permettre l'obtention des 2 battements. Q_1 et Q_2 sont des oscillateurs OSA modèle 8600 (variété 2), dont la stabilité est de l'ordre de 6. 10^{-13} sur 1 s et 10 s après deux semaines de fonctionnement continu.

1ère configuration de mesure : mélangeurs SRA 1, avec + 7 dBm en O.L. et + 1 dBm en RF (→ K = 0,3 V/rad), pas de préamplification avant filtrage, amplificateurs opérationnels OP 27 pour la mise en forme (bruit blanc: 3,5 nV/$\sqrt{\text{Hz}}$). L'écart δt entre les fronts montants des 2 signaux de battement est ajusté à 1 ms environ, et cet ordre de grandeur de δt sera conservé dans toutes les mesures qui suivent. Pour τ = 1 s, on obtient le plancher de stabilité $\sigma_y(\tau = 1\text{ s})$ ≠ $5{,}8\ 10^{-14}$ (100 points) pour f_c = 50 Hz.

2e configuration : idem, mais les amplificateurs opérationnels sont des LT 1028 (bruit blanc à 0,9 nV/$\sqrt{\text{Hz}}$). Pour τ = 1 s, on obtient $\sigma_y(\tau = 1\text{ s})$ ≠ $6{,}2.\ 10^{-14}$ (100 points) pour f_c = 10 Hz et pour f_c = 50 Hz.

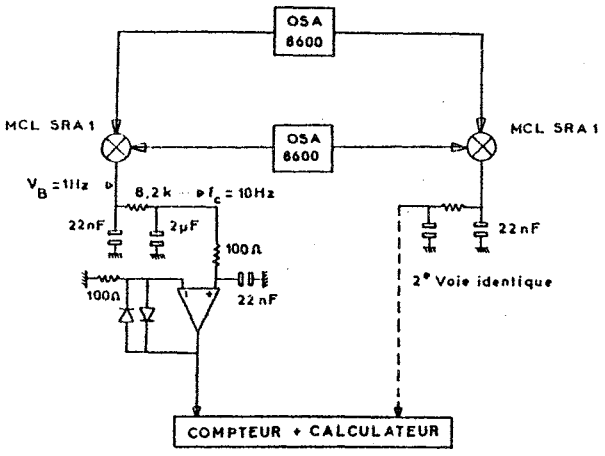


Fig. 5. Test du plancher de mesure sans préamplificateur avant le filtre f_c.

Conclusion surprenante : avec un amplificateur qui est notablement moins bruyant, on trouve un résultat très légèrement moins bon, $\forall$ f_c ! En fait, ceci est parfaitement compréhensible si l'on considère la bande passante des 2 amplificateurs opérationnels et la puissance totale de bruit qui en résulte dans les deux configurations (cf. I.2.2.) avec le gain de la mise en forme (2500 environ), la bande passante est 2 kHz avec l'OP 27 et 30 kHz avec le LT 1028, d'où une puissance (totale) de bruit voisine pour les amplificateurs OP27 et LT 1028. Ceci est parfaitement compatible avec le résultat de la mesure.

Du point de vue ordre de grandeur, la limite théorique calculée en tenant compte du bruit large bande (2 kHz) de l'OP 27, pour f_c = 50 Hz ou pour f_c = 10 Hz, s'établit à :

$$\sigma_y(\tau) \approx \sqrt{\frac{3h_2 f_c}{8\pi\tau^2}} \times \sqrt{1 + \frac{b_{MF}^2 B_{MF} + b_f^2 f_c}{b_M^2 f_c}} \approx 3{,}8.10^{-14}\,\tau^{-1}$$

Cette valeur est tout à fait compatible avec le résultat de mesure obtenu pour τ = 1 s, soit 5,8 10^{-14}.

II.2. Banc de mesure avec préamplificateur avant le filtre f_c

Dans ce dispositif représenté sur la figure 6, les premières mesures ont été effectuées avec des mélangeurs MCL SRA 1.

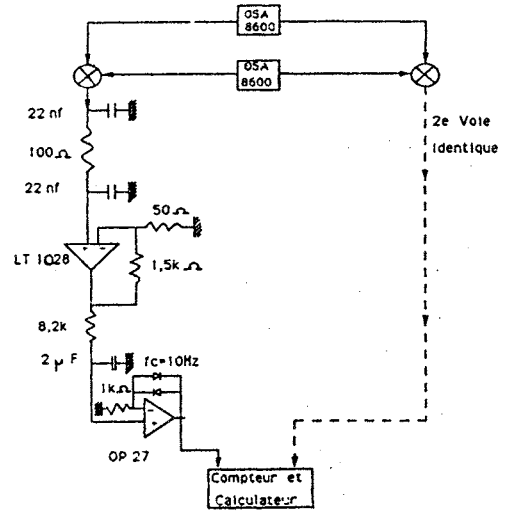


Fig. 6. Test du plancher de mesure avec préamplificateur avant le filtre f_c.

Les niveaux d'attaque étaient + 7 dBm sur les entrées O.L., + 1 dBm sur les entrées RF, procurant un gain K = 0,3 V.rad^{-1} aux mélangeurs. L'écart δt était $\lesssim$1ms (T_B = 1 s).

Dans ces conditions, on constate une amélioration du plancher de mesure, mais le montage réalisé est sensible aux perturbations extérieures : voisinage d'un fer à souder thermostaté Weller en fonctionnement, acoustique, vibrations mécaniques... Les liaisons HF doivent impérativement être réalisées en câble semi rigide et connecteurs vissés de type SMA (RIM) par exemple. Malgré cette dernière précaution, le dispositif reste relativement microphonique et montre une dispersion de mesure sensiblement plus forte que dans les configurations précédentes (II.1.). Les préamplificateurs sont des LT 1028.

1ère configuration : f_c = 16 Hz, $\delta t \sim$ 1 ms
Sur dix enregistrements de 50 points, le meilleur donne un plancher de stabilité à 1,62. 10^{-14} (τ = 1 s), tandis que le plus mauvais se situe à 2,35 10^{-14} (τ = 1 s) : cette dispersion, à la limite des barres d'erreur usuelles, traduit en fait la sensibilité du dispositif aux perturbations extérieures.

2e configuration : f_c = 72 Hz
Sur dix enregistrements de 50 points (τ = 1 s), le meilleur se situe à 2,03. 10^{-14} et le plus mauvais à 3,36. 10^{-14}, ce qui indique toujours une sensibilité à l'environnement.

3e configuration : f_c = 90 Hz, montage stabilisé depuis 30 mn et ambiance très calme.
Sur 200 points, on obtient un plancher $\sigma_y(\tau = 1s)$ = 1.81. 10^{-14}. Ceci confirme la nécessité de travailler en environnement physique calme.

Dans cette dernière configuration, avec 90 Hz de bande, le plancher théorique de stabilité (incluant les résistances de 100 Ω sur les entrées + et -) se situe à 6,5. $10^{-15}\tau^{-1}$ pour le bruit blanc.

4e configuration : idem, sauf niveaux d'attaque des mélangeurs SRA 1 (+ 7 dBm sur toutes entrées) → pas d'amélioration.

Les valeurs mesurées dans ces 2 dernières configurations restent plus élevées que la valeur théorique due au bruit blanc, ce qui indique l'existence probable d'un bruit "flicker de phase" (relativement) fort des mélangeurs. Nous avons donc essayé ensuite les mélangeurs HP 10514, connus pour leur haute performance en bruit flicker [4].

5e configuration : mélangeurs HP 10514 A, avec + 7 dBm sur chaque entrée. Ces mélangeurs sont malheureusement équipés de connecteurs BNC, ce qui a imposé l'utilisation de transitions BNC-SMA pour permettre un câblage semi-rigide. La configuration optimale consistera à utiliser des mélangeurs HP 10514 B dans la version définitive du banc de mesure.

Les mesures effectuées donnent un "plancher" satisfaisant à 5 MHz, mais les résultats sont assez dispersés, pour diverses raisons :
- tout d'abord la présence des connecteurs BNC et des transitions BNC-SMA donnent une microphonie très nette à ce montage
- la présence éventuelle de composantes (parasites) à 50 Hz (et multiples) est un élément très limitatif vis à vis du plancher de mesure : un signal (50 Hz) à - 120 dBc sur le battement limite la stabilité mesurable à environ 10^{-12} sur τ = 1 s, donc l'obtention de 10^{-14} sur 1 s nécessite un niveau de signal parasite à 50 Hz plus faible que - 160 dBc sur les signaux de battement.

Avec le banc de mesure représenté sur la figure 6, utilisant des mélangeurs HP 10514 A attaqués par des signaux 5 MHz (+ 7 dBm en O.L. et en RF), les résultats obtenus dans les conditions les plus favorables, sans thermostat du banc de mesure sont représentés dans le tableau ci-après, pour une bande passante de mesure f_c = 10 Hz et un échantillon de 100 points.

τ	1 s	10 s	100 s	1 000 s
σ_y (plancher)	$7{,}7.10^{-15}$	$1{,}0.10^{-15}$	$1{,}6\ 10^{-16}$	$6{,}5.10^{-17}$

Sur la figure 7 qui représente ces résultats, on constate qu'à 1 000 s, la pente du plancher de stabilité est notablement perturbée, cet effet étant fortement corrélé aux fluctuations thermiques ambiantes.

L'utilisation d'une bande passante plus élevée (80 Hz) procure, sur τ = 1 000 s, un plancher de stabilité légèrement meilleur ($4{,}78.10^{-17}$), ce qui confirme un effet thermique différentiel entre les filtres (f_c) des 2 voies, et la nécessité de thermostater le banc. La stabilité à moyen terme pourrait alors correspondre à la courbe en pointillés qui prolonge la partie rectiligne (5 MHz).

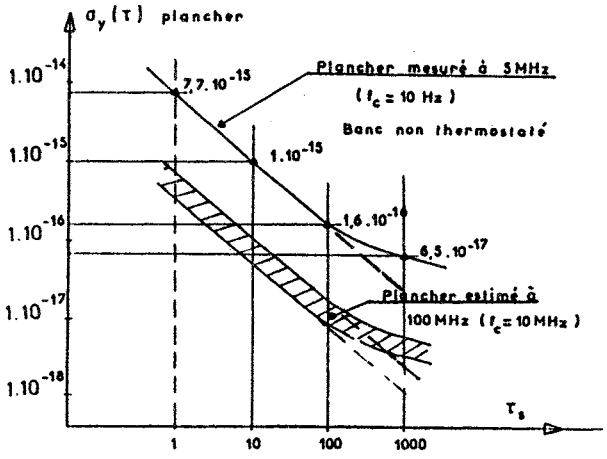


Fig. 7. Plancher de stabilité mesuré à 5 MHz et estimé à 100 MHz.

D'autres perturbations peuvent également intervenir dans cette dégradation : microphonie du montage (→sauts de phase de l'ordre de 10^{-6} période, ce qui correspond à 13 fois la stabilité sur 1 s), bruit d'amplitude en très basse fréquence (variation des offsets des mélangeurs) pour les oscillateurs 5 MHz, interférences électromagnétiques diverses...

Globalement, on constate donc que, dans le domaine du court terme, le plancher de stabilité est actuellement limité aux alentours de 7.10^{-15} à τ=1 s, (f_c = 10 Hz), en raison principalement du bruit flicker des mélangeurs et de perturbations ambiantes "rapides" (vibrations , parasites 50 Hz ou R.F....). En moyen et long terme, l'amélioration des performances nécessite impérativement un thermostat.

III. POSSIBILITE D'AMELIORATION DES PERFORMANCES PAR UTILISATION DE FREQUENCES PLUS ELEVEES

La formule qui permet de calculer le plancher de mesure en variance d'Allan s'écrit pour le système à 2 mélangeurs :

$$\sigma_y(\tau) = \frac{1}{2\,\tau\,T_B\,\nu_o}\ \sqrt{<\Delta_2^2\,(\delta t)>} \qquad (8)$$

où T_B est la période de battement, τ (=nT_B) le temps de mesure, ν_o la fréquence des oscillateurs sous test et $\sqrt{<\Delta_2^2(\delta t)>}$ la valeur R.M.S. des différences secondes sur δt. Cette formule peut prendre la forme (9) :

$$\sigma_y(\tau) = \frac{1}{2n}\ \frac{\nu_B}{\nu_o}\ \frac{\sqrt{<\Delta_2^2(\delta t)>}}{T_B} \qquad (9)$$

où ν_B est la fréquence du battement.

On montre aisément que l'utilisation d'une fréquence de battement ν_B plus faible ne modifie absolument pas la contribution du banc de mesure au plancher de stabilité, contrairement à ce qu'un raisonnement hâtif, mais erroné, pourrait laisser croire.

Par contre, l'utilisation d'une fréquence de signal ν_o plus élevée va réduire d'autant la contribution du banc au plancher de stabilité. En effet dans ces conditions, la perturbation de phase équivalente au bruit du banc ramené à la fréquence ν_o (via le gain K du mélangeur) donnera une densité spectrale relative de phase inversement proportionnelle à ν_o^2, ce qui correspond à une amélioration du "plancher de stabilité" proportionnelle à $1/\nu_o$. Ceci est vrai à condition que les paramètres caractéristiques du banc soient maintenus constants lorsqu'on varie ν_o. C'est le cas pour les amplificateurs B.F., mais pas pour les mélangeurs, dont la pente K (en $V.rad^{-1}$)ne se maintient pas lorsque ν_o passe par exemple de 5 MHz à 100 MHz : dans les conditions de charge représentées sur les figures 5 et 6, les mélangeurs (MCL comme HP) procurent un battement dont la forme d'onde (à 1 Hz) est trapézoïdale pour ν_o = 5 MHz et triangulaire pour ν_o = 100 MHz. Dans ce cas, le gain K est approximativement réduit dans le rapport 1,5. Dans les mélangeurs, il n'est pas certain non plus que la composante "flicker" reste constante au sens de $S_\varphi(f)$, ce qui peut également limiter l'amélioration attendue. Enfin, les caractéristiques des amplificateurs de distribution (éventuellement) intercalés devront rester constantes au sens de $S_\varphi(f)$, ce qui n'est pas totalement certain.

Par ailleurs, l'utilisation opérationnelle d'un tel banc de mesure nécessite de réduire au minimum l'effet de nombreuses perturbations, ce qui impose :

- la réduction de la microphonie par des câblages en semi rigide et des connecteurs de type SMA, associés à une électronique très compacte
- la conception d'un ensemble électronique conforme aux règles de la compatibilité électromagnétique (E.M.C.) : boitiers étanches aux signaux R.F., filtres d'alimentation bien adaptés, suppression des signaux "hors bande" sur les éléments actifs, équipotentialité optimum de l'ensemble du montage...
- l'utilisation éventuelle d'autres mélangeurs mieux adaptés à des fréquences élevées et la détermination des conditions optimales de charge des mélangeurs
- l'utilisation d'un termostat de gain convenable (10 à 100 ?) pour stabiliser en température l'oscillation décalé et l'électronique de traitement des signaux de battement
- l'asservissement de l'oscillateur décalé sur l'une des horloges testées, afin de maintenir constante la période de battement et de réduire l'effet du bruit de l'oscillateur décalé
- l'utilisation éventuelle d'une alimentation par (batteries + régulateurs) pour réduire le "50 Hz" résiduel.

Dans ces conditions les performances prévisibles pour le plancher de mesure entre signaux 100 MHz se situent dans la bande hachurée représentée sur la figure 7, c'est à dire sensiblement 4 à 8. 10^{-16} sur τ = 1 s, 5. 10^{-17} à 1. 10^{-16} sur τ = 10 s, 8. 10^{-18} à 1.6 10^{-17} sur τ = 100 s, quelques 10^{-18} sur τ = 1000 s et probablement mieux au-delà de τ = 1000 s.

CONCLUSION

Le système de mesure à deux mélangeurs permettra sans aucun doute, par ses performances élevées, la mesure de stabilités de fréquence de l'ordre 5. 10^{-15} sur 1 s, 10^{-16} sur 10 s, quelques 10^{-18} sur 1000 s. La limitation éventuelle risque d'être liée aux amplificateurs de distribution, qui devront conserver, à 100 MHz, la densité spectrale S_φ qu'ils sont capables de procurer à 5 MHz, en particulier pour le bruit "flicker de phase". Rien ne semble s'y opposer théoriquement. Les futurs étalons de fréquence pourront donc être évalués à 100 MHz par les bancs de mesure à 2 mélangeurs fonctionnant à température ambiante.

REFERENCES

(1) D.W. Allan and H. Daams, "Picosecond time difference measurement system", in Proc. 29th Annual Frequency Control Symp., May 1975 (pp. 404-411)

(2) C. Audoin, "Frequency Metrology", Metrology and Fundamental Constants (North Holland Publ. Co.) (pp. 171,172,185) 1980

(3) Low Sze-Ming, "Influence of noise of common oscillator in dual-mixer time-difference measurement system", in IEEE Trans. On Instr. and Meas., Vol. IM-35, n° 4, Dec. 1986, pp. 648-651

(4) A. Filimon. Thèse de Docteur-Ingénieur. Etude d'une chaine de multiplication de fréquence à faible bruit. Université Pierre et Marie Curie, Paris VI - 29 Juin 1983 (p. 3.15 à 3.22)

(5) R.A. Eichinger, "Low noise buffer amplifiers and phase comparators for precise time and frequency distribution". Proc. 35th Ann. Frequ. Control Symposium, Fort Monmouth, N.J. (USA), May 1981, pp. 565-575

(6) T.K. Tucker, "Operating and environmental characteristics of Sigma Tau hydrogen masers used in the V.L.B.A.", Proc. PTTI 1988, 29 Nov.-1er Dec. 1988, Vienna, Virginia (USA)

THE REJECTION PROBLEM IN PRECISE PHASE AND FREQUENCY NOISE MEASUREMENT

M. Olivier, A. El Ghazi

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

The recent and rapid improvement of oscillator frequency noise or of passive system phase noise has led to more stringent requirements for phase and frequency measurement systems. The minimum detectable phase or frequency noise performance of a measurement system should be measured with precision. A reference signal must be used to drive the system under test, and the noise of this signal should be eliminated : this can be done using a perfectly balanced measurement system. In fact, we observe a rejection of this noise which is often no more than 20 dB. This can set a limit to the measured performance of the measurement system. We have analyzed the rejection factor for different phase and frequency measurement systems, both passive and active, in the RF and microwave domains. We show the influence of the unbalance and of the phase-frequency relationships for different subsystems, and especially for a phase shifter, which is a fundamental element in an active phase noise measurement system.

Introduction

The noise characteristics of an oscillator can be described, in the spectral domain, in terms of the phase or frequency noise spectra. The phase fluctuations of the oscillator must be detected and amplified before analysis. These operations are usually made using either a passive or active measurement system.

The minimum detectable phase noise performance of these measurement systems must be measured with precision. At the present time, a system noise floor as low as -160 dB can be obtained. Performance is measured with a reference oscillator which drives the system under test. In most of the papers published, we find this statement : "taking into account the symmetry of the system, the noise of the reference oscillator is eliminated" [1].

As a consequence, one could suppose that any kind of oscillator can be used as a reference oscillator : this is not quite true. The rejection of the noise of the oscillator is, in fact, limited.

Some comments concerning this rejection problem are given.

Measurement technique

A phase or frequency noise measurement system uses a reference to detect the phase or frequency fluctuations of the oscillator under test.

We will consider first a passive measurement system, that is, a system which uses a passive resonator as a reference, and then an active system with an oscillator as a reference.

Passive system

The resonator is a discriminator, having a phase-frequency relationship represented by K_d. We have then a classical phase detector of sensibility μ, and an amplification of the detected signal with gain g (Fig. 1).

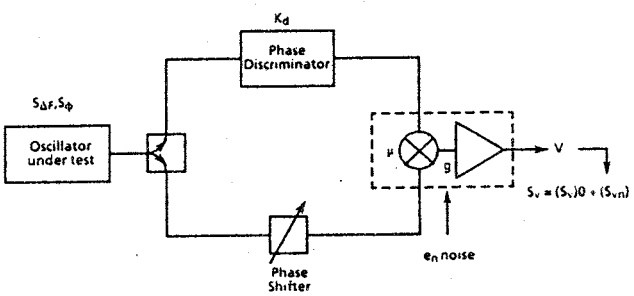


fig. 1a

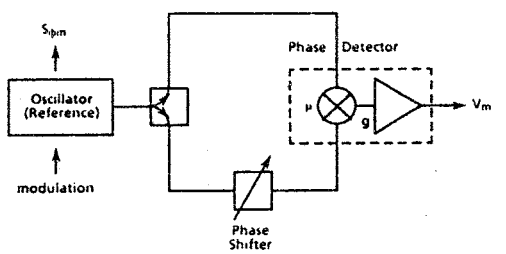


fig. 1b

Fig. 1 : 1a - Passive measurement system
1b - Measurement of the rejection factor

The spectral density S_V of the detected signal $V(t)$ contains a component due to the noise in the system, that is S_{V_n}.

The phase fluctuations of the oscillator under test is related to S_V by the relation :

$$S_\phi(F) = \frac{S_V(F)}{(g\,\mu\,K_d\,F)^2}$$

The noise contribution S_{V_n} to S_V gives the minimum detectable phase noise $(S_\phi)_{min}$ of the measurement system

$$(S_\phi)_{min} = \frac{S_{V_n}}{(g\,\mu\,K_d\,F)^2}$$

The measurement of S_{V_n} uses a symmetrical system. It is important to note that a phase shifter must be used in this system to adjust in phase quadrature the two signals arriving at the phase detector.

If the hypothesis of complete rejection of the noise of the oscillator is not respected, a part of the noise voltage spectrum, $(S_V)_a$ is due to the frequency fluctuations of the reference oscillator, and this is interpreted as a minimum detectable phase noise $(S_{\phi_{min}})_a$

$$\left(S_{\phi_{min}}\right)_a = \frac{\left(S_{V_n}\right)_a}{(g\,\mu\,K_d\,F)^2}$$

This quantity is easily measured with a modulation applied to the reference oscillator (Fig. 1b).The modulation signal can be either a sinusoïdal signal or white noise, its amplitude being high enough that the output noise voltage of the system can be neglected.

We can thus define and measure a rejection factor of the oscillator

$$R_p = \frac{(S_\phi)_{min}}{S_\phi} = \frac{\left(S_{V_n}\right)_a}{(g\,\mu\,K_d\,F)^2\,S_\phi}$$

Active system

A rejection factor can also be defined for an active measurement system. One uses two oscillators maintained in phase quadrature by a phase lock loop, and a phase detector gives, after amplification, a detected signal S_V which is linearly proportional to the phase fluctuation of the oscillator under test (Fig. 2a)

$$S_\phi = \frac{S_V}{g^2\mu^2}$$

This relation is correct for frequencies outside the bandwidth of the phase lock loop. Inside the bandwidth, the system is sensitive to the frequency fluctuations of the oscillator, but a correction can be applied to the signal with a pseudo integrator. We then obtain the same relation as before.

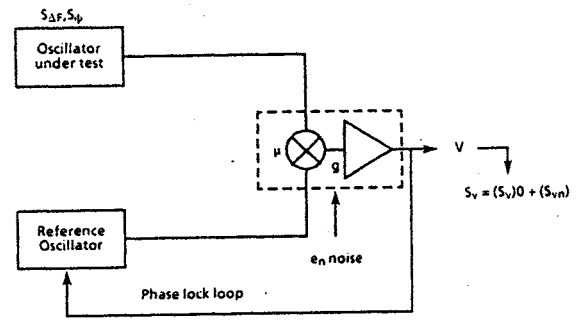


fig. 2a

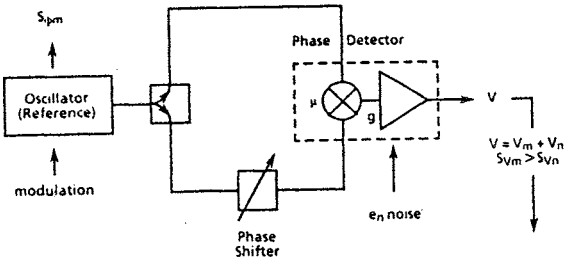


fig. 2b

Fig. 2
2a : Active measurement system
2b : Measurement of the rejection factor

The minimum detectable phase noise spectrum $S_{\phi_{min}}$ is related to the output noise voltage spectrum ; and the contribution of the noise of the oscillator, when not completely rejected, is again described by a rejection factor R_a which is now given by the formula (Fig. 2b)

$$R_a = \frac{S_{V_n}}{g^2\mu^2 S_{\phi_n}}$$

Rejection factor measurement

The origin of the finite rejection factor may be the phase shifter which must be introduced in one of the two arms of the measurement system to obtain the phase-quadrature. A phase shift is always more or less frequency sensitive, and we introduce a coefficient a to caracterize the phase-frequency relationship of the phase shifter.

We can easily relate R_a and R_p to a

$$R_a = a^2F^2 \quad and \quad R_p = \frac{a^2}{K_d^2}$$

The a factor is easily determined for different phase shifters. If we measure the phase variation around 90° due to frequency variations, we obtain an a of the order of 10^{-8} to 10^{-9} rad/Hz at a frequency of 70 MHz and far less (10^{-10}) in the microwave domain (Fig. 3).

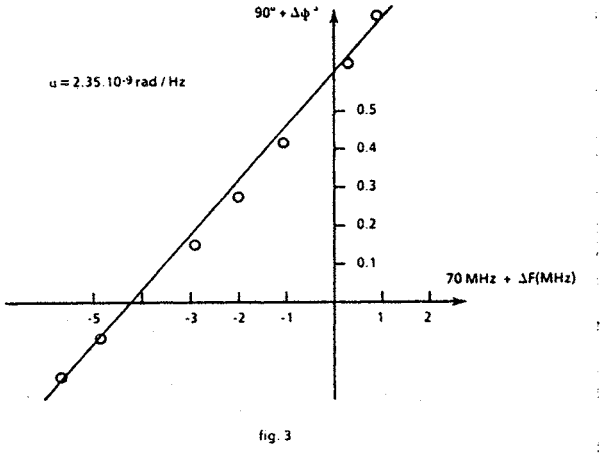


fig. 3

Fig. 3 : a factor measurement

As a consequence, the rejection factor of the reference oscillator for a passive measurement system is of the order of 30 to 40 dB. It can reach 180 dB for an active system at 1 Hz from the carrier, decreasing to 60 dB at a Fourier frequency of 1 MHz.

In the microwave domain, an active phase noise measurement system may use a preliminary mixing of the oscillator under test with a reference oscillator, and the phase detection is performed in the R.F. domain (Fig. 4). In this case, the minimum detectable phase noise is measured with a symmetrical system : two identical microwave mixers and preamplifiers are used. The transfer functions of the preamplifiers are, however, usually different. In this case the rejection of the oscillator can be enhanced by adjusting the I.F. frequency. A rejection of 170 dB has been obtained at 1 Hz of the carrier. However, the minimum detectable phase noise of the measurement system must be measured with a Gunn oscillator whose noise spectrum has been reduced through a phase locking to a high Q cavity (Fig. 5).

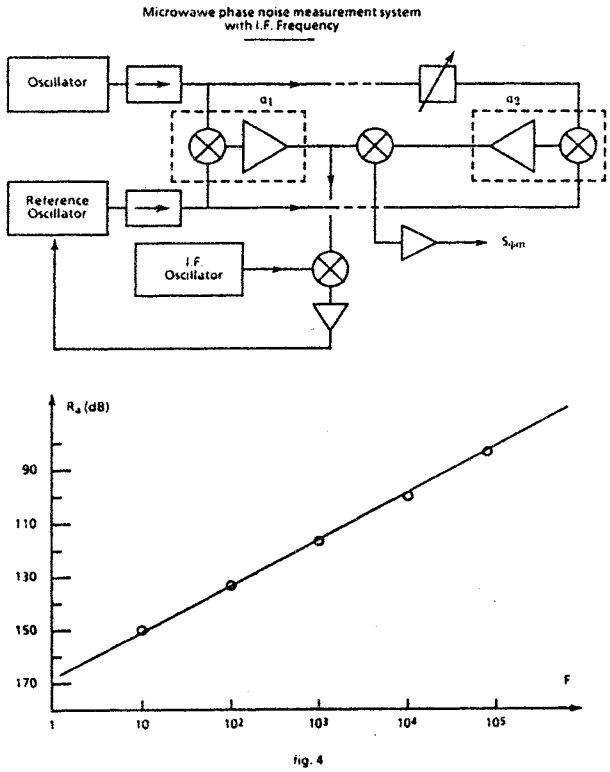
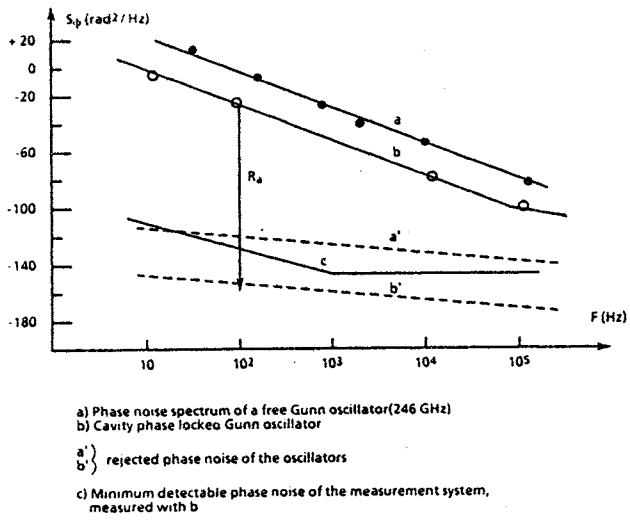


fig. 4

Fig. 4 : Microwave phase noise measurement with IF frequency



a) Phase noise spectrum of a free Gunn oscillator(246 GHz)
b) Cavity phase locked Gunn oscillator

a' } rejected phase noise of the oscillators
b' }

c) Minimum detectable phase noise of the measurement system, measured with b

fig. 5

Fig. 5 : Microwave measurement. Rejection of the oscillator

Rejection problem in resonator noise measurement

There is also a problem of oscillator noise rejection when one measures the frequency fluctuations of a passive device such as a resonator. A symmetrical system is required, and the symmetry of the resonators, with regard to their Q-factors and the resonance frequencies, has to be carefully adjusted (Fig. 6a).

If K_d is the common phase-frequency factor of the two resonators, we obtain the frequency spectrum of the resonator

$$S_{\Delta f_r} = \frac{S_{\Delta V}}{g^2 \mu^2 K_d^2}$$

where g, μ caracterize the phase detection, and $S_{\Delta V}$ is the output voltage spectrum.

One part of the output spectrum, that is $(S_{\Delta V})_o$ is due to the frequency noise $(S_{\Delta f})_o$ of the reference oscillator. This can be attributed to a difference ΔK_d between the characteristics K_d of the two resonators

$$(S_{\Delta V})_o = g^2 \mu^2 \Delta K_d^2 \, (S_{\Delta F})_o$$

Here, $(S_{\Delta V})_o$ is interpreted as a frequency noise of the resonator

$$(S_{\Delta f_r})_{min} = \frac{(S_{\Delta V})_o}{g^2 \mu^2 K_d}$$

The rejection factor of the oscillator R_r can be again defined as the ratio of $(S_{\Delta f_r})_{min} / (S_{\Delta F})_o$

$$R_r = \frac{(\Delta K_d)^2}{K_d}$$

A rejection of 33 dB has been obtained for a measurement of the noise of dielectric resonators [2].

The spectral purity of a 9 GHz dielectric cavity oscillator has been tested with an active measurement system and this oscillator has been phase-locked to an harmonic of a 1 GHz quartz oscillator [3], providing a low noise microwave reference signal for measurement of the noise of dielectric resonators. It can be shown (Fig. 6b) that the rejection factor R_r is too low to permit the use of an unstabilized dielectric cavity oscillator as a reference for the measurement of the resonator noise.

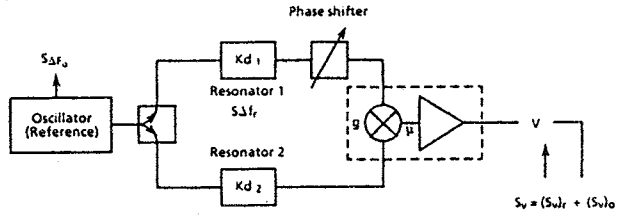
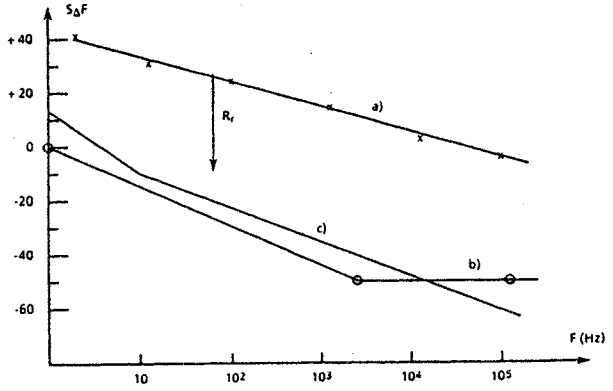


fig. 6a



a) Dielectric cavity oscillator
b) Dielectric cavity oscillator phase locked on a quartz oscillator
c) Dielectric resonator frequency noise

fig. 6b

Fig. 6
6a : Resonator noise measurement system
6b: Noise measurement of resonators and rejection problem

Conclusion

The measurement of the minimum detectable phase or frequency noise performances of a phase noise measurement system has been analyzed. The frequency noise of the reference oscillator used, must be rejected. A rejection factor has been defined and measured. This rejection factor can be as high as 150 dB, especially when one measures the minimum detectable phase noise of an active measurement system in the vicinity of the carrier.

Nevertheless, some care must be taken in the choice of the reference oscillator when one works in the R.F. domain, and when a passive reference system is used.

The rejection of amplitude modulation is a problem which has not been considered here. But, generally, the amplitude noise of a high quality oscillator is negligible compared to the frequency noise.

References

[1] F.L. Walls and A.E. Wainwright, "Measurement of the short term stability of quartz crystal resonators and the implications for crystal oscillator design and applications", IEEE Trans. on Inst. and Meas., vol. IM-24, n° 1, March 1975.
D. Scherer, "The art of phase noise measurement", RF and Microwave Symposium, may 1983.

[2] J.C. Chauvin and M. Olivier, "Oscillateur à quartz à 1 GHz utilisé en synthèse microonde", 1er Forum Européen Temps-Fréquence, Besançon, 1987.

[3] M. Olivier, J. Groslambert, M. Valentin, "Mesure du bruit de phase, par des méthodes actives et passives de sources microondes de haute pureté spectrale", 5e Journées Nationales Microonde, Nice, 22-24 juin 1987.

Coherent frequency synthesis up to 3.7 THz

A. Godone*, M.P. Sassi**, C. Caldera***, E. Bava*

* Istituto Elettrotecnico Nazionale Galileo Ferraris, Torino, Italy
** Istituto di Metrologia G. Colonnetti (CNR), Torino, Italy
*** Politecnico di Torino, Italy

Abstract

We report the absolute frequency measurement of the 81.5 μm rotational transition and of the 263.4 μm inversion transition in the ν_2 vibrational state of $^{14}NH_3$.

The 81.5 μm emission line of the $^{14}NH_3$ laser is useful in frequency measurements up to the infrared because its 8th harmonic is in good coincidence with the CO_2 laser line 10R(30) which is useful for the connection with P(7) CH_4 transition: a simple frequency synthesis chain results in this case, being possible to reach directly the above NH_3 line with a Schottky diode harmonic mixer driven by a 70 GHz klystron.

The klystron is phase-locked, through a very low phase-noise frequency multiplication chain, to a high spectral purity quartz oscillator having a one-sided phase spectral density $S_\phi = 2\times10^{-18}$ rad^2/Hz in its white part.

The beat note between the laser line and the corresponding harmonic of the klystron is observed on the spectrum analyzer: the spectrum turns out to be suitable for frequency measurements with resolution higher than the reproducibility of the laser itself.

The values of the signal-to-thermal noise and of the signal-to-phase noise ratios are in good agreement with other measurements and are the first data concerning a phase-coherent frequency synthesis from a 5 MHz quartz oscillator to the terahertz region.

Both signal-to-noise ratios may be further increased in order to allow precise frequency counting as required in the multiplication chains connecting microwaves with the infrared.

1.- INTRODUCTION

Work in the field of phase-coherent frequency synthesis is in progress, with the aim of realizing a phase-coherent link between the available frequency standards and the optical lasers.

Direct frequency measurements up to the infrared appear to be possible since synthesis chains have been realized, which allow to maintain the phase coherence of the signal up to 30 THz [1].

A new step has recently been done in this direction, with the absolute frequency measurement of the 81.5 μm emission line of the ammonia FIR laser (at 3.7 THz) [2] which yields the first data of phase-coherent frequency synthesis from 5 MHz to the terahertz region.

The above line is directly reached by the 51st harmonic of a 70 GHz klystron, produced by a Schottky diode harmonic mixer.

Furthermore, the ammonia FIR laser tuned on this line can be used as an intermediate transfer oscillator in a rather simple chain reaching the P(7) CH_4 line at 88 THz.

2.- MULTIPLICATION CHAIN

In order to discuss the RF spectrum of the synthesized signal at 3.7 THz it is useful to remind some basic concepts and formulas concerning the phase noise multiplication and the design criteria for the realization of a real FIR synthesizer.

The signal-to-phase noise ratio of a multiplier output is known to decrease according to the formula [1]:

$$S/N_\phi = 1/n^2 \cdot 2/(B_w S_\phi^w(\nu_0)) \cdot (\phi/1-e^{-\phi})^{\alpha/\alpha-1} \cdot e^{-\phi} \qquad (1)$$

where:

- n = multiplication order;
- B_w = filter bandwidth of the spectrum analyzer;
- $S_\phi^w(\nu_0)$ = white phase part of the phase noise spectral density of the quartz oscillator;
- α = high frequency roll-off exponent of $S_\phi(f,\nu)$

and:

$$\phi = \int_0^\infty S_\phi(\nu_0,f)\, df \qquad (2)$$

corresponds to the total phase noise power in the pedestal.

The collapse frequency for which the normalized carrier power is reduced to e^{-1} is given by:

$$\nu^{coll} = \left[\frac{\pi/\alpha}{\sin \pi/\alpha} \cdot B \cdot S_\phi^w(\nu_0)\right]^{-1/2} \cdot \nu_0 \qquad (3)$$

where B is the high frequency cut-off of $S_\phi(f,\nu)$, that is the noise bandwidth of the chain.

B_w has to be much smaller than B.

For $\nu > \nu^{coll}$ the carrier collapses and the noise pedestal width increases as:

$$\Delta\nu_p = 2B\left[\frac{\phi}{1-e^{-\phi}}\right]^{1/(\alpha-1)} \qquad (4)$$

From these formulas strict conditions follow on the starting oscillator and on the noise bandwidth B in order to have a high value of ν^{coll} and a good spectral purity multiplier output signal.

For a quartz crystal oscillator at 5 MHz with $S_\phi^w = 2\times10^{-18}$ rad^2/Hz and a chain noise bandwidth of 20 kHz it is in principle possible to reach a collapse frequency $\nu^{coll} = 33$ THz [1].

The plot of $(S/N)_\phi$ versus synthesized frequency in

this case is shown in Fig. 1: the experimental points up to 10 GHz correspond to previous measurements in different points of the multiplication chain [1].

The synthesizer section from X-band to the FIR is shown in Fig. 2: a harmonic mixer M_1 produces the beat note between the signal of an E-band klystron and the n-th harmonic of the X-band signal. This beat note is then used to phase lock the klystron which in turn is also the local oscillator for the harmonic mixer M_2 to obtain a beat with the FIR signal from the laser.

3.- MEASUREMENT OF AMMONIA 81.5 μm AND 263.4 μm LINES

The 81.5 μm (rotational transition) and the 263.4 μm inversion transition of NH_3 [3] (Fig. 3), have been measured by means of heterodyne technique.

The radiation is emitted by a NH_3 optically pumped FIR laser with dielectric waveguide resonator [2] giving 20 mW at 81.5 μm and 1 mW at 263.4 μm on 3 W from the pump laser.

A μ-size Schottky barrier point contact diode is used as detector and mixer, having 2.2 fF capacitance, ~ 15 Ω resistance and 1.5×10^{17} cm^{-3} epilayer doping.

The junction is coupled to the laser field by a tungsten whisker placed in the focal spot of an f/5 TPX lens at the optimum antenna angle.

Similarly, the radiation of a 70 GHz klystron, representing the local oscillator, is directly irradiated from the open end of a waveguide close to the antenna.

The klystron is phase-locked to a 72 GHz signal synthesized from the chain described above.

The beat signal generated by the Schottky diode is amplified and observed at the spectrum analyzer (Fig. 4): the harmonic mixing number is 51 for the 81.5 μm line [4] and 16 for the 263.4 μm line.

The frequency values obtained are respectively

$$\nu_1 = 1138211.0 \pm 0.5 \text{ MHz and } \nu_2 = 3679592.4 \pm 0.5 \text{ MHz}$$

where the conservative uncertainty of 0.5 MHz takes into account the observed reproducibility of the FIR laser output, which is about ± 0.1 MHz.

The frequency of the 81.5 μm line differs from the value previously reported in the literature [5] by about 1 GHz.

4.- THE SPECTRUM OBSERVED AT 3.7 THz

The RF spectrum of a randomly phase modulated carrier at frequency ν can be approximately described by:

$$S_{RF}(\nu,f) = e^{-\phi_P(\nu)}[\delta(f) + S_\phi(\nu,f)] \qquad (5)$$

For a beat note between two oscillators 1 and 2 one has [6], [1]:

$$S_{RF}^{beat}(\nu_{beat},f) = e^{-(\phi_{P_1}+\phi_{P_2})} \cdot [\delta(f)+S_{\phi_1}(f)+S_{\phi_2}(f)] \qquad (6)$$

Assuming that both S_ϕ and ϕ_P for the FIR laser are much smaller than for the multiplied signal, the analyzed beat note can be considered reproducing the spectrum of the synthesized signal, at least in the white phase-noise region.

The data obtained for the carrier to noise pedestal ratio and for the pedestal width can then be compared to the predicted theoretical limits for the RF spectrum of a synthesized signal [1] reported in Section II (Fig. 1).

Fig. 4 shows the beat note between the 81.5 μm laser line and the 51st harmonic of the klystron as observed at the spectrum analyzer with 10 kHz observation bandwidth. Due to the laser fast linewidth of about 20 kHz the observed signal is reduced by at least 3 dB by the spectrum analyzer bandwidth: this fact is taken into account in Table I.

The values found for $(S/N)_\phi$ are compared to the theoretical prediction (1) in Fig. 1 for both NH_3 lines.

The loss of 7÷10 dB in the $(S/N)_\phi$ value with respect to the theoretical predicted value may be due to a not controlled noise source in the klystron phase-lock loop as well as the unexpected noise pedestal width of 400 kHz.

On the other hand the values of the signal-to-noise floor (S/N) (normalized to 10 mW fir power and 100 kHz B_w) fit well the previous experimental results reported in [7] (Fig. 5).

Both S/N ratios may be further increased in order to allow precise frequency counting.

Table I

	n.	$(S/N)^{10\ mW}_{100\ kHz}$	$(S/N)^{10\ kHz}_{\phi}$	ν_0 (MHz)
263.4 μm	16	58 dB	17 dB	1138211.0±0.5
81.5 μm	51	5 dB	7 dB	3679592.4±0.5

5.- FREQUENCY SYNTHESIS TO HIGHER FREQUENCIES

The 3.7 THz emission line of the $^{14}NH_3$ FIR laser is interesting in designing a frequency synthesis chain project because its eighth harmonic from the measured value turns out to be in coincidence, within 6 GHz, with the 10R(30) CO_2 laser line.

The CO_2 10R(30) line in turn is useful for the connection with the P(7) CH_4 transition.

This fact, together with the possibility demonstrated above of reaching directly the 3.7 THz line from the klystron radiation by means of its 51st harmonic produced by a Schottky diode allows a very simple frequency synthesis chain, whose project is shown in Fig. 6.

REFERENCES

[1] A. De Marchi, A. Godone, and E. Bava - IEEE Trans. IM-30 (1981), 132.

[2] A. Godone, M.P. Sassi, C. Caldera, and E. Bava - Opt. Comm., (to be published).

[3] T.Y. Chang, T.J. Bridges, and E.G. Burkhardt - Appl. Phys. Lett. 17 (1970), 357.

[4] C.O. Weiss, and T. Sakurai - Opt. Comm. 62 (1987), 351.

[5] V.M. Klement'ev, Yu.A. Matyugin, M.V. Nikitin, and B.A, Timchenko - Sov. J. Quantum Electron. 9 (11) (Nov. 1979).

[6] F.L. Walls, and A. De Marchi - IEEE Trans. Instr. Meas. IM-24 (Sept. 1975), 210-217.

[7] C.O. Weiss, and A. Godone - Appl. Phys. B 35 (1984), 199.

CAPTIONS TO FIGURES AND TABLE

Fig. 1: $(S/N)_\phi$ in 1 Hz bandwidth as a function of the synthesized frequency ν.
Experimental points ⊕: Ref. [1], ⧫ this work.

Fig. 2: Experimental set-up used for the heterodyne detection of the FIR radiation.

Fig. 3: Relevant energy level pattern for the 81.5 μm and the 263.4 μm ammonia lines.

Fig. 4: Beat signal between the 51st harmonic of the 72.155 GHz millimeter wave and the 81.5 μm NH_3 line.
Vertical scale: 2 dB/div;
Horizontal scale: 100 kHz/div;
IF-filter width: 10 kHz;
video filter: 10 Hz.

Fig. 5: S/N ratio versus frequency and harmonic numbers.
B_w = 100 kHz, W_{FIR} = 10 mW.
Experimental points ⊕: Ref. [7], ⧫ this work.

Fig. 6: Frequency synthesis chain employing the 81.5 μm NH_3 line.

Table I: Harmonic mixing number, signal-to-thermal noise and signal-to-phase noise ratios and measured frequencies ν_0 of the NH_3 laser 263.4 μm and 81.5 μm lines.

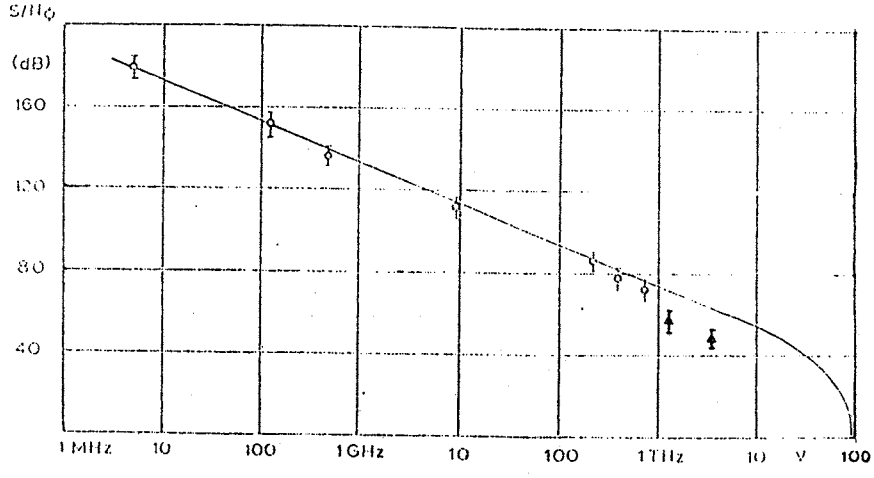


FIG. 1

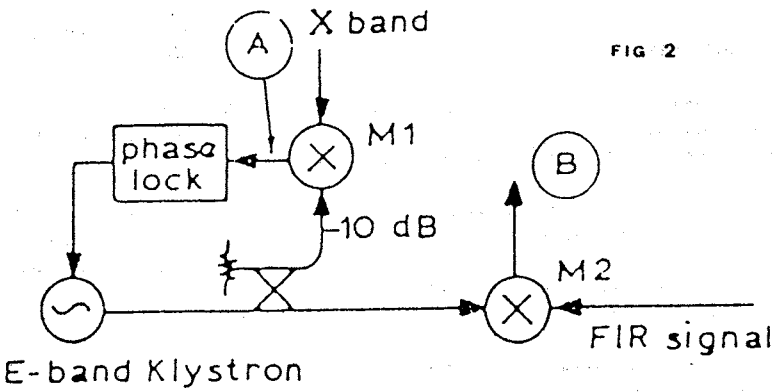


FIG 2

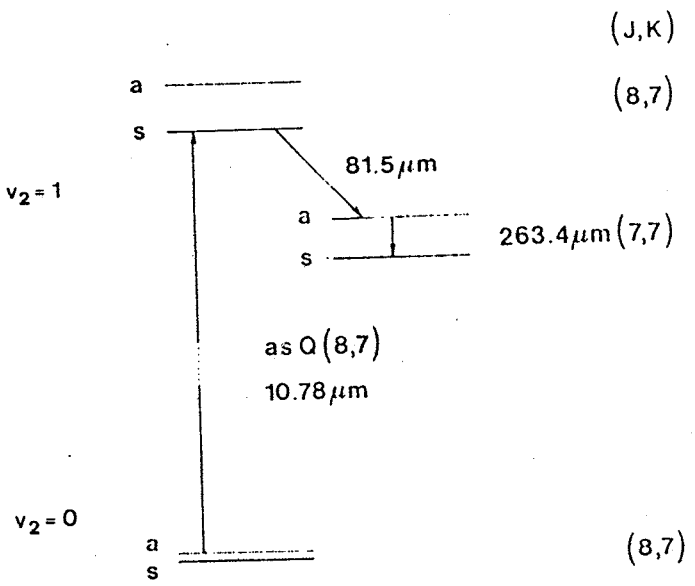
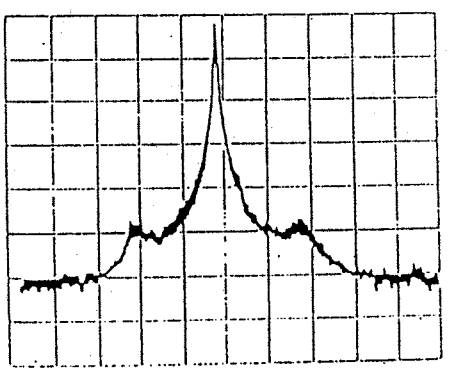


FIG 3

FIG 4



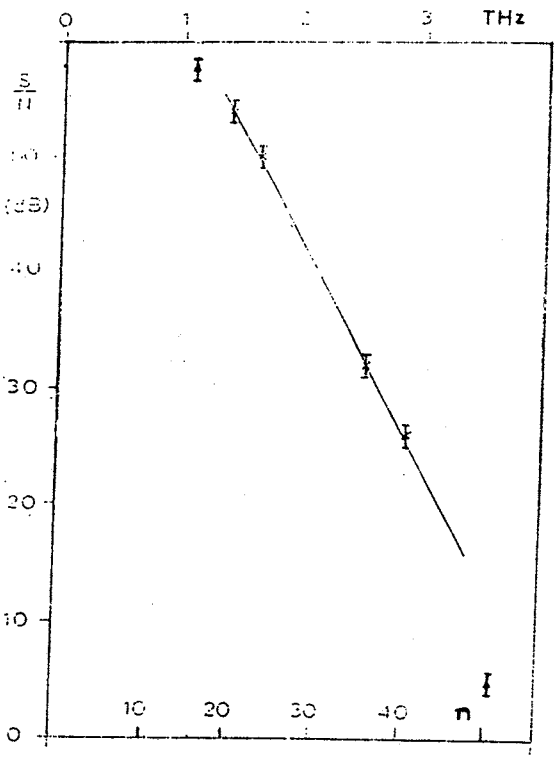

0
1
2
3
THz
S/N
(dB)
40
30
20
10
0
10
20
30
40
n

FIG 5

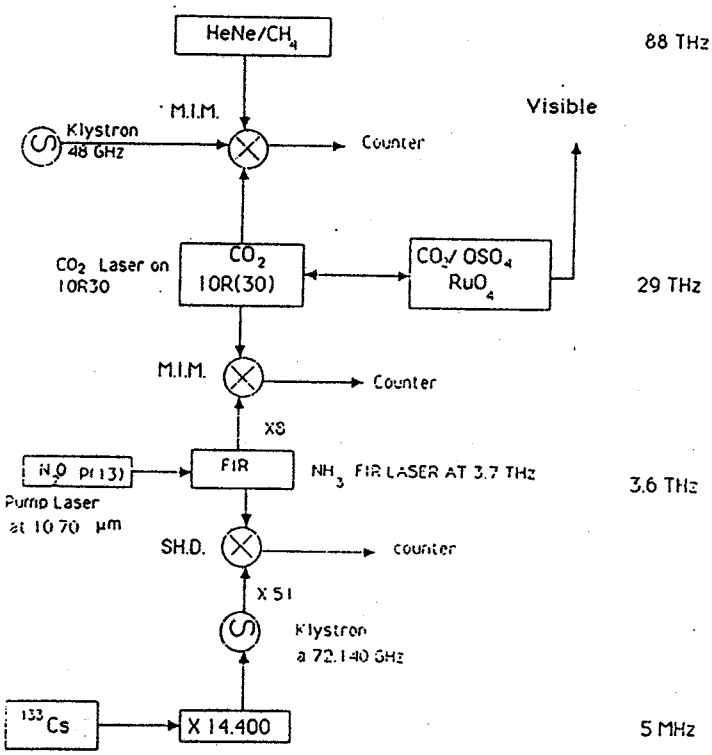

HeNe/CH4
88 THz
M.I.M.
Visible
Klystron
48 GHz
Counter
CO2 Laser on
10R30
CO2
10R(30)
CO2/ OSO4
RuO4
29 THz
M.I.M.
Counter
X8
N2O P(13)
FIR
NH3 FIR LASER AT 3.7 THz
3.6 THz
Pump Laser
at 10.70 μm
SH.D.
counter
X 51
Klystron
a 72.140 GHz
133Cs
X 14.400
5 MHz

FIG 6

THE MANIPULATION OF ATOMIC BEAMS BY LASER LIGHT PRESSURE

W.Ertmer, J.H.Müller, J.Nellessen and K.Sengstock

Institut für Angewandte Physik, Universität Bonn,
Wegelerstr. 8, D-5300 Bonn 1, FRG

Abstract:

During the last few years there were several major experimental developments, which probably will influence the future design of new or improved frequency standards. Among these laser cooling of ions or neutral atoms became most prominent. In this paper we will discuss the techniques for laser cooling and manipulation of neutral atoms (atoms beams) and the resulting benefit for atomic beam frequency standards.

1. Introduction

During the last few years there were three major experimental developments which could influence the future design of novel and/or improved frequency standards: 1) trapping of laser cooled ions [1], 2) frequency stability and accuracy of lasers [2], and 3) laser cooling and manipulation of neutral atoms.
In this paper we will focus on the techniques of laser preparation of atomic beams and its possibilities for the development of future frequency standards. The laser preparation may influence the atomic inner degrees of freedom [3] as well as the phase space density of atomic beams and particularly its mean velocity. The main benefits will be a strongly enhanced signal-to-noise ratio, a much higher Q-value for the frequency discriminating transition, the crucial reduction of the quadratic Doppler effect, and a decreased influence of the 1st order Doppler effect (becoming evident e.g. as phase shifts in atomic beam standard).

2. Cooling of Atomic Beams

The high spectral density of resonant laser light creates light forces sufficient to manipulate the velocity and position of free atoms over large ranges [4]. These forces may also damp the relative motion between free atoms, cooling them down into the μK-region and below [5]. In addition it is possible to trap free atoms at one position of space by radiation pressure only or with the aid of magnetic fields.
Depending on the experimental situation one may select or enhance one of four main contributions to the light force: scattering force, spontaneous force, dipole force in a running wave (gradient force), and forces produced in superimposed laser fields (e.g., standing waves), the intensity and/or polarisation of which is strongly position dependent [6]. In this paper we will restrict the discussion mainly on the application of the scattering force; it should be mentioned, however, that the other light pressure forces or their combinations may improve laser cooling and manipulation considerably [7].
When the laser is tuned into resonance with an atomic transition, the atoms absorb photons and reemit spontaneously photons. These processes are accompanied by the transfer of photon energy $\hbar\omega_l$ (ω_l: frequency of the laser photons in the laboratory frame) and $\hbar\omega_s$ (ω_s: frequency of the spontaneously emitted photon in the moving atomic frame) as well as by the transfer of photon momentum $\hbar\vec{k}_l$ ($\vec{k}_l$: wavevector of the laser photons) and $\hbar\vec{k}_s$ ($\vec{k}_s$: wavevector of the spontaneously emitted photons).
Because of the random direction of emission the momentum recoil by the spontaneously emitted photons averages to zero whereas the atomic momentum change by the absorption of photons adds up constructively to $n\cdot\hbar\vec{k}_l$ (n = number of absorption-emission cycles). The resulting mean velocity change reads $n\cdot\hbar\vec{k}_l/M$ (M = atomic mass). This corresponds to the scattering force $\vec{f}_{sc}$ (atomic momentum change per time):

$$\vec{f}_{sc} = \frac{\hbar\vec{k}_l}{\tau}\,\frac{s}{1+2s} \qquad (1)$$

τ denotes the natural lifetime of the excited atomic state and s denotes the saturation of the transition.
The saturation s is a function of the saturation intensity, the detuning between the Doppler shifted absorption frequency of the moving atom and the laser frequency, the laser beam intensity and intensity distribution, and explicitely a function of time, if the laser frequency is time dependent.
The expression $\tau\cdot(1+2s)/s$ in equ. (1) gives the cycle time.
The scattering force saturates with increasing saturation of the transition to its maximum value

$$\vec{f}_{sc} = \frac{\hbar\vec{k}_l}{2\tau} \qquad (2)$$

Dividing the forces given in equ. (1) and equ.(2) by M gives the corresponding accelerations or decelerations.
The magnitude of the scattering force depends strongly on the resonance condition (otherwise the cycletime becomes to long). This causes two problems:

1. As the atoms change their velocity under the influence of the scattering force, the changing Doppler shift moves them out of resonance.
2. If the excited state decays to a non resonant level, e.g. another fine structrue or hyperfine structure level than the initial one, the scattering force will usually vanish.

During the recent past several experimental solutions to these two basic problems were

discussed [8] and have been realized [9,10], mainly by laser cooling of sodium.
Whereas optical pumping (2.) is a basic problem when using radiation pressure for cooling or manipulation of free atoms the influence of the changing Doppler shift (1.) depends on the experimental situation and is most severe in the case of deceleration or acceleration of atomic beams (s.below).
The basic scheme of atomic beam cooling consists of an atomic beam and a counter-propagating laser beam. When the atoms are decelerated by the successive absorptions of photon momentum from the counterpropagating laser beam, their Doppler shift changes very fast and thus the atoms run out of resonance after the accumulated shift of a few homogeneous line widths. On the other hand the stopping process needs very many photons (e.g. ~20.000, if sodium atoms shall be stopped from an initial velocity of about 600 m/s, using the Na-D-line) and thus optical pumping has to be totally avoided or counteracted.

In the first successful cooling experiments, which produced really a slow atomic sodium beam [9], a longitudinal magnetic field with a decreasing field strength along the beam axis solved both problems. The absorption frequency of atoms entering this magnetic field is Zeeman tuned to higher frequencies compensating the red shift of the atoms counterpropagating to the laser beam. If the magnetic field strength changes along the beam axis in the same manner as the Doppler shift does by the deceleration, the atoms will stay in resonance.
Assuming a constant deceleration a_B within the magnet the velocity v(z) will change (starting with v_0 at z=0) like:

$$v(z) = (v_0^2 - 2a_B z)^{1/2} \qquad (3)$$

The corresponding change of the magnetic field strength B(z) to compensate the z-dependent Doppler shift by the changing Zeeman shift reads then (for details see, e.g. [9]):

$$B(z) = B_0 (1 - 2a_B z/v_0^2)^{1/2} + B_b$$
$$B_b: \text{ bias field} \qquad (4)$$

The difference between the magnetic field strength at both ends of the solenoid defines the magnitude of the velocity interval, which will be slowed down and the laser detuning defines the final velocity of the decelerated atoms. All atoms outside this velocity interval will not be affected. As each location z within the field corresponds, as given in equ. (3), to a resonant velocity, every atom within the affected velocity interval will find its resonant point z and stay in resonance till the exit of the field. All atoms of the decelerated interval will thus have nearly the same velocity at the exit of the magnet. Therefore the resulting velocity distribution is highly compressed (cooled) to some m/s (mK). The final average velocity may be chosen over a broad range including zero.
The magnetic field in combination with the circular polarisation of the cooling beam avoids optical pumping too. The bias field (s.equ.(4)) helps to prevent optical pumping [9].
In successful experiments using this technique for deceleration the velocity distribution of a thermal atomic sodium beam could be compressed to an interval of ≈10m/s corresponding to a temperature below 100 mK [9] with a final mean velocity ranging from a few hundred m/s down to velocity reversal.
The saturation dependent deceleration a(s) reads (s=const.)

$$a(s) = -v_r/\tau \cdot s/(1+2s) \qquad (5)$$

The correponding stopping distance $z_s(s)$ reads

$$z_s(s) = -v_0^2/(2a) = v_0^2/(2a_0) \cdot (2 + 1/s) \qquad (6)$$

with $a_0 = -v_r/\tau$. The stopping time $t_s(s)$ reads

$$t_s(s) = -v_0/a = v_0/a_0 \cdot (2 + 1/s) \qquad (7)$$

The random nature of the numerous spontaneous processes necessary for deceleration will heat the beam. This affects especially the transverse velocity, as the longitudinal velocity is controlled by the velocity dependent scattering force [11].
The resulting transverse heating reads [4]

$$v_t(n) = (n/3)^{1/2} \cdot v_r \qquad (8)$$

with n for the number of resonance cycles.
The final longitudinal velocity width Δv_z reads [11]

$$\Delta v_z = (\hbar\gamma/(2M))^{1/2} \qquad (9)$$

The second general cooling scheme uses fast frequency modulation techniques to compensate optical pumping and to keep the resonance condition (without magnetic fields) [10].
In alkali spectra (e.g. Na) the transition $^2S_{1/2}(F=2) \Rightarrow {}^2P_{3/2}(F=3)$ does not provide a completely ideal two-level system in zero magnetic field; because of the relatively small hyperfine splitting of the upper level and the limited perfection of the circular polarization of the laser light, the atoms can also make the transition $^2S_{1/2}(F=2) \Rightarrow {}^2P_{3/2}(F=2)$ and the upper level can decay to the level $^2S_{1/2}(F=1)$ which is out of resonance. To compensate for this optical pumping a second frequency in the cooling laser beam inducing the transition $^2S_{1/2}(F=1) \Rightarrow {}^2P_{3/2}(F=2)$ can repump the atoms into the level $^2S_{1/2}(F=2)$ via the transition $^2P_{3/2}(F=2) \Rightarrow {}^2S_{1/2}(F=2)$.
This second frequency can be provided as one of two sidebands of a frequency-modulated laser beam; in the case of the sodium D_2 line the difference frequency would be 1712 MHz. For this purpose the laser beam is, for example, sent through an electro-optic phase modulator [10] that is driven at half the desired difference frequency. In case of a sufficient modulation index about 35% of the incoming intensity can be transferred to the first-order sidebands.
The problem of maintaining the resonance condition for the decelerating atoms is solved in this scheme by a fast tuning of the laser frequency synchronously with the rapidly changing Doppler shift. This again can be achieved by electro-optic modulation techniques, if the laser beam with the two frequencies is sent through a second electro-optic modulator, the driving frequency of which is chirped in the right way - in the sodium experiment, for example, from 5 MHz to 1000 MHz within about 1.5 ms. This will produce a pair of sidebands which stays in resonance with the decelerating atoms in both hyperfine sublevels of the ground state, if the carrier frequency is chosen correctly [10].
In the scan method the deceleration a has to

match the scan speed ν_L of the laser frequency; assuming constant deceleration the scan speed is almost constant,

$$\dot{\nu}_L = \frac{|\vec{a}|}{\lambda} \qquad (10)$$

and the frequency varies linearly in time, starting periodically red shifted at the laser frequency ν_s :

$$\nu_L(t) = \nu_s(1+\alpha t) \qquad (11)$$

This frequency ν_s is in resonance with fast atoms v_s at the beginning of each cooling cycle. While the sideband is swept over a frequency interval $\Delta\nu_L$ these fast atoms stay in resonance and slower atoms will get into resonance. Thus at the end of each cooling cycle the corresponding velocity interval $\Delta v = \Delta\nu_L/\lambda$ is compressed into a narrow velocity distribution (s. below) at the final velocity $v_s - \Delta v$.
In our Na-experiments the scan interval was limited by our microwave equipment to ≈1GHz corresponding to a velocity interval of ≈600 m/s. After each cooling scan, which takes about 1 ms, the swept sideband moves very fast within ≈200 μs back into the starting position beginning the next cooling cycle.

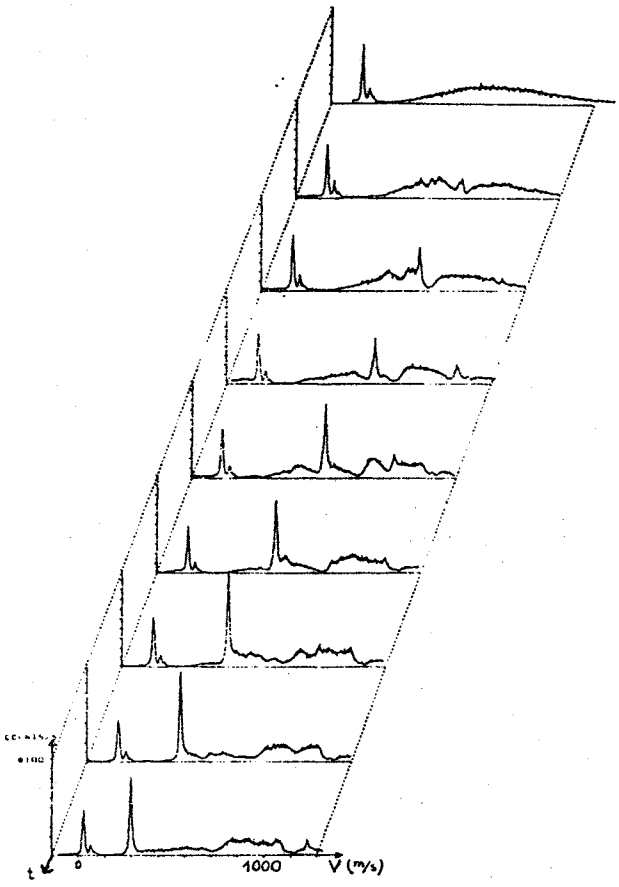


Fig.1 : Evolution of the velocity distribution during the frequency sweep

During this time fresh incoming, uncooled atoms fly only ≈1 m before they get into resonance in the next cooling cycle. Thus all atoms slower than the starting velocity v_s and within Δv may be cooled.
In the first experiments with this second scheme [10] the resulting temperature within the cooled atomic beam was below 50mK with a density of 10^6 atoms per cm^3. As the slow atoms move only a short distance during the cooling cycles, the resulting pile-up of slow atoms forms a nearly constant flow of cold atoms as in the previous scheme. Fig.1 shows a typical experimental result.

3. Laser Manipulation of Atomic Beams

The total atomic beam is not yet monoenergetic, because it still contains the unaffected high velocity wing of the Maxwell-Boltzmann distribution and it is superimposed by the strong cooling laser light. Furthermore the transverse beam temperature of the laser cooled atoms is enlarged by diffusion processes during deceleration.
Therefore the separation of the laser cooled atoms from the residual atomic beam by an improved laser deflection technique provides the right tool to generate a really monoenergetic, slow atomic beam undisturbed by strong laser light [12]. As a by-product it offers transverse cooling in the plane of deflection too.

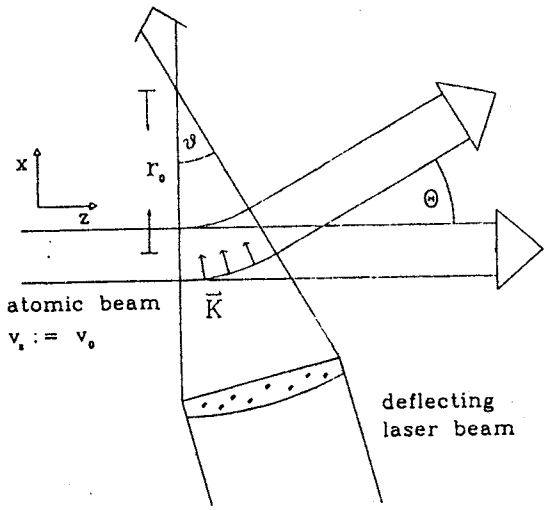


Fig.2: Laser deflection of a slow atomic beam

Fig.2 shows the experimental scheme with the transverse deflecting laser field, formed by a cylindrical lense. If all experimental parameters are chosen correctly, the $\vec{k}$-vector of the deflecting laser field stays perpendicularly to the atomic orbit of propagation balancing the centrifugal force. Therefore the angle of deflection α is given by the angle of divergence of the deflecting light field which may reach more than 40° with one lense.
The radially damped motion is balanced only for a limited initial velocity interval of the atomic beam; thus the laser decelerated, cold part of the atomic beam can be chosen for deflection selectively forming a well separated monoenergetic beam undisturbed by

cooling laser light [12]. The chirp-method is particularly convenient for this type of experiment, as one can use a single dye laser for cooling and deflecting, optimizing the laser frequency for the deflection process. As in the case of atomic beam cooling optical pumping has to be avoided or has to be compensated for.

In the first Na-experiment [12] the deflecting laser beam was split off the cooling beam behind EOM-1, providing the necessary repumping sideband. Fig.3 shows the experimental results for the deflection of a laser cooled sodium beam with an angle of 30-35° [12]. It demonstrates the 100% efficiency of the deflection mechanism.

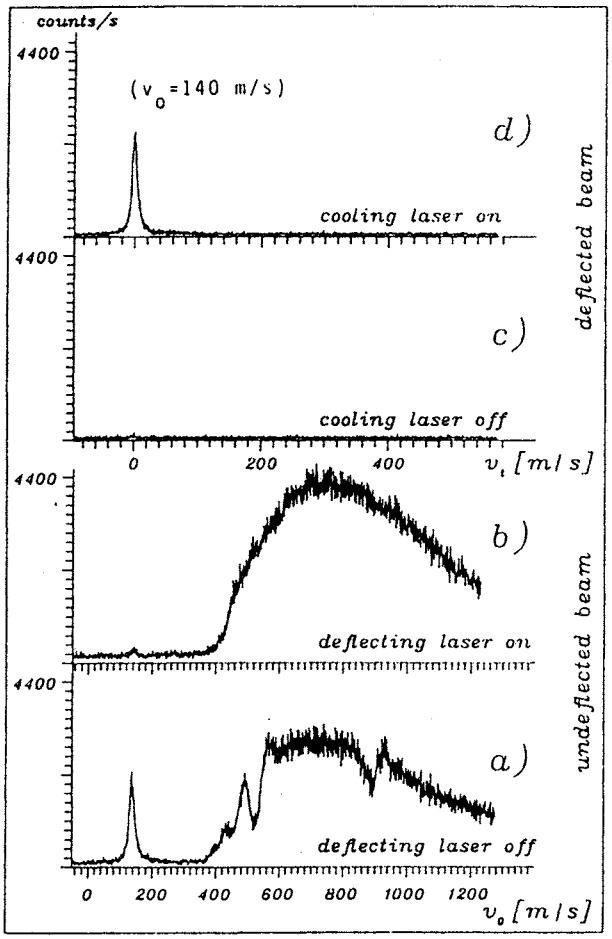


Fig.3: Cooled velocity distributions of the deflected and the undeflected beam

For a further improvement of the beam quality, a one-dimensional "optical molasses" [13] (formed by two counterpropagating laser beams, tuned about half a linewidth below resonance) could be implemented, which collimates the transverse velocity components perpendicular to the plane of deflection too.

A more drastic increase in density can be expected from the application of a focussing scheme analog to the 3D trap scheme in [14] : A 2-dimensional optical molasses setup consisting of pairs of σ^+-σ^- light beams within a superimposed transverse magnetic field from a linear quadrupole [15]. This provides a restoring force to the beam axis. The refocusing and subsequent collimation should yield an atomic beam radius down to the µm region.

Such a dense, monoenergetic atomic beam with a "quantum-limited" divergence (in the case of sodium typically a few mrad for v_0=100 m/s), enables a lot of promissing improvements for "conventional" atomic beam standards as well as for new e.g. optical frequency standards.

4. Discussion of some examples

There are three main groups of elements which offer a level scheme suited for laser cooling and manipulation: alkalines, alkaline earths and noble gases (in their metastable states).

Some candidate atoms for frequency standard research using laser cooling techniques are discussed e.g. in [16] ; the improvement of future frequency standards by the methods presented here are discussed to some extent in [17].

The most promising element is certainly hydrogen (or antihydrogen) with its 1s⇒2s two-photon transition. But this investigation presupposes the solution of the severe technical problem generated by the wavelength of the 1s⇒2p 121.5nm cooling transition [18].

Therefore we suppose that in the near future laser prepared beams of magnesium or calcium will be investigated in regard to their possibilities as microwave or optical frequency standards. In the following we will discuss magnesium for demonstration in more detail:

Mg provides the $^1S_0 \Rightarrow {}^1P^\circ_1$ cooling transition at 285nm as a pure two level system; the 285nm radiation can be generated by frequency doubling of 570nm dye laser radiation. Because of the "heavy" UV-photons and the short live time of the upper cooling level $^1P^\circ_1$ (≈2ns) the deceleration distance for a velocity interval from e.g. 1200m/s to 50m/s is rather short, about 10-15cm, depending on the saturation chosen. This simplifies the experiment by the possibility to use a well tailored short solid magnet (s. eq.(4)) and circular polarized light for the down slowing of the atomic beam [19]. The slow and cold part of the beam (e.g. 50m/s) will be separated (deflected) and collimated to the Doppler limit by a second 285nm laser beam to form a precision atomic beam (transverse velocity spread $\Delta v_T \approx$1m/s and longitudinal $\Delta v_L \approx$1m/s). With the aid of the intercombination line $^1S_0 \Rightarrow {}^3P^\circ_1$ (457 nm) this spreading may be reduced even below the recoil limit $\Delta v = v_r \approx$4cm/s [19,20].

Mg offers two interesting clock transtions [16]: the optical clock at the 457 nm intercombination line and the microwave transition 601 GHz between the metastable states $^3P_1 \Rightarrow {}^3P_0$ [21]. Because of the long nat. lifetime for $^3P_1$ (≈4.6ms [22]) the width of the transitions is about 35Hz (FWHM) giving the very good Q-values of Q(457nm)≈10^{13} and Q(601GHz)≈10^{10}. In case of a thermal Mg-beam (v≈850m/s) the second order Dopplereffect ($\sim v^2$) produces a relative shift $\delta\nu_{2D}/\nu$ of $\delta\nu_{2D}/\nu \approx -4\cdot10^{-12}$ and a relative broadening $\Delta\nu_{2D}/\nu$ of $\Delta\nu_{2D}/\nu \approx 7\cdot10^{-12}$. In case of the laser prepared beam these values are reduced to $\delta\nu_{2D}/\nu \approx -1.5\cdot10^{-14} \pm 5\cdot10^{-16}$ and $\Delta\nu_{2D}/\nu \approx 10^{-15}$ for v=50±1m/s ($\delta\nu_{2D}/\nu \approx -1.4\cdot10^{-14} \pm 4\cdot10^{-17}$ for v=50ms±0.04 m/s).

To reduce the time of flight broadening it is necessary to apply separated fields

(Ramsey-type measurements) interacting with the atomic beam (assuming two fields in the following for simplicity). In order to limit this broadening to the natural linewidth $\Delta\nu_n$ ($^3P_1$, the separation has to be larger than Δl, $\Delta l \approx v/(2\Delta\nu_n)$. For a thermal beam this asks for the totally unrealistic value $l \approx 12$ m. In case of the laser prepared beam (v=50m/s) a separation of $l \approx 70$ cm would be necessary. This is a realistic value for the microwave transition and within a factor of 2-3 for the optical transition. Also the magnetic field influence as well as other systematic effects (see e.g.[23]) may be reduced (controlled) below the 10^{-15} level. Tunable laser sources with a spectral width of about 1 Hz ($\approx 10^{-15}\ \nu$) have been demonstrated recently [24].
The techniques of laser preparation, as described here, will allow the "old dream" of a fountain clock [25]: A slow and cold atomic beam may be bent optically as described above (or magnetically) into a vertical direction. If this beam has an average velocity of about 5 m/s, it will rise in the gravitational field 1.25 m and fall down again, reaching the "starting point" about 1s later. When the interrogating field is located at this starting point, the atoms walk twice through the same field, forming a "fountain Ramsey-fringe geometry". This cancels cavity phase-shift effects and increases the Q-value.
A microwave cavity as well as a laser beam should be used for an optical "fountain Ramsey-fringe geometry", offering a Q-value of the order of 10^{15} for suited transitions.

References

[1] D.J.Wineland, W.M.Itano, R.S.VanDyvk Jr.: Adv.At.Mol.Phys.19. 136 (1983)
[2] Ch.Salomon, D.Hils, J.L.Hall, JOSA B5 1576 (1988)
[3] G.Avila, V.Giordano, V.Candelier, E.de Clercq, G.Theobald, and P.Cérez, Phys. Rev.A36 (1987) 3719
[4] V.S.Letokhov and V.G.Minogin, Phys.Rep.73 (1981) 1
[5] D.J.Wineland, W.Itano, Phys.Rev. A20, 1521 (1979)
[6] J.Dalibard et.al., 11th Conf.on Atomic Physics, Paris, July 1988, to be published (Wiley 1989)
[7] D.E.Pritchard, E.L.Raab, V.Bagnato, C.E.Wieman and R.N.Watts, Phys.Rev.Lett. 57 (1986) 310
[8] See, e.g. Laser-Cooled and Trapped Atoms, Proc. of the Workshop on Spectr. Applications of Slow Atomic Beams, 1983, ed. W.D.Phillips, Natl.Bur.Stand. (U.S.) Spec.Publ.653
[9] J.V.Prodan, A.Migdall, W.D.Phillips, I.So., H.Metcalf, J.Dalibard, Phys.Rev. Lett. 54 (1985) 992;
A.Migdall, J.V. Prodan, W.D.Phillips, T.H.Bergeman, H.J.Metcalf, Phys.Rev.Lett. 54 (1985) 2596
[10] W.Ertmer, R.Blatt, J.L.Hall, M.Zhu, Phys.Rev.Lett.54, 996 (1985);
S.Chu, L.Hollberg, J.E.Bjorkholm, A.Cable, A.Ashkin, Phys.Rev.Lett.55 (1985) 48
[11] H.Wallis and W.Ertmer, J.Phys.B21, 2999 (1988)
[12] J.Nellessen, J.H.Müller, K.Sengstock, W.Ertmer, 1989, submitted to JOSA B
[13] see e.g. S.Chu et.al. in Laser Spectroscopy VIII ed. W.Persson, S.Svanberg
[14] E.L.Raab et.al., Phys.Rev.Lett.59, 2631 (1987)
[15] J.Nellessen, J.Werner, W.Ertmer, to be published
[16] W.Ertmer, R.Blatt, J.L.Hall, Prog.Quant.Electr.8, 249 (1984)
[17] W.Ertmer, S.Penselin, Metrologia 22,195 (1986)
[18] W.Ertmer, H.Wallis, Hyperfine Interactions 44,319 (1988)
[19] G.Hennig, D.Herrendörfer, P.Heldsdörfer, J.H.Müller, K.Sengstock, W.Ertmer, to be published
[20] H.Wallis and W.Ertmer, subm.to JOSA B
[21] E.Bava, A.Godone, G.Giusfredi, C.Novero IEEE J.Quantum Electronics QE-23,455 (1987)
[22] H.S.Kwong, P.L.Smith, W.H.Parkinson Phys.Rev.A25, 2629 (1982)
[23] F.Strumia, Metrologia 8, 85 (1972)
[24] R.Kallenbach, C.Zimmermann, D.H.McIntyre, T.W.Hänsch, R.G.Devoe, Opt. Comm. 70,56 (1989)
[25] J.R.Zacharias, Phys.Rev.94,751 (1954)

SHORT-TERM STABILITY OF A COMMERCIAL OPTICALLY PUMPED Cs TUBE

P. Thomann - F. Hadorn

OSCILLOQUARTZ SA - NEUCHATEL

INTRODUCTION :

This progress report on the development of a commercial -size optically pumped Cesium is divided in two parts. In the first part, devoted to optical pumping efficiency calculations using rate equations, we show how the results for pumping with σ-polarised light can be derived from the more straightforward results for π -polarised light through simple symmetry considerations. These considerations are relevant to the case of pumping in a weak magnetic field. In the second part we report on improvements in the signal-to-noise ratio of the Ramsey signal obtained with the 4-4 σ transition. The reduction of stray light and its associated noise brings the signal-to-noise ratio to 2500 and the calculated short-term stability to $2.5.10^{-11}\,\tau^{-1/2}$.

I OPTICAL PUMPING EFFICIENCY WITH σ AND π POLARISATION.

Various calculations of optical pumping efficiencies have been published in connection with state-selection of Cs atoms [1-5]. The basic approach to this calculation is to solve for the complete density matrix of the problem, including optical, hyperfine and Zeeman coherences [2,4].
This approach is particularly appropriate when the pumping laser light is highly monochromatic (i.e. the correlation time of the laser light is longer than the duration of an optical pumping cycle) because in this case the optical coherences and the hyperfine coherences - in the case of two-laser pumping - must be taken into account.
We consider here the case of optical pumping with one broadband laser tuned to any of the six possible hyperfine transitions.
The "broadband laser" assumption means that the linewidth of the laser is assumed to be larger than the powerbroadened absorption profile of the atomic beam. Commercially available lasers have a linewidth of 40 MHz, which is much larger than the 5 MHz linewidth of the $6^2P_{3/2}$ level of Cs. In addition the spectral power density required to reach optical saturation is at least 0.44 mW/MHz cm^2, depending on the specific hyperfine transition. It can easily be seen that power broadening cannot be reached in a realistic beam diameter with the typical 10 mW of total available laser power. As a consequence the "broadband laser" assumption is justified and the optical coherences can be completely neglected. Another assumption consists in neglecting hyperfine coherences, which is justified in the case of one-laser pumping.
Finally the Zeeman coherences must be considered. Two cases are of particular interest :
a) π -polarised pumping light : in this case Zeeman coherences are never produced by the optical pumping. Since they are absent in the isotropic initial condition, they can be left out of the calculation. The density matrix evolution is then completely described by a set of rate equations for the populations :

$$\dot{\rho}_{F,m} = -\rho_{F,m}\, n_F^{F'}\, \eta_{F,m}^{F',m} + \sum_{m'} q_{F',m'}\, \eta_{F,m}^{F',m'} \left(1 + n_F^{F'} \delta_{mm'}\right) \quad (1)$$

$$\dot{q}_{F',m} = \sum_{F} \left[\rho_{F,m}\, n_F^{F'}\, \eta_{F,m}^{F',m} - q_{F',m} \left(1 + n_F^{F'} \eta_{F,m}^{F',m'}\right) \right] \quad (2)$$

where $\rho_{F,m}$ = population of ground-state $|F,m\rangle$
$q_{F',m}$ = population of excited state $|F',m\rangle$
$n_F^{F'}$ = saturation parameter of the pumping light
$\eta_{F,m}^{F',m'}$ = normalised value of the dipole matrix element squared

$$= \frac{|\langle F',m'|\vec{D}|F,m\rangle|^2\, \omega_{FF'}^3}{3\pi\varepsilon_0 \hbar c^3 \gamma} \qquad \left(\eta_{4,4}^{5,5} = 1\right)$$

γ = spontaneous decay rate = $3{,}3.10^7\,s^{-1}$
$\dot{}$ = derivative with respect to the dimensionless parameter γt.

The number of levels to be included in the differential system above depends on the transition chosen for laser excitation. In two cases, namely F-F' = 3-2 and 4-5, there are only two levels to be included in the calculation (cycling transitions); in all other cases three levels must be included : the two ground levels and the excited level coupled by the pumping light.
Because the light is π -polarised, the equations for states of opposed m-values are identical. Since the initial populations of such pairs of states are also equal, the equations with negative values of m can be dropped which leads to a considerable simplification of the differential system.
Starting with a beam of Cesium atoms with all ground states equally populated, the differential system can be numerically integrated for each of the possible pumping schemes. The final inversion $\delta_T = \rho_{4,0} - \rho_{3,0}$ of the clock states is then deduced from the equilibrium populations.
Results are shown in Table I, column 1-3.

b) σ-polarised pumping light. Zeeman coherences are created by the pumping process. They will be negligible only if the static magnetic field B ("C-field") is strong enough, i.e. if the Larmor frequency is much greater than the inverse of the pumping duration T

$$g_F \mu_B \cdot B \cdot T \gg 1 \qquad (3)$$

With $g_F = 1/4$; $\mu_B = 2\pi.1{,}4$MHz / Gauss and a typical interaction time T = 3μs, the magnetic field required is

$$B \gg 150\text{mGauss}$$

In a device where the C-field is homogeneous over the whole atomic beam, this condition is very likely not to be satisfied and Zeeman coherences should be taken into account.

The case of σ-pumping in low C-field is, on closer examination, essentially identical to π-pumping, as illustrated in fig.1. Since the magnetic field is negligible, the only anisotropy in the system is introduced by the pumping light so that we can, by a suitable rotation of coordinates, bring the quantisation axis into coincidence with the polarisation vector of the incoming light, perform the calculation as above since the polarisation is now π, then rotate back to the original reference frame to obtain the result. Such a procedure involves only rotation matrices R for the F = 3 and F = 4 multiplicities, which are readily evaluated for a $\pi/2$ rotation. The advantage of this procedure is that the effect of Zeeman coherences is taken into account without any modification to the rate equations (1), (2). The specific rotation must be chosen so as to bring the quantisation axis parallel to the polarisation vector: In fig.1, the polarisation vector is in the Oy direction and the rotation is around the Ox axis.
For each multiplicity F = 3 and F = 4, the density matrix of the σ-pumped atoms is expressed by

$$\rho_\sigma^F = R_F^{-1} \rho_\pi^F R_F \qquad (4)$$

where

$$R_F = \exp\left(-i\frac{\pi}{2}\hbar^{-1}F_x\right) \qquad (5)$$

and F_x is the x component of the spin F operator. Using eq. (4), we can express the diagonal elements $P_{F,m}(\sigma)$ of ρ_σ^F as linear combinations of the populations $p_{F,m}(\pi)$:

$$P_{4,0}(\sigma) = \frac{1}{64}\left(35p_{4,4}(\pi) + 20p_{4,2}(\pi) + 9p_{4,0}(\pi)\right) \qquad (6)$$

$$P_{3,0}(\sigma) = \frac{1}{64}\left(40p_{3,3}(\pi) + 24p_{3,1}(\pi)\right), \qquad (7)$$

from which the clock-state inversion
$\delta_\sigma = P_{4,0}(\sigma) - P_{3,0}(\sigma)$ can be readily evaluated.

The equilibrium inversion in low-field σ-pumping is listed in table 1 for all six optical transitions (column 6). For comparison, the corresponding values of the inversions calculated without Zeeman coherences, which are valid in high field, are listed in column 7 (from ref [2]). The values of δ_σ obtained here from π-polarisation results agree with results of standard calculations including Zeeman coherences in low field [5].
The method outlined above establishes a simple connection between the state-selection processes with σ and π polarisation and shows the importance of Zeeman coherences in σ pumping in a weak C-field.
This method can also be applied to evaluate the sensitivity of the equilibrium populations to misalignment between optical and static magnetic fields, since rotation matrices can readily be approximated for small rotation angles. In the definition (5) of the rotation matrix R_F we replace the rotation angle $\pi/2$ by a small angle α and the power series approximation of the exponential is used, to second order in α, to evaluate ρ_α^F as function of ρ_π as in eq.(4).
For one-laser pumping, sensitivity to misalignment between laser and C-field is fairly low, the two most sensitive cases being the 3-3(π) and 4-4(π) transitions, for which the inversion decreases by 12% and 46%, respectively, for α=0.1 radian. A higher sensitivity for these two specific transitions is probably linked to the zero optical transition probabilities $\eta_{3,0}^{3,0}$ and $\eta_{4,0}^{4,0}$ which have an important effect on the pumping efficiency only if α=0. Sensitivity of the inversion in the case of σ pumping can be evaluated in a similar fashion. The relative change of inversion is less than 3% (for α=.1) for all 6 transitions.

II EXPERIMENTAL STATUS.

The experimental apparatus represented schematically in fig 2. A single laser is used for pumping and detection. A rotatable halfwave plate placed between the laser and a polarising beam splitter allows to vary the relative amount of power incident in the two regions of interaction with the atomic beam. The polarisation of both the pump and probe beams is independently controlled by two additional half-wave plates. The Ramsey cavity is made of an E-bend waveguide section; the length of the cavity is 13 cm. The total length of the atomic beam is 30 cm. A hot wire filament is used to calibrate the atomic beam flux. The collecting optics is made of two concave mirrors, one elliptical and one spherical, facing each other to form a "closed" cavity. The interaction region is centered at one focus of the elliptical mirror while a hole in the spherical mirror and the magnetic shield are at the second focus. The detector is placed as near to the hole as possible. Taking into account the holes in the mirrors for the laser beam and atomic beams, the reflectivity of the mirrors(60%) and the non optimal position of the detector, the overall collection efficiency is estimated to be 25%. In a previous stage of the experiment [6,7] we found that the signal-to-noise ratio of the Ramsey signal was limited by intensity fluctuations of the background light scattered towards the detector by imperfections of the optical system. The addition of light baffles along the laser beam in front of the the interaction region prevents light diffracted out of the parallel beam from reaching the detector. An efficient light trap which uses a reflecting surface instead of a diffusing surface as a first absorbing medium [8] also greatly reduces the amount of scattered light. The background level amounts to 10^{-6} of the incident laser power, and is comparable to the fluorescent signal. Thus intensity fluctuations from background light are a negligible contribution to the total noise (fig 4). With this improvement, the signal-to-noise ratio in a 1Hz bandwith is now 2500. From this value and from the measured linewidth of 950Hz of the Ramsey signal (fig 3), we calculate a short-term stability of $2.5.10^{-11}\ \tau^{-1/2}$.
The main contribution to the total noise is associated with the fluorescence signal. Other contributions, which can be independently measured (such as the intensity noise from background light and detector noise) or derived from measured parameters (such as shot noise from background and signal), amount together to a value which is 4 times smaller than the measured fluorescence noise (fig 4). This suggests laser frequency noise as the most probable cause for the observed fluorescence noise. We have indeed measured a considerable amount of excess frequency noise in the diode laser used in these experiments : The lower curve in fig 5 is the laser frequency noise one would expect from the measured injection current noise converted into frequency noise (3GHz/mA). The upper curve shows the measured frequency noise, using a Cs absorption cell as a frequency discriminator. At the Fourier frequency f_m = 137Hz used for the quartz oscillator frequency servo, the measured noise excedes the white noise by a factor of 5.

As a conclusion we note that in the present state, an optically-pumped Cs tube of commercial size shows a combined S/N ratio and linewidth leading to the same short-term stability as existing magnetic-deflection tubes. Considerable improvement is possible in particu- through a better control of the fluorescence noise.

REFERENCES :

[1] M. Arditi, et al. , J. Phys. D11, 2465 (1978)
[2] E. de Clercq et al. J. Phys. 45, 239 (1984)
[3] G. Avila et al., Phys. Rev 36, 379 (1987)
[4] E. de Clercq, P. Mangin, these Proceedings
[5] G. Théobald et al. Proc of the 4th Symposium on Frequency Standards and Metrology Ancona, 5-9 sept. 1988, P.110 (Springer, 1989)
[6] P. Thomann et al. Proceeding of the 2nd EFTF P 513-520 (March 1988)
[7] P. Thomann et al. Proc. of the 4th Symposium on Frequency Standards and Metrology, Ancona 5-9 sept 1988, p 392-394 (Springer 1989)
[8] E.J.Breneman, Applied Optics 20 , 1118 (1981)

This work was supported by the European Space Agency (ESTEC)

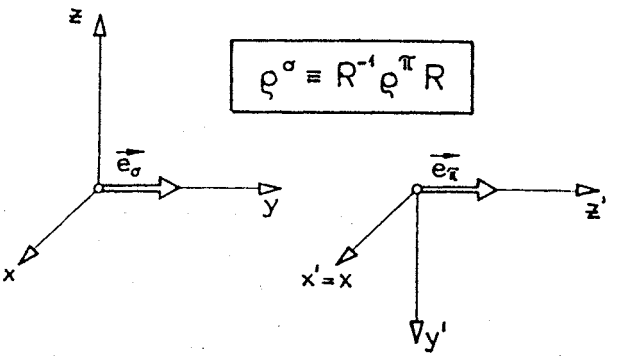


Figure 1. Rotation of coordinates showing the correspondance between π and σ pumping in a weak magnetic field; $\vec{e}$ = polarisation vector of laser $\vec{E}$-field.

HYPERFINE LEVELS COUPLED BY LASER F-F'	π POLARISATION			σ POLARISATION			
	CLOCK-STATE POPULATIONS		INVERSION	CLOCK-STATE POPULATIONS		INVERSION	
	p(4,0)	p(3,0)	δ_{π}	p(4,0)	p(3,0)	δ_{σ} low field	δ_{σ} high field
3-2 (1)	.063	0	+.063	.063	.137	-.074	
3-3	.073	.159	-.086	.098	.001	+.097	0.122
3-4	.137	0	+.137	.097	0	+.097	0.099
4-3	0	.121	-.121	0	.078	-.078	-.154
4-4	.116	.090	+.026	.016	.130	-.113	-.155
4-5 (1)	.199	.063	+.136	.041	.063	-.022	
column	1	2	3	4	5	6	7 (2)

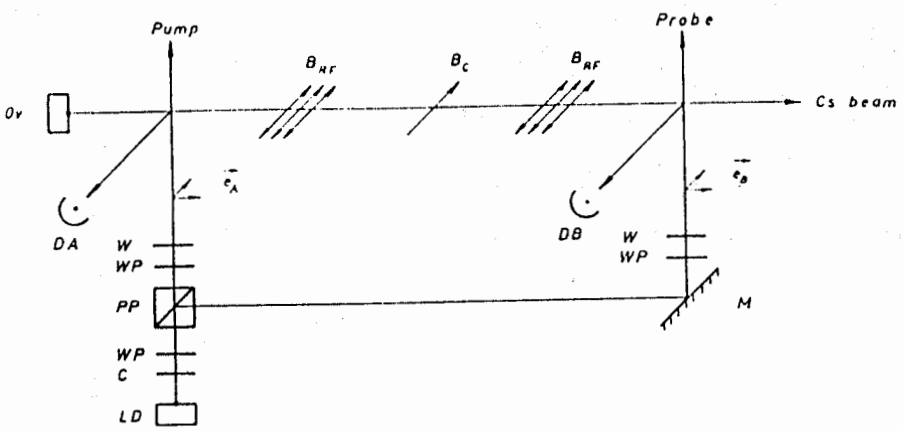
(1) cycling transitions

(2) from ref [2] (no Zeeman coherences)

TABLE 1

Equilibrium populations of clock states and clock inversions after optical pumping of the ground levels of cesium by one broadband laser

(D_2 line; λ = 852.1 nm)



Ov	: Cs oven	Probe	: probe laser beam (B)	WP	: waveplate
Pump	: pump laser veam (A)	DA, DB	: photodetectors	PP	: polirizing prism
B_{RF}	: microwave interaction regions	$e_{A,B}$	: possible polarization of laser beams	M	: mirror
B_C	: C-field	W	: windows in vacuum envelope	C	: collimating optics
				LD	: Laser diode

FIGURE 2

Schematic diagram of optically pumped Cs tube (one laser)

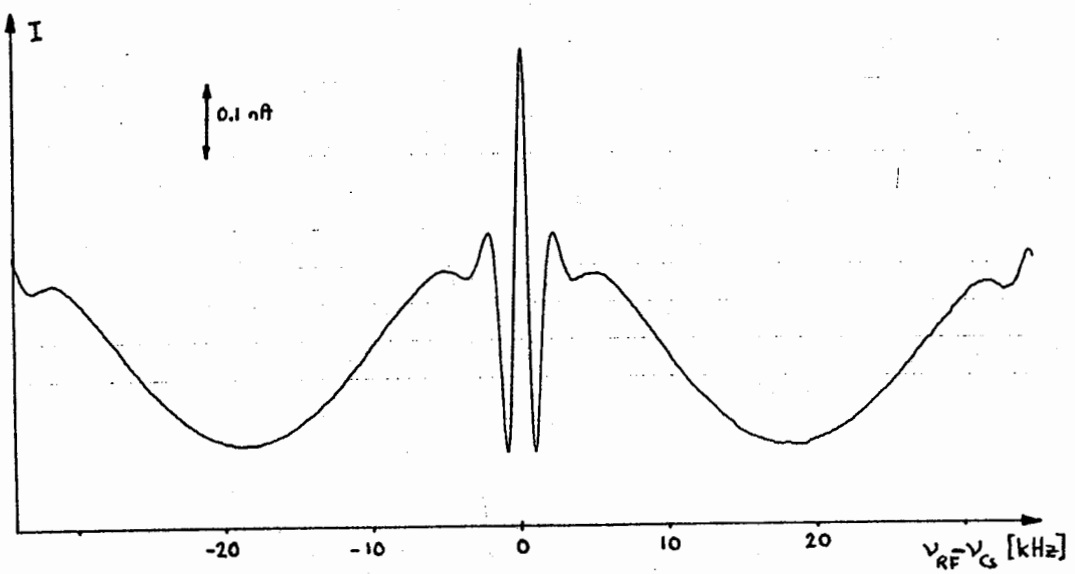


FIGURE 3

Ramsey signal (4-4σ transition for pumping and detection). Background level is 1.6nA

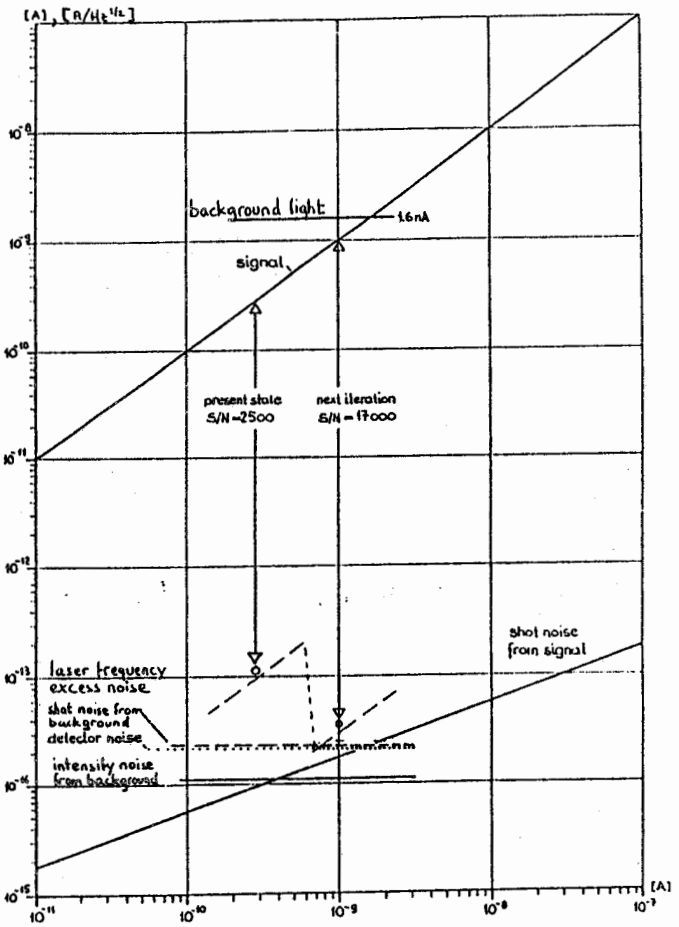


FIGURE 4. Graphical representation at the contributions to S/N ratio, present and projected.

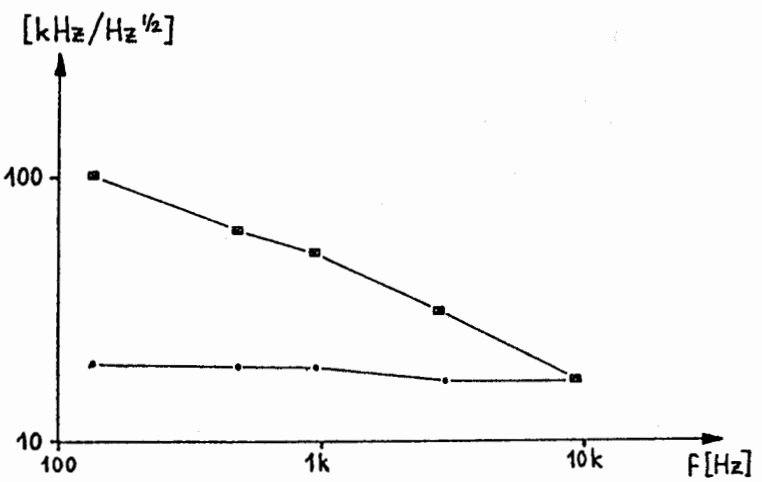


FIGURE 5. Laser frequency noise vs Fourier frequency
lower curve : calculated from experimental injection current noise
upper curve : measurement

FLUORESCENCE ANISOTROPY EFFECTS IN AN OPTICALLY PUMPED CESIUM BEAM RESONATOR

G. THEOBALD, N. DIMARCQ, V. GIORDANO

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

ABSTRACT

In an optically pumped cesium beam resonator, the clock signal-to-noise ratio depends on the collection efficiency of the detecting set-up. When the optical set-up defines a limited collection solid angle, anisotropy properties of the fluorescence emission may be revealed. Consequently the collection efficiency depends on the direction of the magnetic field with respect to the axis of symmetry of the collection set-up and on the excitation characteristics such as the optical transitions and laser polarizations.

I. INTRODUCTION

Since an optically pumped cesium beam frequency standard using laser diodes was proposed [1], a lot of basic experiments on this new method have been carried out. They confirm the great capability of optical pumping methods to improve the frequency performances of cesium clocks [2-7].

In all the reported experiments, the clock signal-to-noise ratio is limited by several noise sources which are :

(1) the detector noise and the current converter noise

(2) the fluorescence noise linked to the laser frequency noise [3,8]

In the present state of the art, the highest clock signal-to-noise ratio was obtained in a basic experiment using a single laser diode for pumping and detection purposes tuned to the $3 \rightarrow 3$ transition and σ polarized. The reported value is 10^4 and the signal-to-noise ratio is limited by the fluorescence noise [9]. Generally, the signal-to-noise ratio is critically related to the photon collection efficiency γ , which represents the proportion of fluorescence photons effectively collected to the number of photons emitted in the whole space. When γ is rather low, typically of the order of 10 % with a collection optics using aspherical lenses, its value depends not only on geometrical parameters such as the solid angle accepted by the collection set-up but also on the excitation scheme characterized by the atomic transition and laser polarization. The latter dependence is due to anisotropy properties of the fluorescence emission and we have investigated the influence of the anisotropy on the signal available in an optically pumped cesium beam resonator.

In this paper we first present the different configurations for the optical detection of the clock signal. Then we describe how the anisotropy could modify the collection efficiency. Finally, we point out the particular situations where some excitation schemes take advantage of anisotropy.

II. CONFIGURATIONS FOR THE DETECTION OF THE CLOCK SIGNAL

In an optically pumped cesium beam tube, three privilegied direction may be defined as :

- the laser beam axis Ox
- the atomic beam axis Oy
- the axis of symmetry of the optical collection set-up Oz.

The static magnetic field $\vec{B}$ which fixes the quantization axis may be directed along one of these three axes. Consequently we can classify the collection configurations encountered in operating machines into (Fig. 1) :

(a) $\vec{B}$ along the atomic axis [9,10]
(b) $\vec{B}$ along the laser beam axis [6]
(c) $\vec{B}$ along the axis of the collection set-up [2,4,5,7,11].

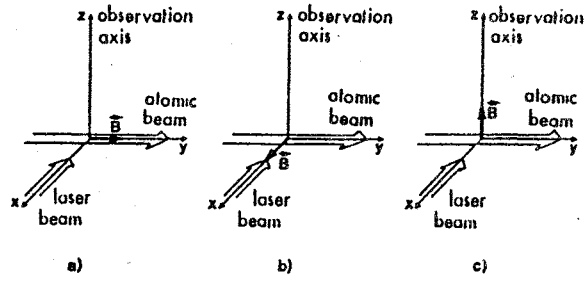


Fig. 1. Schematic collection configurations encountered in operating machines.

The γ value results from the technique used to collect the photons. When large mirrors are used [4,9,10], γ can reach values as high as 50 %. Although the use of lenses provides lower γ values [5,7], this technology offers the advantage of reducing appreciably the laser stray light.

An anisotropic contribution to the emitted fluorescence arises due to the existence of a preferred direction for the laser excitation. It is found that the anisotropic part is cancelled when the fluorescence is observed in the whole space. Such is almost the case of a collection achieved by mirrors with a large efficiency. However, when collection is made inside a restricted portion of the space, anisotropic features of the fluorescence may be revealed. It is therefore of interest to study their influence on the detection of the clock signal.

III. ANISOTROPY OF THE FLUORESCENCE LIGHT. GENERAL CONSIDERATIONS

The laser excitation transfers cesium atoms on the excited state. Subsequent spontaneous decay is observed. Due to the laser excitation, the spontaneous fluorescence emission exhibits anisotropy features resulting from the creation of alignment or quadrupolar momentum in the excited state [12].

Let us define $I(z', d\Omega)$ the number of photons of any polarization emitted by one atom and collected in unit time by an optical set-up defining a small solid angle $d\Omega$ centered about the z'-axis (Fig. 2). In presence of alignment, $I(z', d\Omega)$ is no more proportional to the number of atoms transferred in the excited state. It can be written as [13] :

$$I(z', d\Omega) = I_o \frac{d\Omega}{4\pi} \left[1 - (1 - \frac{3}{2} \sin^2 \theta) W_2 \right] . \qquad (1)$$

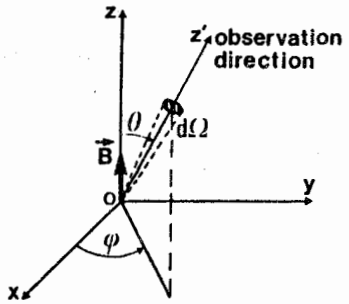


Fig. 2. Axis and angles definitions for fluorescence emission.

$I_0\, d\Omega/4\pi$ is the fluorescence yield collected in unit time when the emission is assumed isotropic. The second term into brackets is the anisotropic correction factor ; it is the product of two terms. The first one is a geometrical factor related to the polar angle $\theta = (\vec{B}, \vec{Oz})$ (Fig. 2). The second one W_2 is called the depolarization coefficient. It is usually introduced to account for the excitation anisotropy. W_2 is a function of the atomic parameters in the excited state, namely the populations and the longitudinal alignment. It is important to notice that relation (1) is only valid when Zeeman coherences are assumed negligible.

A general expression for W_2 relative to the emission from a multiplet F_e in the excited state is derived as [13] :

$$W_2 = \left[\frac{\sum_{M=-F_e}^{F_e} 3M^2 p_M}{\sum_{M=-F_e}^{F_e} p_M} - F_e(F_e+1) \right] \times K_e \qquad (2)$$

The term into brackets represents the alignment contribution due to population imbalance between the Zeeman sublevels in the excited state. It depends on excitation characteristics, namely the chosen transition, the laser polarization and intensity, by means of the population p_M of the Zeeman sublevel $|F_e, M\rangle$. K_e is a constant depending only on the desexcitation properties of the F_e multiplet state. Its value is tabulated in Table 1 for the dipolar electric transitions of cesium D_2 line at 852 nm.

F_e	K_e
5	-0.01111
4	$+7.14\ 10^{-3}$
3	+0.02778
2	-0.02381

Table 1. K_e values for each multiplet F_e of the cesium D_2-line.

The weak value of K_e can be compensated by an important population imbalance to give a non negligible depolarization coefficient and therefore, a noticeable anisotropy correction in Eq. (1).

It should be noted that $K_e \equiv 0$ for the cesium D_1 line at 894 nm. Consequently there is no anisotropy effect in the fluorescence signal for this excitation line which looks promising for optical pumping and detection in a cesium beam resonator. Indeed, in a single laser basic configuration, the $4 \rightarrow 3\ \pi$ scheme of the D_1 line has been predicted to give signal performances comparable to the ones obtained with the $3 \rightarrow 3\ \sigma$ scheme of the D_2 line [14].

IV. APPLICATION TO A PARTICULAR SITUATION

The situation corresponding to Fig. 1c is depicted on Fig. 3. Optical detection is achieved by two aspherical lenses which focus the light onto a large area silicon photodiode. This set-up collects fluorescence photons inside a solid angle Ω_0 defined by a $\theta_0 = 22.5°$ aperture angle about the quantization axis. A spherical mirror allows to double the collection yield.

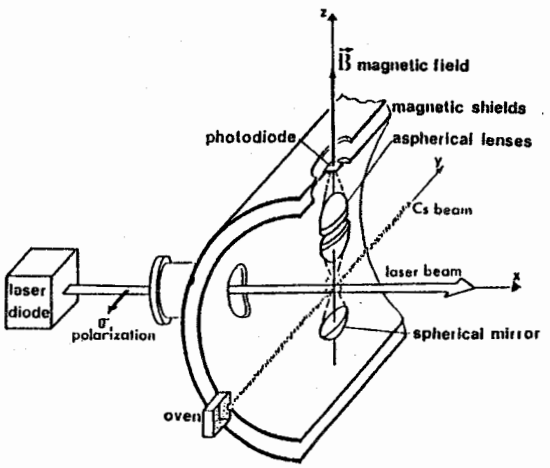


Fig. 3. Typical arrangement for the light-atom interaction region.

Integration of Eq. (1) over the angle φ between 0 and 2π and over the angle θ between 0 and θ_0 gives the fluorescence yield I_{θ_0} collected in unit time inside the solid angle $\Omega_0 = 2\pi\ (1 - \cos\theta_0)$:

$$I_{\theta_0} = I_0 \frac{2\Omega_0}{4\pi} \left[1 - \frac{W_2}{2} (\cos^2\theta_0 + \cos\theta_0) \right] \qquad (3)$$

The factor 2 takes into account the mirror contribution to I_{θ_0}.

As the light-atom interaction time is T, the total number of available photons is L_{θ_0} which is given by :

$$L_{\theta_0} = \int_0^T I_{\theta_0}\, dt \qquad (4)$$

L_{θ_0} has been numerically calculated for the six optical transitions of the cesium D_2 line under the following conditions : the broadband laser intensity is 3 mW cm^{-2} and the interaction time T = 20 µs. L_{θ_0} values are reported in Table 2 and compared to the fluorescence yield $L^{(o)}_{\theta_0}$ calculated when the photon emission is considered as isotropic.

excitation scheme	L_{θ_0} [ph/at]	$L^{(o)}_{\theta_0}$ [ph/at]
$3 \rightarrow 3\ \sigma$	0.144	0.133
$3 \rightarrow 3\ \pi$	0.076	0.083
$4 \rightarrow 3\ \sigma$	0.059	0.057
$4 \rightarrow 3\ \pi$	0.041	0.041
$4 \rightarrow 4\ \sigma$	0.106	0.102
$4 \rightarrow 4\ \pi$	0.077	0.077
$3 \rightarrow 4\ \sigma$	0.055	0.056
$3 \rightarrow 4\ \pi$	0.061	0.057
$3 \rightarrow 2\ \sigma$	1.023	1.070
$3 \rightarrow 2\ \pi$	0.082	0.075
$4 \rightarrow 5\ \sigma$	3.95	3.03
$4 \rightarrow 5\ \pi$	2.42	3.25

Table 2. Fluorescence yield for excitation schemes of the cesium D_2 line. L_{θ_0} and $L^{(o)}_{\theta_0}$ correspond to the anisotropic and isotropic cases respectively.

Table 2 shows quite clearly that the 4 → 5 and 3 → 3 transitions are strongly affected by anisotropy whereas the other excitation schemes do not give rise to significant changes in fluorescence. When the fluorescence emission is assumed isotropic the collection efficiency is the same for all the excitation schemes. Taking into account anisotropy features it may be seen from Table 2 that the collection efficiency depends on the involved transition and laser polarization. Compared to the isotropic case, its value increases by an amount of 31 % and 8 % when 4 → 5 σ and 3 → 3 σ schemes are respectively considered.

In order to compare theoretical predictions with the experiment, we have studied the variation of the fluorescence as a function of the laser polarization in the case of the 4 → 5 transition. This gives a good verification of the anisotropy contribution to the photon emission. Let us define δ , the angle between the laser polarization and the magnetic field. The predicted evolution of the fluorescence signal L_{θ_0} as a function of δ is reported on Fig. 4a for two θ_0 laser intensities. The dashed line represents $L_{\theta_0}(o)$. Fig. 4b shows the experimental variation of L_{θ_0} versus δ when the laser intensity is 1 mW cm^{-2} typically. The theoretical ratio $L_{\theta_0}(\delta = 90°)/ L_{\theta_0}(\delta = 0°)$ is found to be 1.5 and agrees quite well with the experimental one, 1.45. We have also verified that the measured collection efficiency is larger for the 4 → 5 σ scheme than for the 4 → 5 π one.

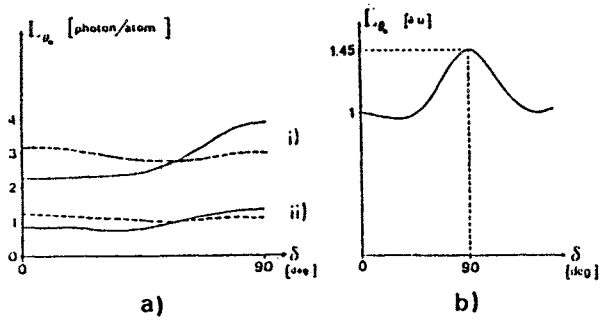


Fig. 4. Fluorescence yield for the 4 → 5 transition as a function of the laser polarization :
a) predicted evolution for two laser intensities
i) 3 mW cm^{-2}
ii) 1 mW cm^{-2}
The dashed line represents the isotropic case
b) Experimental variation for a 1 mW cm^{-2} laser intensity.

V. CONCLUSION

We have presented a study of the anisotropic character of the fluorescence of cesium atoms resulting from the creation of alignment in the excited state under the action of a linearly polarized laser light tuned to the D_2 line.

The modifications brought to the collected fluorescence depends strongly on the choice of the involved transition and laser polarization when the collection efficiency γ is low, for instance in a collection set-up with lenses. We have more specifically considered the situation where the collecting set-up has a cylindrical symmetry about the magnetic field axis. It is noteworthy that in this case, the (4 → 5 σ) and (3 → 3 σ) schemes take advantage of the anisotropy. This conclusion is of evident interest because the two mentioned excitation schemes appear as the most efficient for detection purposes in optically pumped cesium beam resonators operating in a transverse magnetic field. For example, in the basic experiment with a single laser [5,15], the (3 → 3 σ) scheme leads to the best short term frequency stability. In more sophisticated resonators, detection with the (4 → 5 σ) scheme has been proposed to improve the clock signal detection[3, 15]. The present study confirms the great interest of using these two transitions in resonators when the detection of the clock signal is performed along the direction of the magnetic field.

One additional point to recall is that anisotropy effects are smeared out when the collection efficiency reaches values larger than 30 %. This is the case where the collection optics is designed with two mirrors facing each other.

Acknowledgements

The authors are grateful to the Bureau National de Métrologie for its financial support.

1. J.L. Picqué, Metrologia 13 (1977), 115
2. T. Mc Clelland, I. Pascaru, J. Zacharski, N.H. Tran and M. Meirs, Proceedings of the 41st Annual Symposium on Frequency Control, Philadelphia, USA (1987) p. 59
3. S. Oshima, Y. Koga, Y. Nakadan, L. Hollberg and R. Drullinger, Proceedings of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988) p. 531
4. P. Thomann, H. Schweda and G. Busca, Proceedings of the 2nd European Frequency and Time Forum, Neuchâtel Switzerland (1988), p. 513
5. V. Candelier, V. Giordano, A. Hamel, G. Théobald, P. Cérez and C. Audoin, Proceedings of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988), p. 483
6. Wang Yi Qiu, Proceedings of the 4th Symposium on Frequency Standards and Metrology, Ancona, Italy (1988). To be published
7. E. de Clercq, A. Clairon, B. Dahmani and A.H. Gérard, Proceedings of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988), p. 499
8. M. de Labachelerie, K. Diomandé and N. Dimarcq, Proceedings of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988), p. 547
9. A. Hamel, P. Petit, G. Théobald, P. Cérez and C. Audoin, Proceedings of the 3rd European Frequency and Time Forum, Besançon, France (1989). This issue.
10. R.E. Drullinger, J. Shirley, D.J. Glaze and L.W. Hollberg, Proceedings of the 40th Annual Symposium on Frequency Control, Philadelphia, USA (1986), p. 428
11. S. Oshima, Y. Nakadan and Y. Koga, IEEE Trans., IM (1988). To be published
12. A. Omont, Progr. Quantum Electronics, 5 (1977), 69
13. G. Théobald, N. Dimarcq, V. Giordano, A. Hamel and P. Cérez, J. of Appl. Physics. Submitted for publication
14. G. Théobald et al., Optics Comm.Submitted for publication
15. G. Théobald, V.Giordano, M. de Labachelerie, A. Hamel, N. Dimarcq, P. Cérez and C. Audoin, Proceedings of the 42nd Annual Symposium on Frequency Control, Baltimore, USA (1988), p. 496

PROGRESS IN AN OPTICALLY PUMPED CESIUM BEAM RESONATOR OPERATING WITH A LONGITUDINAL MAGNETIC FIELD

A. HAMEL, P. PETIT, G. THEOBALD, P. CEREZ, C. AUDOIN

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

ABSTRACT

We have previously shown (1) that coherent population trapping may occur in the two optical interaction regions of the resonator at low static magnetic field. This effect, which induces a decrease of the clock signal amplitude, can be cancelled by the use of a larger magnetic field (3×10^{-5} Tesla) in the optical interaction regions. We have thus designed a resonator having 3 magnetic field regions in which the field structure is longitudinal for sake of easier realization. We present this resonator and the preliminary results obtained from such a machine.

INTRODUCTION

In order to improve the long term frequency performances of atomic beam frequency standards, the microwave interrogation of the atoms must take place in a low magnetic field amplitude region. Actually, the frequency fluctuation df related to a given C-field fluctuation dB is proportional to the C-field value B. We have :

$$\frac{df}{dB} = 855 \times 10^8 B_o ,$$

where the frequency is expressed in Hertz and the magnetic induction in Tesla.

However, in conventional Cs beam tubes using magnet state selectors, the asymmetrical population of Zeeman sublevels ($m_F \neq 0$) forces to set the C-field to a typical value of 6×10^{-6} T in order to minimize the Rabi-pulling effect.

In optically pumped Cs beam resonators, optical pumping performed with strictly linearly polarized light leads to a symmetrical microwave spectrum, which thus eliminates the Rabi pulling frequency shift. In such a machine it is possible to reduce drastically the C-field amplitude.

Unfortunately, drawback arises due to the non linear Hanle effect (1). It reduces the efficiency of optical pumping in weak magnetic fields so we had to design a new cesium beam tube with three magnetic field regions.

In Section I, we recall the consequences of non linear Hanle effect on the clock signal features and suggest a proper magnetic field structure for an efficient optically pumped Cs beam resonator.

In Section II, the experimental set-up Cs III is described. Finally, results and short term stability obtained from this new machine are given in Section III.

I. CONSEQUENCES OF THE HANLE EFFECT ON THE DESIGN OF A CESIUM BEAM TUBE

When applying optical pumping techniques to the Cs beam frequency standard, we are mainly interested in two physical quantities which determine the clock signal features i.e. : its amplitude and the background noise. Those quantities are :

- the fractional population difference between the $m_F = 0$ ground state hyperfine sublevels obtained after the optical preparation :

$$\Delta n_{oo} = \frac{n(|\, F = 4,\ m_F = 0\,\rangle) - n(|\, F = 3,\ m_F = 0\,\rangle)}{n},$$

where n is the total population of cesium atoms.

- the number L of fluorescence photons emitted by the cesium beam under the action of the laser. These fluorescence photons are collected in the detection region to provide the clock signal.

It has been shown (2) that the transition $3 \rightarrow 3$, σ polarized light is the most interesting one. It maximizes the product $\Delta n_{oo} \times L$. But, when operating in weak magnetic field i.e. $B < 3 \times 10^{-5}$ T (300 mG), successive absorption and stimulated emission of photons create $\Delta m_F = \pm 2$ Zeeman coherences in the ground state of Cs atoms. Since $\Delta m_F = \pm 2$ coherences are coupled by the laser to the population of the excited state, it results in a decrease of Δn_{oo} and L when the C-field value goes to zero (3). Consequently, the peak to valley amplitude of the clock signal decreases, whereas the continuous background noise due to the unpumped atoms increases. Figures 1a and 1b report predicted Δn_{oo} and L values versus static magnetic induction amplitude B for the 4 Cs pumping transitions. The atom-light interaction time and laser intensity were assumed equal to 12 µs and 2 mW cm^{-2}, respectively.

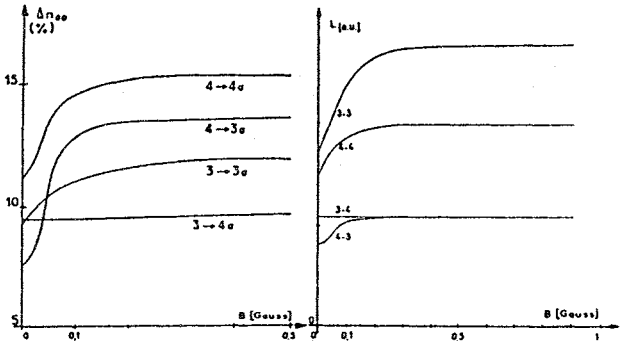


Fig. 1. Population difference (a) and fluorescence yield (b) versus the amplitude of magnetic induction.

One can see that for the $3 \rightarrow 3$ σ transition, the value of B should be kept above 3×10^{-5} T in order to retain full pumping and fluorescence efficiency.

To take advantage of the best of the optical pumping technique, i.e. large signal to noise ratio and small C-field value, we have designed a new experimental Cs beam resonator with the following magnetic structure.

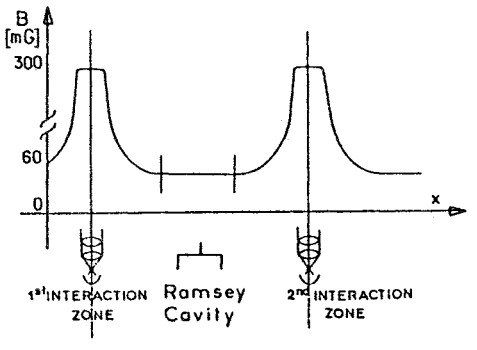


Fig. 2. Magnetic field profile along the atomic beam axis.

II. EXPERIMENTAL SET-UP

Beside the original magnetic field configuration, other features have been improved : the cavity feeding and the light collecting device.

a) Magnetic field

In this machine the magnetic field is directed along the atomic beam. The advantages of such a configuration are the following :

- easier practical realization
- as seen below, it leads to a narrower Rabi pedestal in the microwave spectrum.

The C-field B_0 in the RF interaction region is generated by a solenoïd which surrounds the Ramsey cavity. The stronger magnetic field needed in the optical pumping and detection regions is produced by sets of twin flat coils.

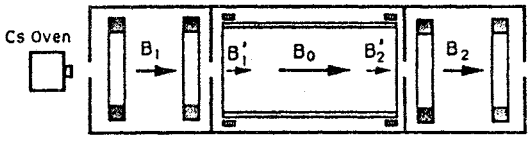


Fig. 3. Magnetic field generation in Cs III.

Each region is magnetically insulated from the other by a mu-metal partition wall. Furthermore, to improve the uniformity of the magnetic induction in the RF interaction region, two compensation coils are placed at both ends of the solenoïd. They generate two weak magnetic inductions B'_1 and B'_2 that cancel the influence of the stronger field in the adjacent regions. Figure 4 shows an example of static magnetic induction pattern along the atomic beam path.

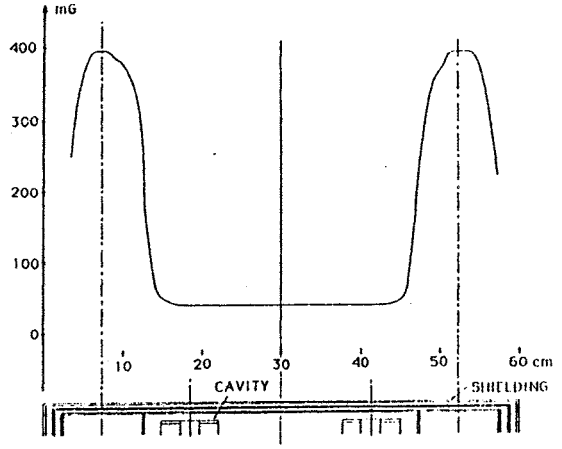


Fig. 4. Experimental magnetic induction profile in Cs III.

b) Microwave cavity

Our cavity shown in figure 5 is home-machined and is made of hard aluminium alloy 6061. This solution has been retained in order to get a better dimensional precision on the arms length. Due to the longitudinal magnetic field, the cavity is bent in the H-plane. It operates on the TE 0, 1, 13 mode. The dimensions have been choosen to remain compatible with an industrial realization. The inner drift distance is 21 cm ; this yields a line width of 500 Hz.

In order to eliminate the usual hole in the magnetic shielding and to avoid a field perturbation along the beam path, we have developed a coaxial feeding of the cavity. For the same purpose, the coupling arm is enclosed in the evacuated tube placed on the side of the main U-shaped wave guide to avoid stopping the beam. Great care has been taken to accurately tune the cavity since once it is in the tube no further frequency adjustement can be made. The loaded Q of our cavity is about 800 which requires an increased excitation power but minimizes the cavity pulling effect.

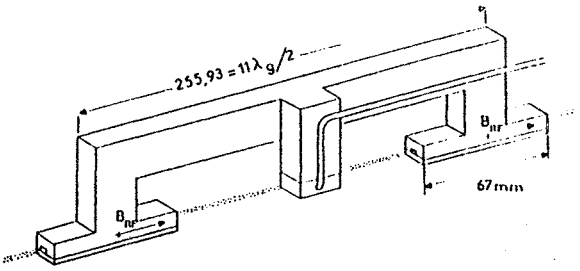


Fig. 5. Ramsey cavity design.

c) The optical collecting device

Our photon collecting system is made of two aluminium mirrors which surround the light-interaction region. The geometric collection efficiency has been measured and reaches 48 % for a 100 mm^2 detector area. A set of diaphragms has been placed along the laser beam path in order to minimize the stray light. This stray light comes mainly from the laser incident power. It induces a continuous background and a spurious which decreases the detection signal to noise ratio. It is mainly due to the poor geometric quality of the laser light beam delivered by conventional laser diodes, and to scattered light from the windows. We have also designed an amplification electronic device which displays a dark current noise of 2×10^{-14} A x $Hz^{-1/2}$. It converts the photon flux into a proportional voltage.

III. EXPERIMENTAL RESULTS

We have investigated the properties of our new atomic cesium beam resonator Cs III. We shall present here the most distinctive results. In the following experiments, a single laser diode is used for optical pumping and optical detection. It is tuned to the $3 \rightarrow 3\sigma$ transition.

a) Pumping efficiency

With a three independent magnetic field regions Cs III is well suited to investigate the beam preparation efficiency versus the magnetic field amplitude. We have recorded the variation of the fluorescence signal $\mathcal{F}$ in the detection region versus B_1, i.e. magnetic induction in the pumping region. This signal comes from unpumped atoms in this region. Let $\mathcal{F}_0$ the fluorescence of the beam when no pumping occurs in the preparation region. Figure 6 shows the ratio $\mathcal{F}/\mathcal{F}_0$, which accounts directly for the percentage of unpumped atoms. We see that when B_1 is very small, $\mathcal{F}/\mathcal{F}_0$ is about 8 % and that complete preparation is only achieved when B_1 is increased above 3×10^{-5} T (300 mG). In this experiment B_0 and B_2 are set respectively to 4×10^{-6} T and 5×10^{-5} T.

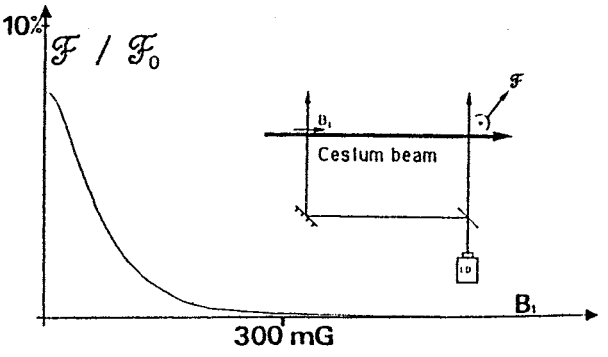


Fig. 6. Percentage of unpumped atoms versus B_1.

Furthermore the uncomplete atomic beam preparation induces a dramatic enhancement of the noise level with the fluorescence signal $\mathcal{F}$, as can be seen on figure 7. This effect can be explained by noting that when the pumping is no longer complete, the fluorescence yield of the atomic beam depends on the laser frequency fluctuations.

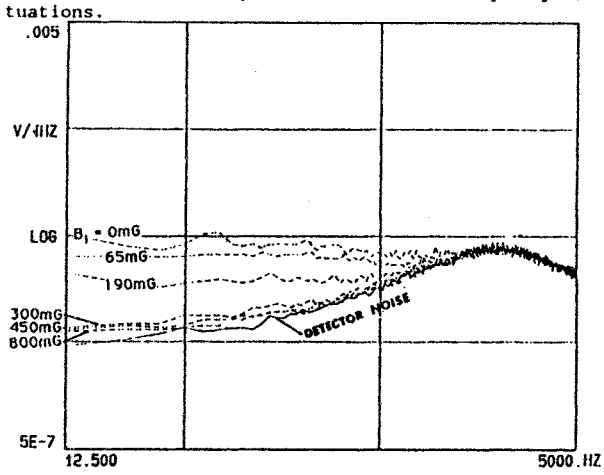


Fig. 7. Photocurrent fluctuations spectrum for several B_1 values.

b) Microwave spectrum

In the following results optical pumping and detection are performed within a magnetic induction of 5×10^{-5} T (500 mG) to achieve full efficiency of the two processes. Figure 8 shows the microwave spectrum of our Cs beam resonator Cs III when the value of the magnetic induction B_o is set to 4×10^{-6} T (40 mG). We note first that the Ramsey patterns of all the frequency dependent transitions are very well resolved. Secondly, they are well centered on their Rabi pedestal. This accounts directly for the magnetic field uniformity along the beam path in the whole RF interaction region. This has been made possible by adjusting the field amplitude with the compensation coils.

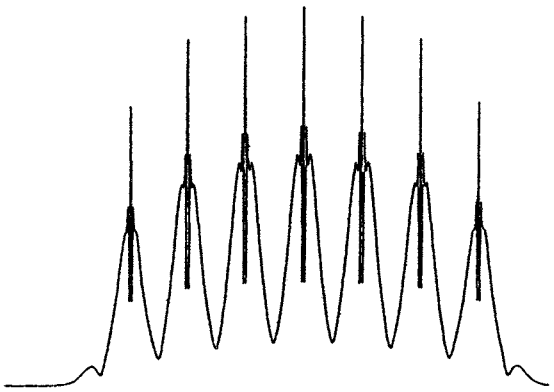


Fig. 8. Microwave spectrum B_o : 4×10^{-6} T (40 mG).

c) Central Ramsey pattern

Figure 9 shows the Ramsey pattern obtained for the $F = 4$, $m_F = 0$, $F = 3$, $m_F = 0$ transition. The central fringe is 500 Hz wide. I_b represents the background of stray light, which is very small. The amplitude signal to noise ratio is defined as the peak to valley current difference divided by the standard deviation of the peak current. It reaches 11.000 in a 1 Hz noise bandwidth.

We have obtained this signal to noise ratio with an oven temperature of about 115° C. The effective oven emissive area is about 5 mm^2. It is of interest to note that the Rabi pedestal is narrower than with conventional E-plane cavities. It accounts for the larger interaction time in each arm. Furthermore, the smoother variation of the microwave field along the beam path reduces the amplitude of the aisles of the Rabi pedestal, which is favorable to a decrease of the Rabi pulling, if the latter exists.

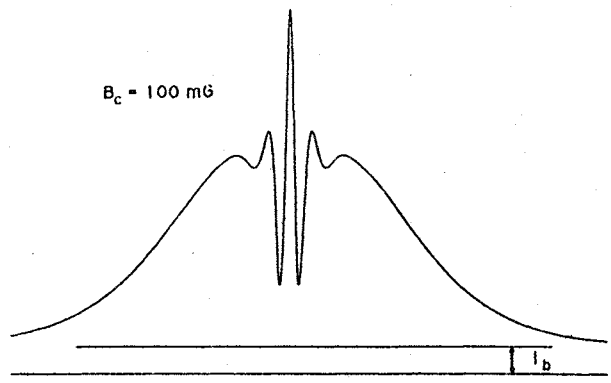


Fig. 9. $|F = 4, m_F = 0\rangle \rightarrow |F = 3, m_F = 0\rangle$ resonance pattern.

d) Very low magnetic induction

Thanks to the three separate fields configuration, it is possible to set the longitudinal magnetic induction in the RF interaction region to very small values, while keeping the magnitude of the induction at 5×10^{-5} T (500 mG) in the optical interaction regions. The two compensation coils are used to correct the field near each end of the solenoïd. As a results, the magnetic induction B_o can have a mean value adjusted closely equal to zero, while its deviation can be made very small. In such conditions there is no Zeeman splitting any longer. All the atoms of the beam undergo the hyperfine transition. As shown on figure 10, we thus see that the obtained Ramsey pattern is about 6 times higher with a slightly increased width. The usual fringe is displayed for comparison purpose.

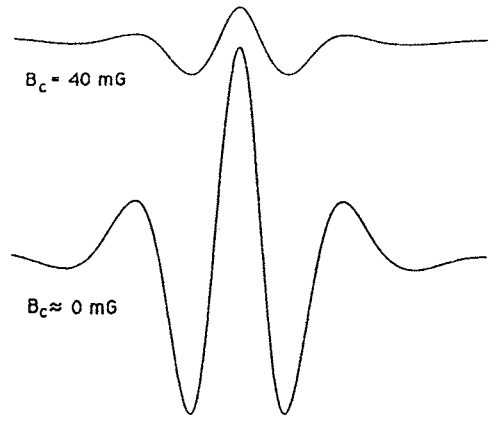


Fig. 10. Ramsey pattern at 4×10^{-6} T (a) and at very low magnetic induction (b).

e) Stability measurements

The frequency stability of the frequency controlled 10 MHz quartz oscillator has been measured with respect to a hydrogen maser. The result is $\sigma_y(\tau) = 2.1 \times 10^{-12}\ \tau^{-1/2}$ for 1 s $< \tau <$ 50 s. It is similar to the frequency stability obtained in the Cs II device [4], but the presently reported result has been achieved with a beam intensity approximately 4 times smaller.

CONCLUSION

We have designed a new optically pumped cesium beam resonator which allows to optimize the three interaction processes. We have confirmed experimentally, with the simplest possible set-up using a single laser, that optical pumping methods can improve the short term frequency stability of Cs beam frequency standards.

Furthermore, we have demonstrated that they can be efficiently implemented in the presence of a weak magnetic field in the RF interaction region, a condition favorable to long term frequency stability.

Acknowledgements

We are grateful to STEI and BNM for their financial support. We are indebted to V. Candelier for his help during frequency stability measurements.

REFERENCES

(1) G. Théobald, P. Cérez, N. Dimarcq, V. Giordano, Proceedings of the IVth Symposium on Metrology and Frequency Standards, Ancona, Italy (1988). To be published

(2) G. Avila, V. Giordano, V. Candelier, E. de Clercq, G. Théobald, P. Cérez, Phys. Rev. A36, 3719 (1987)

(3) G. Théobald, N. Dimarcq, V. Giordano, P. Cérez, Optics Comm. Accepted for publication

(4) V. Candelier, V. Giordano, A. Hamel, G. Théobald, P. Cérez, C. Audoin, Proceedings of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (March 1988), p. 483

OPTICAL PUMPING OF A CESIUM BEAM AND POPULATION TRAPPING

E. DE CLERCQ AND P. MANGIN
Laboratoire Primaire du Temps et des Fréquences
61 avenue de l'Observatoire - 75014 Paris - France

I - INTRODUCTION

In a cesium beam frequency standard, it is very important to get a large difference of population between the two clock-levels in order to obtain after the microwave resonance a large signal and consequently a good short-term stability. Optical pumping with the different possible transitions allows us numerous ways to create this population difference. This has been already studied both theoretically and experimentally [1] for broadband lasers and a sufficiently high magnetic field. However it has been shown that in other cases coherence effects can occur preventing a perfect pumping [2, 3, 4]. So, we have undertaken a numerical calculation of the atomic evolution during optical pumping process with monochromatic lasers and low magnetic field. Here we describe first our theoretical model, then we give some results for the one laser pumping scheme and finally for the two lasers pumping scheme.

II - THEORETICAL MODEL

1) Theoretical Model

We consider an atomic beam propagating along Ox. The atoms are irradiated by one or two lasers beams of pulsation ω_1 and ω_2 propagating along Oy and having a linear polarization $\mathbf{e_1}$ and $\mathbf{e_2}$. A static magnetic field B is applied in the Oz direction (See Fig. 1). Each polarization vector $\mathbf{e_i}$ (i = 1 or 2) make an angle θ_i with Oz. If $\theta_i = 0°$ the polarization is called π polarization, if $\theta_i = 90°$ it is called σ. In this work circular polarizations are not considered. Of course, our results are also valid if the magnetic field B is applied along the atomic beam since, in fact, our model do not take into account the relative directions of the atomic beam and of the magnetic field.

Fig. 2 shows the first energy levels of the cesium atom. Each laser can be tuned on a transition between a hyperfine level of the ground state and a hyperfine level of the $6^2P_{3/2}$ state (D2 line) or a hyperfine level of the $6^2P_{1/2}$ state (D1 line). The Larmor frequency is indicated beside each level.

We use a semi-classical model in which the laser fields are classically written and the atomic system is quantum mechanically treated. We use the density matrix formalism [5]. Let ρ the density matrix representing the

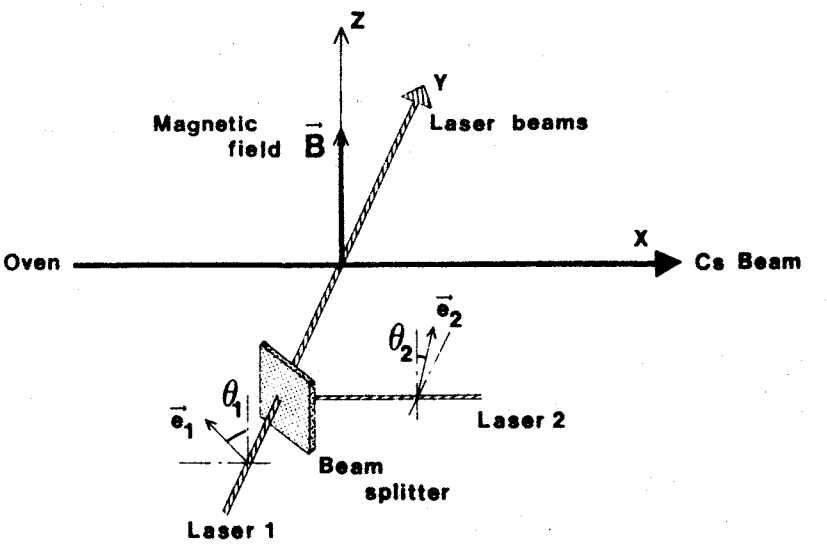


Fig. 1. Experimental geometry.

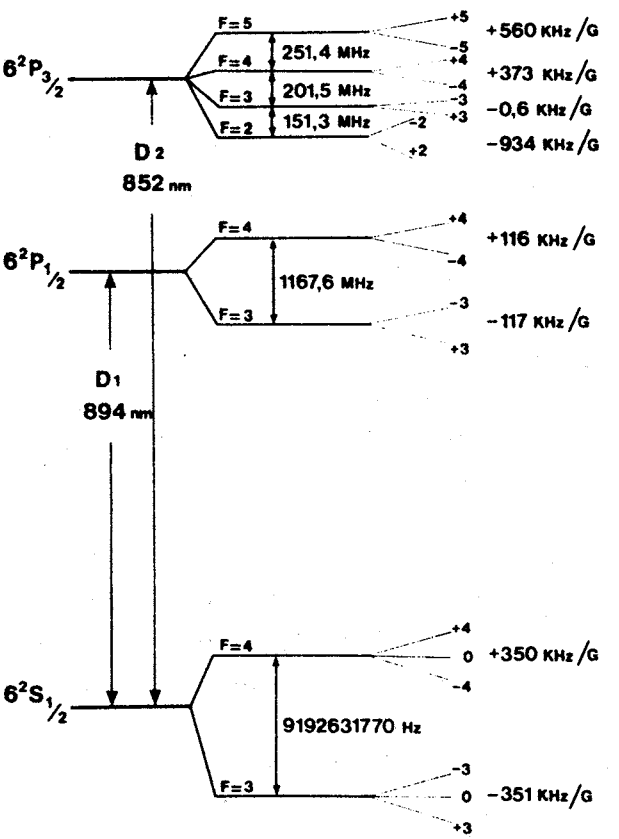


Fig. 2. First energy levels of cesium atom.

ensemble of atoms in the illuminated zone of the atomic beam. The time evolution of ρ is given by the Liouville equation :

$$\frac{d}{dt}\rho = 1/i\hbar\,[H, \rho] + R \qquad (1)$$

where we have added the R term to describe the effect of spontaneous emission. H is the total hamiltonian of the system, we can write :

$$H = Ho + Hz + V \qquad (2)$$

where Ho is the unperturbed hamiltonian, Hz is the Zeeman hamiltonian and V is the interaction term with the electromagnetic field

$$Hz = -\mu.\mathbf{B} \qquad (3)$$

where μ is the magnetic dipole moment of the atom. The interaction term is :

$$V = -\mathbf{D}.\mathbf{E} \qquad (4)$$

with **D** the electric dipole moment and **E** the total electric field : $\mathbf{E} = \mathbf{E_1} + \mathbf{E_2}$. Laser beams are described as monochromatic plane waves, then :

$$\mathbf{E}_i(t) = 1/2\ Ei(e^{i\omega_i t}\,\mathbf{e}_i + e^{-i\omega_i t}\,\mathbf{e}^*_i) \qquad (5)$$

For each density matrix element ρ_{ij}, equation (1) yields :

$$d/dt\ \rho_{ij} = -i(\omega_{ij} + \omega_{zij})\rho_{ij} + 1/i\hbar \sum_k (V_{ik}\,\rho_{kj} - \rho_{ik}\,V_{kj}) + R_{ij} \qquad (6)$$

with $\omega_{ij} = (E_i - E_j)/\hbar$, E_i is the energy of level i ; $\omega_{zij} = \omega_{zi} - \omega_{zj}$ and ω_{zi} is the Larmor pulsation for the level i. If we put $\rho_{ij} = \sigma_{ij}\, e^{-i\omega_{ij}t}$ and using the rotating wave approximation, we obtain :

$$d/dt\ \sigma_{ij} = -i\omega_{zij}\,\sigma_{ij}$$

$$+ i/2\hbar \sum_k \{\, [E_1(\mathbf{D}.\mathbf{e}_1)_{ik}\, e^{i\Delta\omega_{ik,1}t} + E_2(\mathbf{D}.\mathbf{e}_2)_{ik}\, e^{i\Delta\omega_{ik,2}t}]\,\sigma_{kj}$$

$$- \sigma_{ik}[E_1(\mathbf{D}.\mathbf{e}_1)_{kj}\, e^{i\Delta\omega_{kj,1}t} + E_2(\mathbf{D}.\mathbf{e}_2)_{kj}\, e^{i\Delta\omega_{kj,2}t}]\,\}$$

$$+ R_{ij} \qquad (7)$$

where $\Delta\omega_{ik,1} = \omega_{ik} - \omega_1$ if $\omega_{ik} > 0$ and $\omega_{ik} + \omega_1$ if $\omega_{ik} < 0$, the same for $\Delta\omega_{ik,2}$

If i and j are two levels of the excited states

$$R_{ij} = -1/\tau\ \sigma_{ij} \qquad (8)$$

τ is the lifetime of the excited state, τ = 32 ns. If i is a level of the excited states and j a level of the ground state, or the reverse, (σ_{ij} is an optical coherence) : $R_{ij} = -1/2\tau\ \sigma_{ij}$ (9)

If i and j belong to the ground state :

$$R_{ij} = 1/\tau \sum_{\substack{e,e' \\ m_e - m_{e'} = m_i - m_j}} a_{ei}\, a_{e'j}\, \sigma_{ee'}\, e^{i(\omega_{ie} - \omega_{je'})t} \qquad (10)$$

where e and e' are excited levels and the coefficients $a_{ei}(a_{e'j})$ are linked to Clebsch-Gordan coefficients and are defined in [2].

If we assume that initially the 16 Zeeman sub-levels of the ground state are equally populated and that all the other matrix elements are null, so we can calculate the evolution of the density matrix.

2) Numerical Calculation

48 Zeeman sub-levels are involved in the D2 line which lead to a density matrix of dimension 48 x 48 or a set of 2304 coupled linear differential equations. Using hermiticy of the density matrix we can reduce this number to 1176. We use then a Runge-Kutta method to compute the time evolution of the matrix elements. The calculation is performed with a variable step Δt. If the off-resonance hyperfine levels are taken into account in the calculation oscillating terms appear in the equations and we have to use a very little Δt step leading to a very long computation time. Then we have not taken into account off-resonance levels in our calculation. This approximation is justified when the number of absorbed photons is not too long. With these conditions, the calculation of the density matrix evolution during 5 μs needs about 1 hour of CPU time on a Micro-Vax II computer.

III - ONE LASER INTERACTION

We have calculated the population of each magnetic sublevel as a function of time. This has been done for all possible transitions with σ or π polarization. In an atomic clock we are especially interested by the difference of population between the two clock-levels : (F = 4, m = 0) and (F = 3, m = 0) of the ground state. We define the fractional population difference Δn as :

$$\Delta n = \left| \frac{n(F = 4,\ m = 0) - n(F = 3,\ m = 0)}{n} \right| \qquad (11)$$

where n is the total population (= 1 in our case). We have calculated Δn after an interaction time T equal to 5 μs, with a laser intensity I = 3 mw/cm^2 and for three values of the magnetic field : B ~ to 0, B = 0.1 G and B = 1G. The results are shown in Table 1 for one laser tuned to a transition of the D_1 line and in Table 2 for one laser tuned to a transition of the D_2 line.

Table 1 : Population difference Δn obtained for one laser pumping with the D_1 line
Δn in percent, T = 5 μs, I = 3 mw/cm^2, the symbol - means unchanged.

Transition	B ~ oG	B = 0,1G	B = 1G
3-4 π	14,4	-	-
4-4 σ	12,3	14,5	15,6
3-3 σ	10,6	11,4	11,9
4-3 π	10,0	-	-
3-4 σ	9,6	-	9,7
4-3 σ	1,8	2,9	16,7
3-3 π	0	-	-
4-4 π	0	-	-

We can see on table I and II that : (i) for σ transitions Δn changes with increasing magnetic field, (ii) for a sufficiently high magnetic field our results are in good agreement with a standard rate equations model [1].

The behaviour in weak magnetic field can be explained by Zeeman coherences effects [5, 6]. In a weak magnetic field, a first interaction with the σ polarized laser removes the atom from a ground state IF, m_F > (absorption process) and puts it in a coherent superposition of m_F - 1 and m_F + 1 sublevels of the excited level. A second interaction with the laser (stimulated emission process) brings back the atom to the ground state in a coherent superposition of m_F - 2, m_F and m_F + 2 states creating so $\Delta m = \pm 2$ and $\Delta m = \pm 4$ coherences. After a few cycles, we obtain coherences between Zeeman sublevels with $|\Delta m| = 2, 4, 6$ or when it is possible 8. This prevents further absorption processes and the atomic system becomes transparent. The Zeeman coherences decrease with increasing values of B due to the Larmor precession and the detuning of the components of the optical line with respect to the laser frequency. The populations of the Zeeman sublevels as the coherences will exhibit a resonant behaviour near B = 0. This effect is known as stimulated level crossing [7], saturation resonance [5, 6] or non-linear Hanle effect [8].

The histograms of Figure 3 shows the populations of the Zeeman sublevels of the ground state in two cases : after a 4-4 σ pumping and after a 3-3 σ pumping. In these two cases the effect of coherences is clearly visible on the populations distribution.

Table 2 : Population difference Δn obtained for one laser pumping with the D_2 line.
Δn in percent, T = 5 μs, I = 3 mw/cm², the symbol - means unchanged. Results of ref [1] are calculated with a different model

Transition	B ~ oG	B = 0,1G	B = 1G	Ref [1]**
3-4 π	13,7	-	-	13,7
4-3 π	12,1	-	-	12,1
4-4 σ	11,5	14,3	15,6	15,5
3-3 σ	9,8	10,7	12,6	12,2
3-4 σ	9,7	-	10,1	9,9
3-3 π	8,8	-	-	8,5
4-3 σ	7,8	10,6	15,6	15,4
4-5 π*	7,6	-	-	
3-2 σ*	7,4	-	2,7	
3-2 π	6,3	-	-	
4-5 σ*	3,4	3,5	5,3	
4-4 π	2,9	-	-	2,5

* For these transitions, there is still an important fraction of the population in the excited level at t = 5 μs

* * Results of ref. 1 are calculated with slightly different parameters (T=3 μs, I = 10 mw/cm²). It is not important if the steady-state is reached.

Figure 4a shows the shape of the resonance of Δn for different interaction times for a 4-4 σ pumping. As soon as B is not null the steady state is reached when the F = 4 level is empty. Lower is the magnetic field and longer is the time needed to reach this state. So the resonance width decreases with increasing interaction time. Figure 4b shows the same resonance with different laser intensities. The decrease at high intensity can be explained by stimulated emission effects.

Of course, the same resonances can be observed on the fluorescence light as in ref [3].

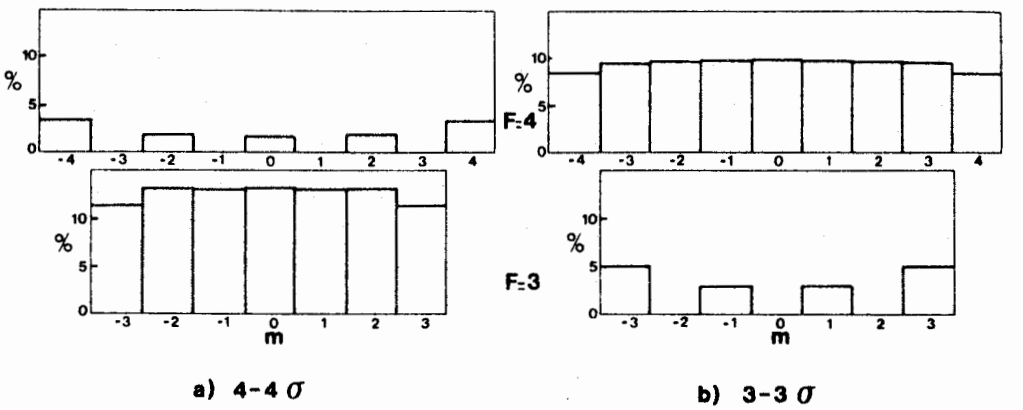


Fig. 3. Population of Zeeman sublevels in zero magnetic field at steady state, a) 4-4 σ pumping, b) 3-3 σ pumping. The two transitions are D_2 line transitions, T = 5 μs, I = 3mW/cm^2.

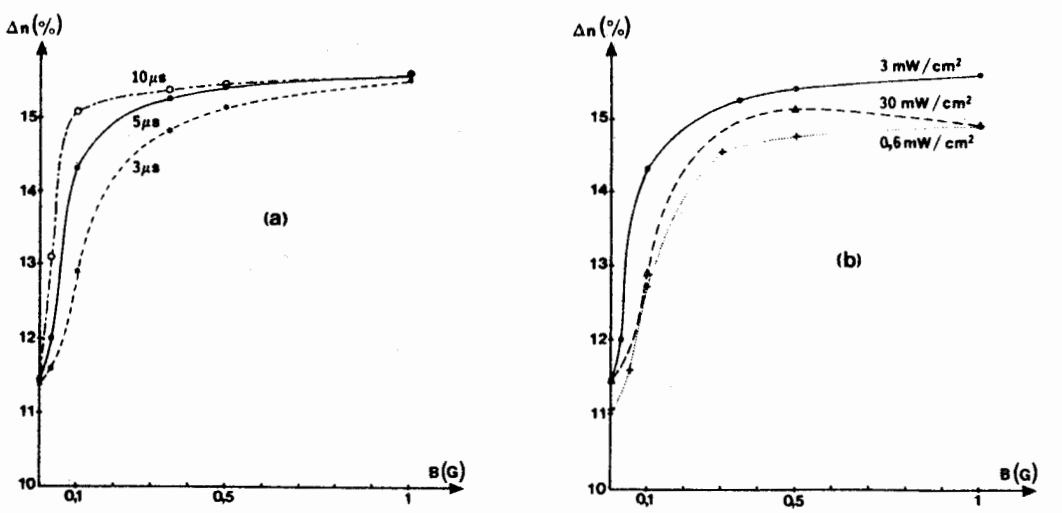


Fig. 4. Fractional population difference versus the magnetic field. One laser pumping with the 4-4 σ transition of the D_2 line : a) for different interaction times, I = 3mW/cm^2 ; b) for different laser intensities, T = 5 μs.

IV - TWO-LASERS INTERACTION

Schemes using two lasers in the optical pumping region should transfer all the atoms on a single clock-level [9]. In these schemes a laser emptys a ground-state hyperfine level and the other laser, which must be tuned to a F-F π polarized transition, depletes all the sublevels of the other hyperfine level exept the F, $m_F = 0$ sublevel. So after a few cycles all the atoms should be trapped in a single F, $m_F = 0$ sublevel. However, in a low magnetic field if one laser is σ polarized, it will induce Non-Linear Hanle Effect as seen above preventing a perfect pumping. We can use also two π polarized lasers since π transitions do not create Zeeman coherences. In Cesium there are only two possibilities : (i) a laser is tuned on the 3-3 π transition and the other on the 4-4 π transitions. The two F = 3, m = 0 and F = 4, m = 0 sublevels are then trapping levels and the population difference is far from 100%. (ii) the two transitions share the same excited level : 3-4 π and 4-4 π or 3-3 π and 4-3 π. Unfortunately in these schemes the atom is put in a coherent superposition of F = 4, m_F and F = 3, m_F ground state sublevels (hyperfine coherence) which prevents further absorption processes [10] and the population difference remains rather low [2].

Table 3 shows the population differences obtained with an optical pumping using two lasers tuned on D_2 line transitions. The results are given for three values of the magnetic field. Table 4 is the same as Table 3 for the D_1 line.

Table 3 : Population difference Δn obtained for two laser pumping. The lasers are tuned on the D_2 line. Δn in percent, T = 20 μs, I = 3 mW/cm^2 for each laser, the symbol - means unchanged. The transitions are listed by order of efficiency in null magnetic field.

Transition	B ~ oG	B = 0,1G	B = 1G
3-3 π , 4-4 σ	57,4	97	100
3-3 π , 4-4 π	38,2	–	–
3-3 σ , 4-4 π	37,3	77	98
3-4 π , 4-4 π	14,3	21	97
3-3 π , 4-3 π	14,2	20	47
3-4 σ , 4-4 π	11,3	20	81
3-3 π , 4-3 σ	8,1	22	100

Table 4 : Population difference Δn obtained for two laser pumping. The lasers are tuned on the D_1 line. Same conditions and same remarks as in table 3.

Transition	B ~ 0	B = 1G
3-3 σ , 4-4 π	53,2	96
3-3 π , 4-4 σ	47,6	97
3-4 π , 4-4 π	14,5	88
3-4 σ , 4-4 π	10,8	66
3-3 π , 4-3 π	9,8	18
3-3 π , 4-3 σ	1,7	79
3-3 π , 4-4 π	0	0

As for the one laser pumping case, the coherences effect decreases with increasing values of the magnetic field due to the different values taken by the Larmor frequencies. This is true also for hyperfine coherences as we can see with the 3-4 π, 4-4 π transition. In this scheme we get no Zeeman coherences because the two polarizations are π, so it is a very good example to study the hyperfine coherences behaviour. The fractional population difference versus one laser detuning in a zero magnetic field is shown in Figure 5 for two pumping schemes : 3-4 π, 4-4 π and 3-3 π, 4-4 σ for comparison purpose. In the first case as soon as a laser is detuned the excitation of hyperfine coherences is no more resonant and Δn increases rapidly. In the second case, on the contrary Δn decreases with increasing detuning because Δn is limited by Zeeman coherences and no by hyperfine coherences and the laser detuning has no effect on Zeeman coherences. Of course, Δn decreases when the detuning becomes larger than the natural width.

Another way to avoid creation of hyperfine coherences is the use of one laser tuned to the D_2 line while the other laser is tuned to the D_1 line. Figure 6 shows the time evolution of the fractional population difference Δn for different pumping schemes : (1) the most efficient schemes using two lasers tuned on the same line, 3-3 π, 4-4 σ with the D_2 line and 3-3 σ, 4-4 π with the D_1 line. Δn does not reach 60%. (2) the 3-4 π, 4-4 π scheme with the two lasers tuned to the D_2 line or to the D_1 line. Δn reaches 14%. (3) the same scheme 3-4 π, 4-4 π but with one laser tuned to the D_2 line and the other tuned to the D_1 line. Δn reaches 100%, all the atoms are in a single clock-

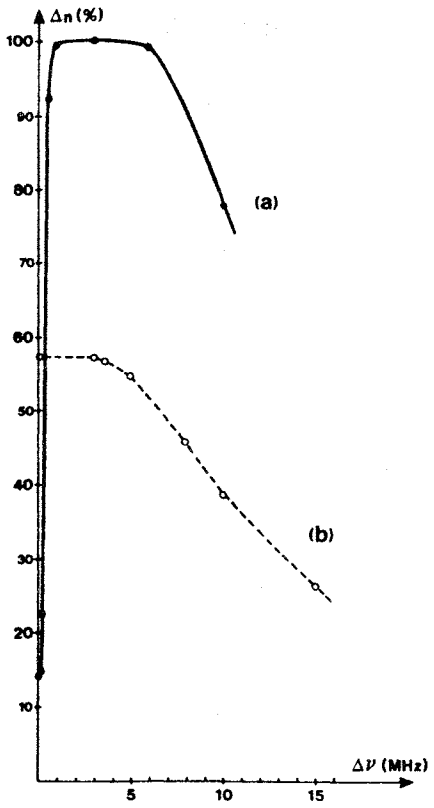


Fig.5. Fractional population difference Δn versus one laser detuning in zero magnetic field. The two lasers are tuned to the D_2 line :
a) 3-4 π, 4-4 π scheme ; the 3-4 π laser is detuned ;
b) 3-3 π, 4-4 σ scheme, the 3-3π laser is detuned. T = 20 μs, I = 3 mW/cm^2.

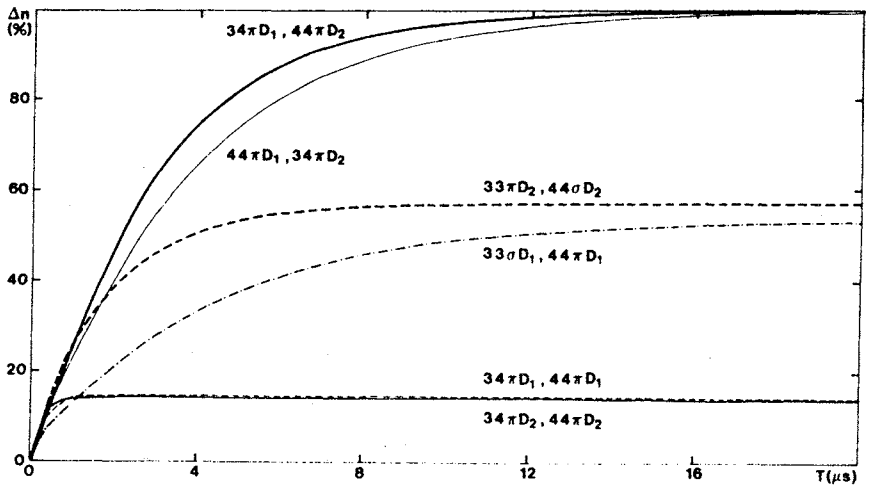


Fig. 6. Time evolution of the fractional population difference Δn for a few two-lasers pumping schemes in a zero magnetic field. (I = 3 mW/cm^2 for each laser).

level. These two last schemes are the most efficient in zero magnetic field and are the only-ones preventing coherences creation.

V - CONCLUSION

We have shown that optical pumping process can be partially inhibited in weak magnetic field by coherences effect : Zeeman coherences (Non-Linear Hanle Effect) with σ polarized transitions or hyperfine coherences (black lines [10]) with two π polarized transition. In the one laser scheme the Zeeman coherences build-up can be avoided by use of a strong magnetic field (B > 0,5G) in the optical interaction region or by using a π polarized transition although the population difference obtained is not the largest possible. In order to achieve a complete pumping with a two-lasers scheme we can also use a strong magnetic field or two π polarized transitions with two broad-band uncorrelated lasers [2]. However in this case the noise and the pumping time are greatly enhanced. Another interesting possibility is to detune one laser frequency. The simplest way use the Doppler effect due to divergence of the atomic beam and of the laser beams with two counter propagating laser beams. Finally, the best way to achieve a complete pumping in weak magnetic field seems to make use of two π polarized beams resonant with the D_1 and D_2 lines. It is worth noting that if Zeeman or hyperfine coherences are created by optical pumping they could modify the microwave interaction or the optical interaction in the detection region. This point must be carefully investigated in order to prevent new error sources in an optically pumped cesium clock.

Acknowledgments

We wish to thank Pascal Blondé for his helpful assistance in use and management of the computer , and Catherine Boulland for help in preparing this manuscript.

REFERENCES

[1] G. Avila, V. Giordano, V. Candelier, E. de Clercq, G. Theobald and P. Cerez, Phys. Rev. A36, 3719 (1987).

[2] E. de Clercq, M. de Labachelerie, G. Avila, P. Cerez and M. Tetu, J. Phys. (Paris) 45, 239 (1984).

[3] G. Theobald, P. Cerez, N. Dimarcq, V. Giordano, Fourth Symposium on Frequency Standard and Metrology, Ancona, September 1988.

[4] C. Jacques and P. Tremblay, Proceedings of the 42nd Annual Frequency Control Symposium, Baltimore, 1988.
P. Tremblay and C. Jacques, Fourth Symposium on Frequency Standard and Metrology, Ancona, 4-8 September 1988.

[5] C. Cohen-Tannoudji in Frontiers in Laser Spectroscopy, Les Houches 1975, edited by R. Balian, S. Haroche and S. Liberman (North Holland, Amsterdam).

[6] C. Cohen-Tannoudji in Atomic Physic 4, edited by G. Zu Putlitz, E.W. Weber and A. Winnacker, Plenum (New York 1975), p. 589.

[7] M.S. Feld, A. Sanchez, A. Javan and B.J. Feldman, Coll. Int. du CNRS, Publ. n° 217 (Paris 1974), p. 87.

[8] N. Beverini, K. Ernst, M. Inguscio and F. Strumia, Appl. Phys. B37, 17 (1985).

[9] L.S. Cutler, US patent 4, 425, 653.

[10] G. Alzetta, A. Gozzini, L. Moi and G. Orriols, Nuovo Cimento 36B, 5 (1976).
E. Arimondo and G. Orriols, Lett. Nuovo Cimento 17, 333 (1976).

RESONATEUR MINIATURE A ENTREE ET SORTIE OPTIQUES

E. DIEULESAINT, D. ROYER, Ph. de HOND
Lab. d'Acoustoélectricité de l'Université P.& M. Curie
10, rue Vauquelin, 75231 Paris Cedex 05
et
B. MORBIEU
Division Aérospatiale de la Société Crouzet
25, rue Jules Védrines, 26027 Valence.

MINIATURE RESONATOR WITH OPTICAL INPUT AND OUTPUT

Electrical connections render piezoelectric resonators vulnerable to electromagnetic disturbances. A way of overcoming this vulnerability is to excite the resonator by photothermal effect i.e. by a light beam whose intensity is modulated, and to detect its vibration optically. These optical methods of excitation and detection have the additional advantage of requiring no mechanical contact.

The object of our experiments is a miniature resonator. It consists of a thin plate of quartz 0,25 x 0,70 x 5 mm^3 vibrating in a flexural mode at 60 kHz and constitutes an essential part of a pressure sensor. The optical source is a laser diode (20 mW) with an optical fiber. Several techniques of detection have been tested: homodyne interferometer, knife wedge, rotation of polarization. The last one has been found the most attractive. It is very stable and the tolerance for positioning the probe with respect to the resonator is quite acceptable.

1 - INTRODUCTION.

L'intérêt d'exciter et de détecter les vibrations de résonateurs mécaniques par des moyens optiques tient à plusieurs raisons:

-i) les matériaux, doués de piézoélectricité et aptes, par leurs dimensions, coupes, coefficients de qualité,..., à constituer des résonateurs sont rares, en dehors du quartz ($Si O_2$) et du tantalate de lithium ($Li Ta O_3$). Il est donc utile d'étudier d'autres matériaux, possédant des qualités mécaniques comparables à celles de ces deux cristaux, comme le corindon (monocristal d'alumine: $Al_2 O_3$) ou le grenat d'yttrium et d'aluminium (YAG). Or un moyen relativement simple d'exciter des résonateurs taillés dans des cristaux non piézoélectriques est l'effet photothermique: génération périodique de déformations par chauffage local à l'aide d'un faisceau lumineux dont l'intensité est modulée [1-6].

-ii) L'excitation par effet piézoélectrique ne présente pas que des avantages. D'une part, la soudure fil - cristal charge mécaniquement le résonateur; d'autre part, la boucle source électrique - fils - cristal rend le résonateur sensible aux perturbations électromagnétiques. L'excitation et la détection optiques qui n'exigent aucun contact mécanique ne présentent pas ces inconvénients.

L'objet de cet article est de décrire une expérience relative à un résonateur qui est l'élément essentiel d'un capteur de pression Crouzet. Ce résonateur en quartz, normalement excité, dans les dispositifs développés jusqu'à présent, par une tension électrique, est,ici, mis en vibrations par un faisceau lumineux. Trois méthodes de détection ont été essayées. La meilleure est celle de la rotation de polarisation. Avant de procéder à ces essais, nous avons relevé la déformée et la réponse fréquentielle du résonateur, en fonctionnement normal, à l'aide d'une sonde optique interférométrique.

2 - STRUCTURE DU CAPTEUR DE PRESSION.

Le principe de fonctionnement de ce capteur [7] est illustré par la figure 1. Le résonateur, assimilable à une poutre encastrée à ses deux extrémités, vibre sur un mode de flexion. La pression à mesurer engendre par l'intermédiaire du soufflet et du fléau une force, suivant l'axe du barreau, qui modifie la fréquence de résonance. La basse pression (vide primaire), dans l'enceinte, diminue l'amortissement dû à l'air et sert de référence. Les pertes d'énergie vers le support sont réduites grâce à deux masselottes mobiles en rotation. Masselottes et lames élastiques forment un système dont les caractéristiques sont calculées pour annuler l'effort tranchant dans les éléments de liaison avec les supports. Les dimensions du résonateur en quartz (de coupe Z, l'axe du barreau est parallèle à l'axe cristallographique X) sont indiquées sur la figure 1.

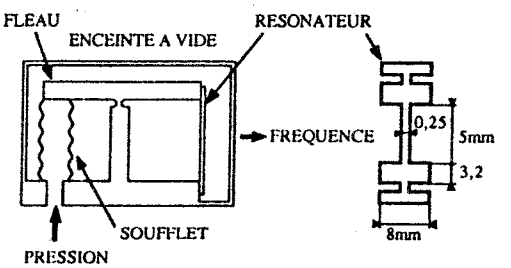


Fig. 1. Capteur de pression Crouzet (type 51). La fréquence du résonateur qui vibre en flexion dépend de la force engendrée, suivant son axe, par la pression dans le soufflet.

3 - RELEVE OPTIQUE DE LA DEFORMEE ET DE LA REPONSE FREQUENTIELLE DU RESONATEUR.

En vue de connaître l'ordre de grandeur des déplacements en fonctionnement normal et de comparer leur distribution théorique et réelle, nous avons relevé la déformée du résonateur lorsqu'il est excité piézoélectriquement. Ce relevé a été effectué à l'aide de la sonde interférométrique [8,9], de résolution 10^{-4} Å/$\sqrt{Hz}$, mise au point au laboratoire (figure 2). Le quartz était collé sur un support de capteur (dépourvu de soufflet) placé à l'intérieur d'une enceinte à vide primaire munie de hublots.

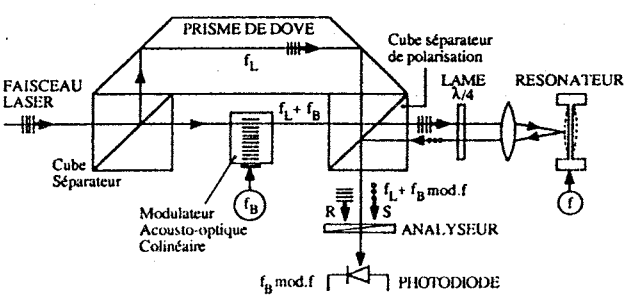


Fig. 2. Sonde optique hétérodyne utilisée pour relever la déformée du résonateur et sa réponse fréquentielle.

Les points expérimentaux portés sur la figure 3 sont très proches de la courbe prévue analytiquement [10]. L'amplitude du déplacement au centre est 120 Angström. Les deux noeuds de vibration se situent aux points prévus.

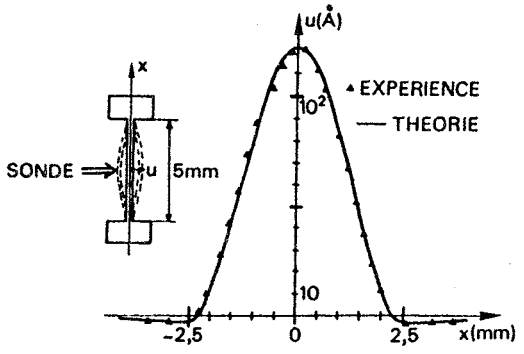


Fig. 3. Déformée du résonateur relevée par sonde interférométrique hétérodyne. Le déplacement mécanique au centre est: 120 Å.

La figure 4 représente la réponse fréquentielle du résonateur. L'amplitude et la phase de la vibration ont été mesurées au centre de la lame. La fréquence de résonance (59 345 Hz) diffère peu de celle prévue (59 320 Hz). Le coefficient de qualité (Q) déduit de cette réponse est 15 000. Les expériences menées avec trois autres résonateurs montrent que la valeur de ce coefficient dépend des conditions de collage sur le support (15 000 < Q < 60 000).

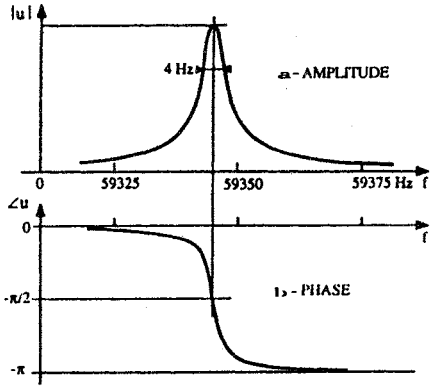


Fig. 4. Amplitude (a) et phase (b) de la réponse fréquentielle du résonateur excité piezoélectriquement (f_r = 59 345 Hz) sous vide primaire ($p = 10^{-2}$ Torr).

4 - EXCITATION PHOTOTHERMIQUE

Un résonateur de coefficient de qualité environ 60 000 et de fréquence 49 330 Hz a été excité par une diode laser (810 nm, puissance maximale: 50 mW) munie d'une fibre optique à modes multiples (Ø: 100 µm), suivant le schéma de la figure 5. Le faisceau, à la sortie de la fibre, est focalisé, à l'aide d'une lentille, soit au centre (point A) soit à l'extrémité (point B) du résonateur. Pour une puissance lumineuse moyenne en sortie de fibre d'environ 20 mW l'amplitude de la vibration est 450 Å si la lame est chauffée en A et 730 Å si la lame est chauffée en B. Le rapport de ces amplitudes est égal à celui des moments fléchissants en ces deux points [10].

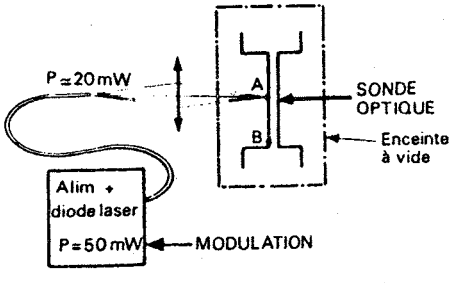


Fig. 5. Excitation par diode laser et fibre optique. Le déplacement, mesuré par la sonde interférométrique de la figure 2, est maximal lorsque le point d'impact est B.

5 - DETECTION OPTIQUE

La sonde hétérodyne, utilisée pour mesurer les déplacements, est un instrument de laboratoire, difficilement miniaturisable à l'échelle d'un résonateur. Aussi avons nous essayé d'autres moyens de détection: interféromètre homodyne, technique du couteau, interaction acousto-optique (rotation de polarisation).

Deux interféromètres homodynes ont été réalisés. Leurs schémas sont représentés sur les figures 6 et 7. Dans le premier, le faisceau incident, polarisé horizontalement, est divisé en deux parties par un prisme de Wollaston [11]. La première partie (R) est focalisée sur l'extrémité du résonateur (noeud de vibration), la deuxième partie (S) sur le centre. Les polarisations de ces deux faisceaux sont perpendiculaires. Chacune est décalée de π/4 par rapport à celle du faisceau incident. Après réflexion sur la lame, les faisceaux se recombinent dans le prisme de Wollaston en gardant leur polarisation respective et interfèrent sur la photodiode.

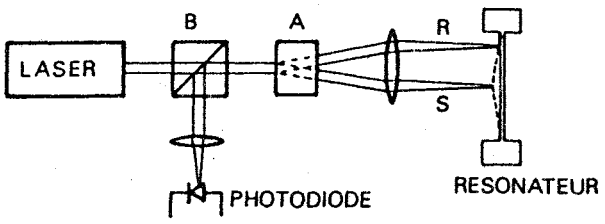


Fig. 6. Détection de la vibration par interférométrie différentielle entre le faisceau de référence R et le faisceau sonde S. A est un prisme de Wollaston. B est un cube séparateur de polarisation.

Dans le deuxième interféromètre, le faisceau laser source est introduit dans une des fibres monomodes du coupleur [12]. La partie de la puissance optique réfléchie à l'autre extrémité de la fibre fait office de référence; la partie réfléchie par l'objet dans la fibre subit une modulation de phase. Comme précédemment, l'interférence des deux faisceaux sur la photodiode transforme la modulation de phase en modulation d'intensité.

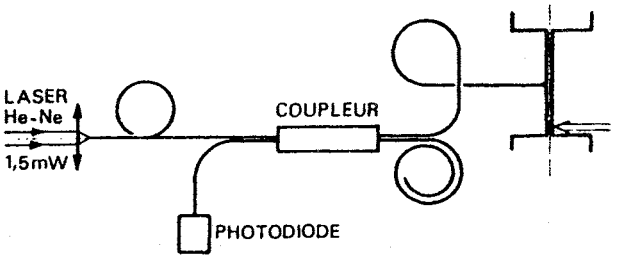


Fig. 7. Detection du déplacement par interféromètre à fibres optiques.

La vibration du résonateur, excité par diode laser, a été détectée avec les deux interféromètres mais le signal est très instable et d'une sensibilité excessive aux parasites.

La *technique du couteau* n'impose aucune contrainte sur l'état de la surface de la lame, contrairement aux méthodes interférométriques. Elle fournit pour des déplacements de plusieurs centaines d'Angström un signal d'amplitude satisfaisante (résistance de charge de la diode 1kΩ → 0,2 mV avec un rapport signal / bruit de 30 dB) mais très vite détérioré par les moindres perturbations mécaniques.

La méthode consistant à tirer profit de la *biréfringence* périodique engendrée par les contraintes au cours de la vibration (figure 8) s'est révélée la meilleure: le résonateur est placé entre un analyseur et un polariseur croisés de façon qu'au repos la photodiode ne reçoive aucune lumière. Lorsque le résonateur vibre, les contraintes induisent une biréfringence qui fait tourner la polarisation du faisceau lumineux; la photodiode est éclairée périodiquement. La rotation est maximale lorsque le faisceau traverse la zone à contraintes élevées, près d'une masselotte.

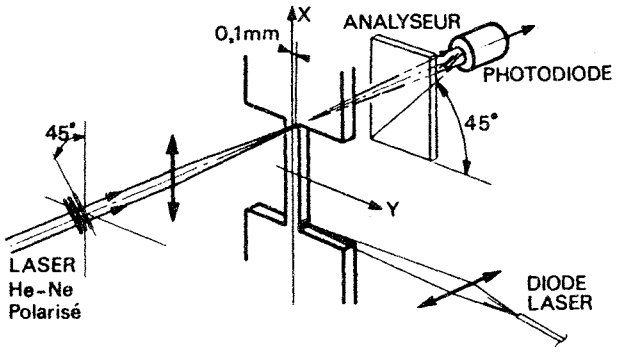


Fig. 8. Détection de la vibration par rotation de polarisation du faisceau sonde.

La figure 9 représente le signal fourni par la photodiode pour une amplitude de vibration de l'ordre de 100 Å. Ce signal est en phase ou en opposition de phase, par rapport à la tension de commande de la diode laser, suivant que le faisceau sonde traverse la lame d'un coté ou de l'autre de la fibre moyenne. Il est très stable, quasiment insensible aux chocs et vibrations extérieures.

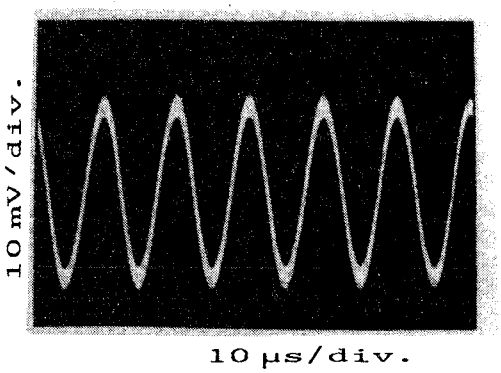


Fig. 9. Signal à la sortie de la photodiode, produit par la rotation de polarisation du faisceau sonde. f=58 186 Hz, gain 300, bande passante 10 kHz.

La comparaison, sur la figure 9, des signaux fournis par cette technique et celle du couteau est instructive. La deuxième requiert une précision de l'ordre de 10 µm quant à la position relative de la tache focale par rapport à l'arête de la lame. La détection par rotation de polarisation est cinq fois moins exigeante (≅ 50 µm).

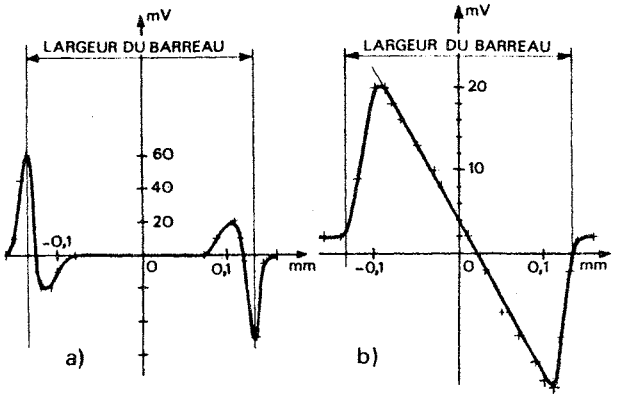


Fig. 10. Variation de l'amplitude du signal en fonction de la position relative du point de focalisation du faisceau sonde. (a) méthode du couteau, (b) rotation de polarisation.

CONCLUSION

Ces expériences, portant sur des résonateurs miniatures, vibrant sur un mode de flexion, à une fréquence comprise entre 40 et 60 kHz, montrent l'intérêt des techniques optiques dans le domaine des résonateurs.

- La déformée d'un résonateur a été relevée à l'aide d'une sonde optique, sensible et stable. La distribution réelle des déplacements a ainsi été comparée avec la distribution prévue, le rôle de masselottes de découplage a été vérifié.
- Plusieurs résonateurs ont été excités par un faisceau lumineux guidé par une fibre optique avec des amplitudes de vibration du même ordre de grandeur que celles obtenues en fonctionnement piézoélectrique.
- La vibration a été détectée par trois méthodes. La meilleure est basée sur la rotation de polarisation d'un faisceau qui traverse la lame de quartz. Elle est stable, la tolérance de positionnement du faisceau par rapport au cristal est tout à fait acceptable.

Ces techniques utilisées ici avec un résonateur en quartz, élément d'un capteur de pression, s'appliquent naturellement à des résonateurs taillés dans d'autres matériaux [13].

REMERCIEMENTS

Les auteurs tiennent à remercier Hugues Duchaussoy de la DRET, Paris, qui a initié et soutenu cette coopération entre le laboratoire d'Acoustoélectricité de l'Université P.& M. Curie et la société Crouzet.

REFERENCES

1 - E. DIEULESAINT, D. ROYER et C. BONNEFOY,"Mechanical excitation of a membrane by an optical beam", IEEE Ultrasonics Symposium Proc., pp. 802-805 (1981).

2 - E. DIEULESAINT, D. ROYER and X. SERVAJAN, "Sapphire resonator", Electron. Lett., **22**, pp. 266-268 (1986).

3 - E. DIEULESAINT, D. MAZEROLLE and D. ROYER, "YAG resonator", Electron. Lett., **23**, pp. 581-582 (1987).

4 - S. VENKATESH, and B. CULSHAW, " Optically actived vibrations in a micromachined silica structure", Electron. Lett., **21**, pp. 315-317 (1985).

5 - R.M. LANGDON and D.L. DOWE,"Photoacoustic oscillator sensors", Proceedings of the SPIE Conference on Fibre optic sensor II, pp.86-93, the Hague (1987).

6 - M.V. ANDRES, K.W.H. FOULDS and M.J. TUDOR, "Optical activation of a silicon vibrating sensor", Electron. Lett.,**22**, pp. 1097-1099 (1986).

7 - G. de SORBIER, " Capteur de pression à quartz vibrant type 51" Actes du 1er Forum Européen Temps-Fréquence, pp.179-183 Besançon (1987).

8 - D. ROYER, E. DIEULESAINT and Y. MARTIN, "Improved version of a polarized beam heterodyne interferometer", IEEE Ultrasonics Symposium Proc., pp. 432-435 (1985).

9 - D. ROYER and E. DIEULESAINT, " Optical probing of the mechanical impulse response of a transducer", Appl. Phys.Lett., **49**, pp. 1056-1058 (1986).

10-Ph. de HOND, " Résonateur mécanique miniature à entrée et sortie optiques", Thèse d'Ingénieur CNAM, Paris (1988).

11-H.K. HEINRICH, B.R. HEMENWAY, K.A. Mc GRODDY and D.M. BLOOM, "Measurement of real-time digital signals in silicon bipolar junction transistor using a noninvasive optical probe", Electron. Lett., **22**, pp. 650-652 (1986).

12-N.A.D. STOKES, R.M.A. FATAH and S. VENKATESH,"Self excited vibrations of optical microresonators", Electron. Lett., **24**, pp. 777-778 (1988).

13-B. MORBIEU, E. DIEULESAINT et D. ROYER, "Résonateur vibrant à excitation et détection optiques à usage de capteur", Brevet Français N° 88 12413 déposé le 19-09-1988.

SPECIAL MINATURISED SC CUT QUARTZ CRYSTAL RESONATOR FOR SEVERE ENVIRONMENTS APPLICATIONS

Richard DELAITE

L.C.E.P./E.N.S.M.M - Route de Gray - 25050 BESANCON CEDEX - FRANCE

Groupe de Recherche en Génie Thermique - 90016 - BELFORT CEDEX - FRANCE

SUMMARY

Space applications require piezoelectric resonators with excellent stability under environmental conditions which usually cause frequency perturbations.

For improved long term drift and short term stabilities under accelerations and external pressure variations a BVA SC cut 10 MHz 3rd overtone specially mounted crystal is proposed.

Several parameters of quartz fixation including number and location of quartz bridges need a theoretical study which has already been done in the AT cut case. We now present a stress sensitiveness coefficient calculation which connect frequency variations to effective stresses in an SC cut crystal.

The supporting structure between quartz resonator and case also needs special care and study. A rigid peripheral structure associated to elastic fixation is used. Computer aided design is used to turn theoretical results into pratical realization.

Measurements and tests on four resonators are performed. Results are presented.

RESUME

Pour les applications spatiales, il est nécessaire d'utiliser des résonateurs piézoélectriques ayant les meilleures stabilités ; les conditions d'environnement sont la cause des perturbations de fréquence.

Pour obtenir des stabilités à long terme ainsi que des stabilités à court terme dans le domaine de l'accélération ou des variations de pression extérieure, nous proposons de combiner l'utilisation d'un cristal de coupe SC, 10 MHz P3, monté en B.V.A. avec une structure périphérique spécifique.

Le choix du nombre de ponts reliant la plaque vibrante à la couronne de quartz et leur position relative au repère cristallographique est lié aux résultats de l'étude théorique systématique des actions appliquées à un cristal de coupe SC. De la même manière que pour la coupe AT, nous définissons à présent, pour un cristal de coupe SC, le coefficient de sensibilité déduit de l'étude de la variation de fréquence fonction des différentes contraintes effectivement appliquées.

Le support reliant l'empilage de quartz au boîtier a nécessité une étude spéciale : une structure périphérique rigide associée à des fixations élastiques est utilisée. L'optimisation des résultats de cette recherche, avec l'aide de la conception assistée par ordinateur, permet la réalisation pratique de cette structure.

Des mesures et des tests sont obtenus sur les quatre premiers résonateurs de ce type ; les résultats sont présentés.

1 - INTRODUCTION

Under environmental conditions a quartz B.V.A. resonator (1) undergoes a change of frequency. This change is linked to effects which are directly applicable to quartz crystal plate and which function according to the type of disturbance, the type of support, the number of connecting points between the crystal and the support and also according to their position in relationship to quartz crystallographic reference.

At the level of the connecting points to the thin disk of crystal, the resulting effects correspond to : radial compression, a shearing at the edge and a uniformed bending of the plate.

First of all one can study the influence of these effects in the case of the self-suspended (2) SC cut quartz crystal with three or four connecting bridges symmetrically distributed.

Following this, from these results and the knowledge from work already done in this field (3)(4), it is possible to deduce the outline structure of a new resonator with weak simultaneous sensitivity.

In order to verify the theoretical results that have been found it is necessary to make and assemble the first resonators of this new generation, so that they can be tested.

2 - INFLUENCE OF EFFECTS DIRECTLY APPLICABLE TO CRYSTAL

2.1/ The crystal shearing

Experiments have been carried out to interfer with the static shearing of resonant quartz in the conventional resonators (5).

In the case of the suspending quartz resonators of type Q.A.S. secured by two bridges and with supports (6)(7), the consequence of a diametrical

compression of the ring is the appearance of stress at the level of the crystal, which is a function of the angle formed by the axis of compression and the axis of the two briges.

In certain particular cases or for certain applications, it is possible to determine the specific solutions for reducing - or cancelling - the influence of the shearing.

In order to determine the general manner of the way in which the crystal shearing disturbs its frequency one limits oneself to the study of the thin disks of quartz which have the vibration of the thickness of the shearing. Neglecting the compressibility of the crystal, the frequency change due to the stress distribution can be given by the equation **(8)(9)** :

$$\frac{\overline{f_k} - f_k}{f_k} = \frac{-1}{2c_{66}} \left(\overline{T_2} + \sum_{i=1}^{i=6} k_i \overline{T_i}\right) \qquad (1)$$

where the barred symbol is relative to the final state, c_{66} is the elastic coefficient corresponding to the vibration, the $\overline{T}_{i's}$ the static stresses and $k_{i's}$ are functions of stiffness and compliance for the second and third order of a given cut crystal.

In case of the SC cut, all coefficients k_i are not null. For AT cut only k_5 and k_6 are null.

When a crystal quartz is uniquely subjected to the moments on the lateral faces or in the plan of the plate, the static stresses of the shearing are distributed around the crystal. These stresses are not null in the central part. The relative change of frequency can be written thus :

$$\frac{\overline{f_k} - f_k}{f_k} = \frac{-1}{2c_{66}} (k_4\overline{T}_4 + k_5\overline{T}_5 + k_6\overline{T}_6) \qquad (2)$$

In order to characterize the influence of these phenomena of the shearing on the frequency of the crystal resonator, a shearing sensitivity coefficient "Q_c" may be introduced by the following definition. The shearing stresses are directly linked, for each stress component, to the characteristics of the support of the crystal ; the coefficient Q_c takes the general form of the equation :

$$Q_c = \frac{1}{2c_{66}} \left(\sum_{i=4}^{i=6} k_i t'_i\right) \qquad (3)$$

where t'_i is a coefficient function of the characteristics of the support and of the type of connection to the crystal.

In order to eliminate the influence of the shearing on the frequency of the resonator, it is necessary to cancel out the coefficient of sensibility Q_c. One solution consists in making null the values that the coefficient t'_i can take. To do this, it is sufficient to bind the vibration plate to its environment by means of a non elastic connection which cannot transmit to it any moments. In the case of the crystal, bound by three or four quartz bridges of short length and a rigid support, it is like a monolithic block.

This principle is also retained in the study of the flexion of the crystal as well as for research into a new peripheral structure of the resonator.

2.2/ Compression of the crystal

Previous studies have shown the advantages and disadvantages of systems with crystals maintained by two equal and opposite forces acting along a diameter **(10)(11)**, or by 2α angular-sectors **(8)** or by integral peripheral bonding **(12)(13)**.

In the case of a self-suspended quartz crystal in a B.V.A. system, utilising three or four connection points distributed symmetrically and inducing a compression of a crystal by the application of forces on the edges, the stresses at the centre of the crystal are equal and can be written in the form :

$$T_1 = T_3 = T, \quad \forall \Psi \qquad (4)$$

where T is a function of h_o half the thickness of the crystal, R its radius and F the intensity of the forces applied.

From equation (1) the frequency change due to the stress distribution previously studied is obtained from the equation :

$$\frac{\overline{f_k} - f_k}{f_k} = \frac{-1}{2c_{66}} (k_1 + k_3)T \qquad (5)$$

In order to characterise the influence of the phenomena of compression on the frequency of the crystal resonator, one can introduce a sensitivity coefficient of compression "Q_F" such that :

$$Q_F = \frac{1}{2c_{66}} (k_1 + k_3) \qquad (6)$$

This coefficient is independent of the azimuth Ψ of the applied forces and uniquely dependent on the coefficients of stiffness and compliance for the second and third order of the crystal cut utilized.

At a temperature of 25°C, one obtains $Q_F = (0.85).10^{-11}\ N^{-1}m^2$ for a SC cut crystal sustaining symmetrical compression in three or four points.

2.3/ Flexion of the crystal

Assuming that the crystal resonator is bound to its environment by a non elastic connection and that it has the form of a plate with circular planes, it bends when it is subjected to an axial acceleration.

In the case of one maintened by three or four symmetrically positioned bridges, the maximum stresses of flexion in the central part of the disk may be expressed by the equation :

$$T_1 = T_3 = T, \forall \Psi \qquad (7)$$

Applying axial acceleration Γ_2, one obtains the flexion of the cristal plate. The stresses are the function of axial acceleration such that **(9)** :

$$T = t_i \frac{\rho R^2}{2h_o} \Gamma_2 \qquad (8)$$

where ρ is the volumique mass of quartz and t_i a not null coefficient dependent on the type of connection (t_i = 3.95 with three bridges and t_i = 3 with four bridges).

Taking equation (1) together with the hypothesis of proportionality **(9)** in the change of frequency in function of the stresses which are not null T_1 and T_3, one obtaines the equation :

$$\frac{f_k - f_k}{f_k} \alpha \frac{-1}{2c_{66}} (k_1 + k_3)\, t_i \frac{\rho R^2}{2h_o} \Gamma_2 \qquad (9)$$

In order to characterise influence on the self-suspended SC cut quartz crystal, if is convenient to introduce a sensitivity coefficient of flexion "Q_a" defined as :

$$Q_a = \frac{1}{2c_{66}} (k_1 + k_3)\, t_i \qquad (10)$$

The frequency change is proportional to the acceleration applied as well as the sensitivity coefficient Q_a ; the latter cannot cancel itself and remains constant, having a position relative to the connecting bridges. The coefficient Q_a is function of coefficients of stiffness and compliance of the crystal cut utilized.

At the temperature of 25°C, the coefficient Q_a of SC cut takes the value $(34).10^{-11}$ $N^{-1}m^2$ if the crystal is maintened by three bridges and $(25.5).10^{-11}$ $N^{-1}m^2$ if it is maintened by four bridges.

The relative change of frequency corresponding to 1 g axial acceleration is respectively $(3.1).10^{-11}$/g for three bridges and $(2.4).10^{-11}$/g for four bridges. These values correspond to the limiting values below wich, with this type of configuration, it is not possible to go. Only the compensation affects at the level of the crystal can be envisaged in order to reduce again the axial acceleration sensitivity of the resonators ; this solution show, nevertheless, how extremely it is to make it work.

3 - CONCEPTION AND PUTTING INTO PRACTICE A RESONATOR WITH WEAK SIMULTANEOUS SENSITIVITY

If one decides to cancel out the possibilities of the crystal shearing phenomena by choosing a rigid support, the sensitivity of the crystal resonator is no longer part of connection but only quantified by the compression and flexion phenomena of the crystal.

On the case of B.V.A. resonator system **(1)(14)** equiped with a SC cut quartz crystal, the theoretical study has shown that the quartz self-suspension made with four symmetrically positionned connecting points is preferential. This realisation will give a reduced simultaneous sensitivity, for a quartz crystal with a vibration frequency of the thickness of the shearing, when the effects of compression are directly applied and when all the quartz plate is amenable to bending.

The way to obtain a radial compression result which is limited and uniformed at the level of the connection bridges with the crystal, the ring of the quartz is maintained on the periphery with an associated clamped axial.

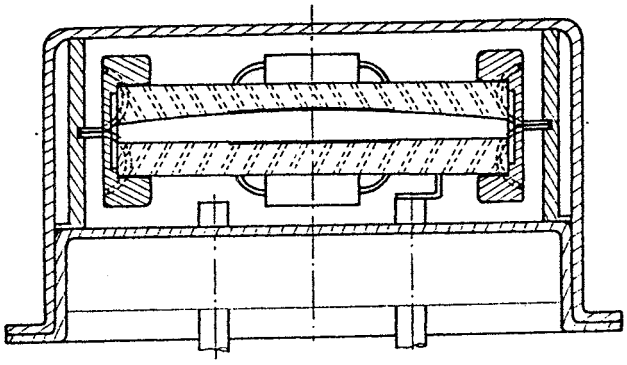
At the time of a variation of temperature or of external pressure or of an application of acceleration, the resulting distortion of the peripheral structure is defined by only one transmission of weak variation of stress on the clamped axial, necessary for maintaining the ring in position. The radial compression stress, which results, is reduced by the slight convexity of the crystal, but it is transmitted directly to the level of the four quartz bridges. These bridges being positionned so as to obtain a null sensitivity at the time of compression ; the frequency will then remain stable.

This result is admissable even for a null defect of position at the time when the bridges come into effect, in relationship with the symmetry and quartz crystallographic reference. In practice, a tolerance exists and the influence of which has already been shown **(15)** ; the self-suspended quartz system limits this defect and its influence in comparaison to the conventional resonator model.

The peripheral structure is composed of two rigid quartz half-supports and four rigid nippers associated to a symetrical suspension with eight springs, which have three degree of liberty in translation with equal stiffness. A ring is filled into the tip of the hood to remove all contact from the suspension of the distorted part of the holder. The influence of the outside pressure is stopped with the choice of the strap-ring connection in the median plan of the resonator.

All these elements of the resonator are mounted in an improved ageing holder **(16)** with a cold-weld enclosure and chimney to obtain a long term stability. Computer aided design is used to turn theoretical results into practice (Figure 1).

Coupe



Detail of symmetric internal structure of miniaturized piezoelectric high stability resonator : the plan-convexe cristal with two non adherent electrodes, the peripheral structure with two rigid quartz half-supports, four rigid nippers, a symetrical spring and a ring mounted in a cold-weld enclosure with chimney.

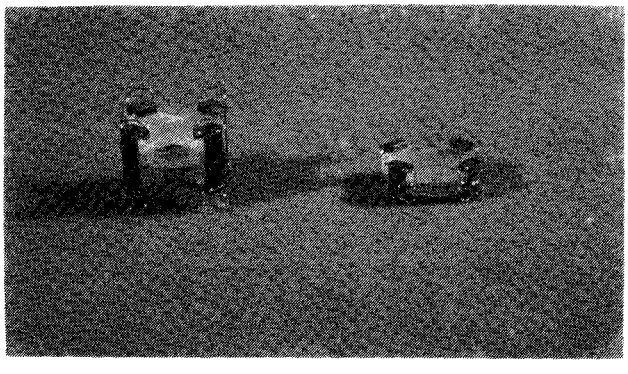
Figure 1

Measurements and tests have been performed on the first four resonators. During three months, the long term stability is contained between $(1).10^{-11}$ and $(2).10^{-11}$. For the quality coefficient, the minimum value is $(1.15).10^{6}$ and the maximum value is $(1.31).10^{6}$.

Mechanical sensitivities of pressure and acceleration have been measured for two bars and two g of variation in the aerospatial agency : Office National d'Etudes et de Recherches Aérospatiales. For the influence of external pressure the lowest sensitivity is $(1).10^{-11}$ for 1 bar variation. For the acceleration frequency variation, the minimum sensitivity is $(1).10^{-11}$ for an axial acceleration of 1 g and $(3).10^{-11}$ for a radial acceleration of 1 g in the crystal plane.

This specific structure, which has recently been protected by a patent **(16)**, has the advantage of being able to be used in any type of crystal resonator. The proposal here is to make within this frame work a B.V.A. system with a self-suspended crystal **(2)** and two non-adherent electrodes **(17)**.

The photograph presented, shows two versions of the internal structure with quartz crystal and electro-plates : the usual version resonator and the miniaturized version resonator.



4 - CONCLUSION

The influence of effects directly applicable to a crystal resonator have been demonstrated. The study has been concerned with the phenomena relating to compression, to flexion and to a shearing from a vibrating plate of quartz crystal. Thus, the anisotropy of the quartz is taken into account in order to determine the three sensitivity coefficients applicable to a self-suspended SC cut quartz crystal.

Turning to the possibility of the crystal shearing, one solution which has been given is no longer to generate them when external disturbances are applied to the resonator.

For the phenomena of flexion and compression particular solutions are put forward to obtain better stabilities. Thus, it is possible to have a crystal -resonator with reduced sensitivity whatever the environmental conditions.

Starting from configuration of the crystal which has been given a miniaturized peripheral structure to reduce obstruction is determined in order to limit to a maximum its influence and the transmission of stresses to the vibrating quartz disk.

This paper shows the interest in making work the necessary research for the conception of a new product. The general study of miniaturized piezoelectric high stability resonators consists in putting forward the evidence for different phenomena in a way that will characterise and in the end give a full conception of a new system with better performance.

The tests on prototypes confirm the interest of such proceeding.

ACKNOWLEDGMENT

We wish to thank Aerospatial Agency O.N.E.R.A. 92322 - Chatillon Cedex (France) for tests.

REFERENCES

(1) BESSON R., "A new piezoelectric resonator design". Proceeding of 30th Annual Frequency Control Symposium, Fort Monmouth, 1976.

(2) BESSON R. "Résonateur piézoélectrique à cristal auto-suspendu". Brevet d'invention n° 78 022 61, 1978.

(3) DELAITE R. et VALENTIN J-P, "Compensation des effets mécaniques dans un cristal de coupe SC". Revue de Physique Appliquée, février 1986.

(4) DELAITE R., "Une nouvelle génération de résonateurs piézoélectriques miniaturisés de haute stabilité". Conférence, actes du Premier Forum Européen Temps-Fréquence, Besançon 18, 19 et 20 mars 1987.

(5) JANIAUD D., "Modélisation de l'influence d'une accélération sur la fréquence des résonateurs à quartz". Thèse de Docteur-Ingénieur, faculté des Sciences et des Techniques de Besançon, 1978.

(6) AUBRY J-P, CRAVEUR J-C, DEYZAC F., FRAGNEAU M. et JANIAUD D., "Analyse des mécanismes de sensibilités accélérométriques et barométriques des résonateurs à quartz de type QAS". Actes du Premier Forum Européen Temps-Fréquence, Besançon, 18, 19 et 20 mars 1987.

(7) AUBRY J-P, BONNOTTE B., CRAVEUR J-C, DEYZAC F. et JANIAUD D., "Optimisation de l'orienation des résonateurs auto-suspendus". Actes du Second European Frequency and Time Forum, Neuchâtel, 16, 17 et 18 mars 1988.

(8) DELAITE R. and VALENTIN J-P, "Analysis of force-frequency effect of piezoelectric circular resonators strained by edge angular sector-stresses". Il Nuovo Cimento (Europhysics Journal), Volume 6 D, N° 5, pp 574-584, Dicembre 1985.

(9) DELAITE R., "Influence de la structure de maintien sur la sensibilité accélérométrique et barométrique des résonateurs à quartz de haute stabilité". Thèse de Doctorat n° 005, Université de Franche-Comté, Juin 1985.

(10) BALLATO A.D. and BECHMANN R., "Effect of initial stress in vibrating quartz plates". US. Army Signal Res. and Dev. Lab. Fort Monmouth N.J. Proceedings of the IRE, Feb. 1960, pp 261-262.

(11) RATAJSKI J.M., "The force sensitivity of AT cut quartz cristals". International Business Machines Corporation. Federal Systems Division. Electronics Systems Center, Owego, New-York, 1968.

(12) NAKAZAWA M., LUKASZEK T. and BALLATO A., "Force and acceleration Frequency effects in grooved and ring supported resonators". Proceedings of 35th Annual Frequency Control Symposium, Fort Monmouth, may 1981.

(13) DELAITE R., "Résonateur piézoélectriques à extremum de sensibilité vis-à-vis des contraintes extérieures de pression". Brevet d'invention N° 85 090 58, 14 juin 1985.

(14) BESSON R., "Résonateur à quartz à électrodes non adhérentes au cristal". Brevet d'invention N° 76 010 35, 9 janvier 1976.

(15) DELAITE R., "Précision de réalisation pour résonateurs à quartz HQ". Revue de Physique Appliquée N° 24, pp 83-92, janvier 1989.

(16) BESSON R., DELAITE R., RENARD G., VALENTIN J-P et VATERKOWSKI J-L, "Dispositif démontable de support d'un résonateur piézoélectrique à l'intérieur d'un boîtier". Brevet d'invention N° 88 141 97, 28 octobre 1988.

(17) BESSON R., "A new electrodeless resonator design" Proceedings of 31rst Annual Frequency Control Symposium, Fort Monmouth, 1977.

An investigation on dislocation and channel effects on acoustic waves in quartz resonators by stroboscopic X-ray topography

Y.ZHENG, A.ZARKA & B.CAPELLE

Laboratoire de Minéralogie-Cristallographie,
Universités P. & M. CURIE (Paris VI) et Paris VII, unité associée au CNRS,
4 place Jussieu, 75252 Paris Cedex 05, France

and

J.DETAINT & J.SCHWARTZEL

Centre national d'Etudes des Télécommunications,
PAB/BAG/MCT, 92220 Bagneux, France

Abstract

AT-cut quartz resonators containing dislocations or channels have been studied by the stroboscopic X-ray topography using the synchrotron radiation (LURE, Orsay, France). Contrasts obtained on topographs of vibrating resonators in presence of dislocations show that the acoustic wave is perturbed in the vicinity of dislocations. Simulated images are presented in order to show the existence of the interaction between the acoustic wave and dislocations. The comparison between the experimental images and simulated ones indicates that non-linear terms have to be considered to describe the acoustic deformation field in the vicinity of dislocations. Similar experimental contrasts have also been obtained in vibrating resonators containing channels. These contrasts suggest that the influence of channels may be more important than that of dislocations.

Introduction

Time resolved (1ns) X-ray topography has been performed to observe the acoustic vibrational states in quartz resonators. The principles of the experiments (Fig.1) consists on the exact synchronization between the acoustic vibration and the synchrotron radiation [1,2]. Results already obtained have shown that this technique is very performant to determine low vibration amplitudes (0-7Å) through the interference fringes [2,3]. The proposal here is to clarify whether the dislocation (and channel) effects on acoustic waves can be characterized by X-ray topography.

Topography on dislocations

When an acoustic wave and a dislocation are both present in a crystal, the displacement field can be written in a general form as following:

$$u(r,t) = u_a(r,t) + u_d(r) + u_i(r,t)$$

where $u_a(r,t)$ is the acoustic displacement without dislocation, $u_d(r)$ the static dislocation displacement without acoustic wave and $u_i(r,t)$ the displacement resulted of the interaction between the acoustic wave and a dislocation. For linear piezoelectricity theory, no interaction is predicted and $u_i(r,t)$ is equal to zero.

To clarify the existence or not of the interaction term $u_i(r,t)$, a dislocation was selected

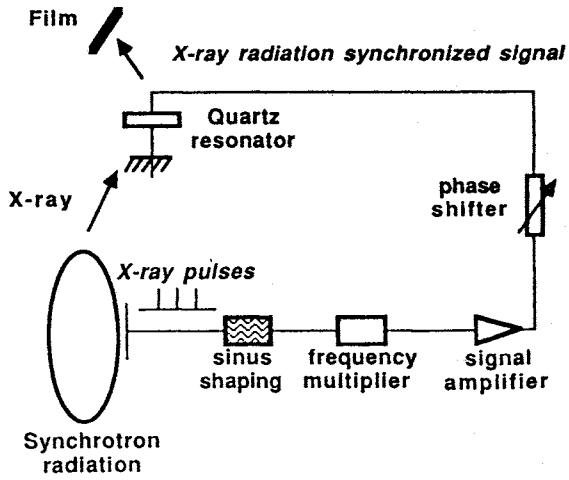


Fig.1 Principles of stroboscopic X-ray topography.

from a synthetic AT-cut quartz resonator operating on fifth overtone thickness-shear mode (Fig.2).

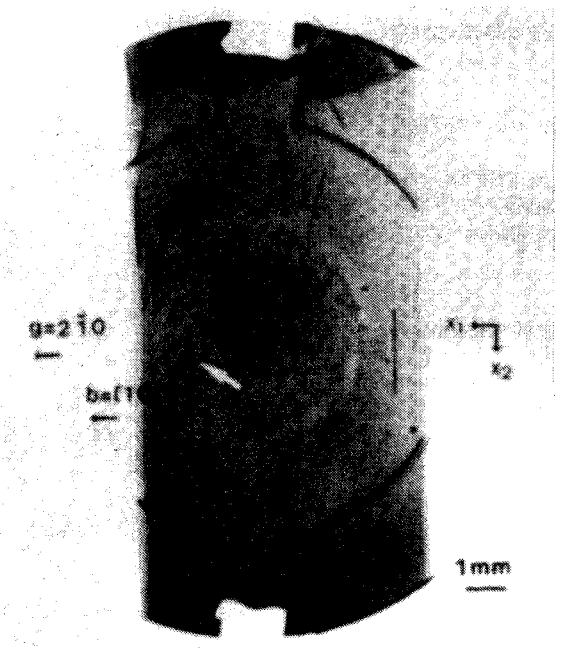


Fig.2 A traverse topograph of a quartz resonator containing dislocations. The dislocation studied by section topographs is indicated by an arrow.

The Burgers vector **b** of the dislocation was identified by X-ray section topographs, **b**=[100] for the present case. It is an almost edge dislocation. Stroboscopic X-ray section topographs were taken with the reflection vector $\mathbf{g}=2\bar{1}0$ which is parallel to the dislocation Burgers vector **b**=[100] and to the acoustic displacement vector $\mathbf{u_a}(\mathbf{r},t)$. The Bragg angle θ_B used is 8.3406° which corresponds to a seclected X-ray wavelength λ of 0.7127Å. The X-ray extinction distance Λ_0 is equal to 96.396μm.

Figure 3a represents a section topograph of the dislocation without acoustic vibration. Two arrows indicate roughly the zone over which the dislocation displacement $\mathbf{u_d}(\mathbf{r})$ (more precisely the induced misorientation of the reflecting planes) can be seen by the X-ray beam. The corresponding simulated image is represented in figure 3b where the dislocation displacement $\mathbf{u_d}(\mathbf{r})$ was calculated from the linear anisotropic elastic theory [4]. It is noticed that most characteristical contrasts of the dislocation image can be simulated. Some differences, namely near the dislocation core image, exist and this is probably due to the polychromatism of the X-ray incident beam [5,6] and other experimental conditions not taken into account in the calculation. When an acoustic wave is excited in the resonator, the stroboscopic section topograph (Fig. 3d) shows two types of contrasts: a) on the upper part of the image far from the dislocation, interference fringes due to the acoustic wave can be seen and they seem to be not altered by the presence of the dislocation; b) near the dislocation on the lower part, a drastical change is recorded in comparison with figure 3a and fringes dues to the acoustic wave are completly destroyed. In order to analyse this image, two simulations were performed. The first simulated image (Fig. 3c) corresponds to the case where only the acoustic wave is considered. This case permits to know the zone where the acoustic wave is not affected by the presence of the dislocation. The interesting point is that, in comparison with the dislocation alone case (Figure 3a), the dislocation influence zone is enlarged (indicated by arrows in figure 3d). The acoustic displacement $\mathbf{u_a}(\mathbf{r},t)$ used in the calculation was from the linear piezoelectricity theory [7,8]. It can be seen that this acoustic deformation model is a very good approximation far from the dislocation. The second simulated image (Fig. 3e) took into account both the dislocation displacement $\mathbf{u_d}(\mathbf{r})$ and the acoustic displacement $\mathbf{u_a}(\mathbf{r},t)$ used preceedingly. This calculation allows to know whether a simple linear addition $\mathbf{u}(\mathbf{r},t) = \mathbf{u_a}(\mathbf{r},t) + \mathbf{u_d}(\mathbf{r})$ of two displacements permits or not to reproduce most features of the experimental image. It can be easily seen that this image is very different of the experimental one over all the zone in the vicinity of the dislocation. It means that the interaction term $\mathbf{u_i}(\mathbf{r},t)$ must be considered near the dislocation and linear theories are not suffisant for the present case. This point is important because it indicates a different behaviour of the acoustic wave around the dislocation and it suggests the existence of a dislocation associated vibration component in the resonator. An exact calculation of $\mathbf{u_i}(\mathbf{r},t)$ should be difficult and its action range around the dislocation can be roughly estimated to be 90μm for the present case. In a general case, $\mathbf{u_i}(\mathbf{r},t)$ should depend a lot on the dislocation geometry also.

Topography on channels

Stroboscopic experiments were also performed on synthetic AT-cut quartz resonators containing channels and operating on the third overtone thickness-shear mode. The X-ray reflection parameters θ_B, λ and Λ_0 taken were similar to previously and the electrical excitation were increased progressively. Figures 4a-f show a series of stroboscopic X-ray section topographs on a channel in function of the vibration amplitudes. For the case of low amplitude acoustic vibrations ($u_0 \leq 10$Å), interference fringes due to the acoustic wave can be seen (Fig. 4b-d) and they are strongly

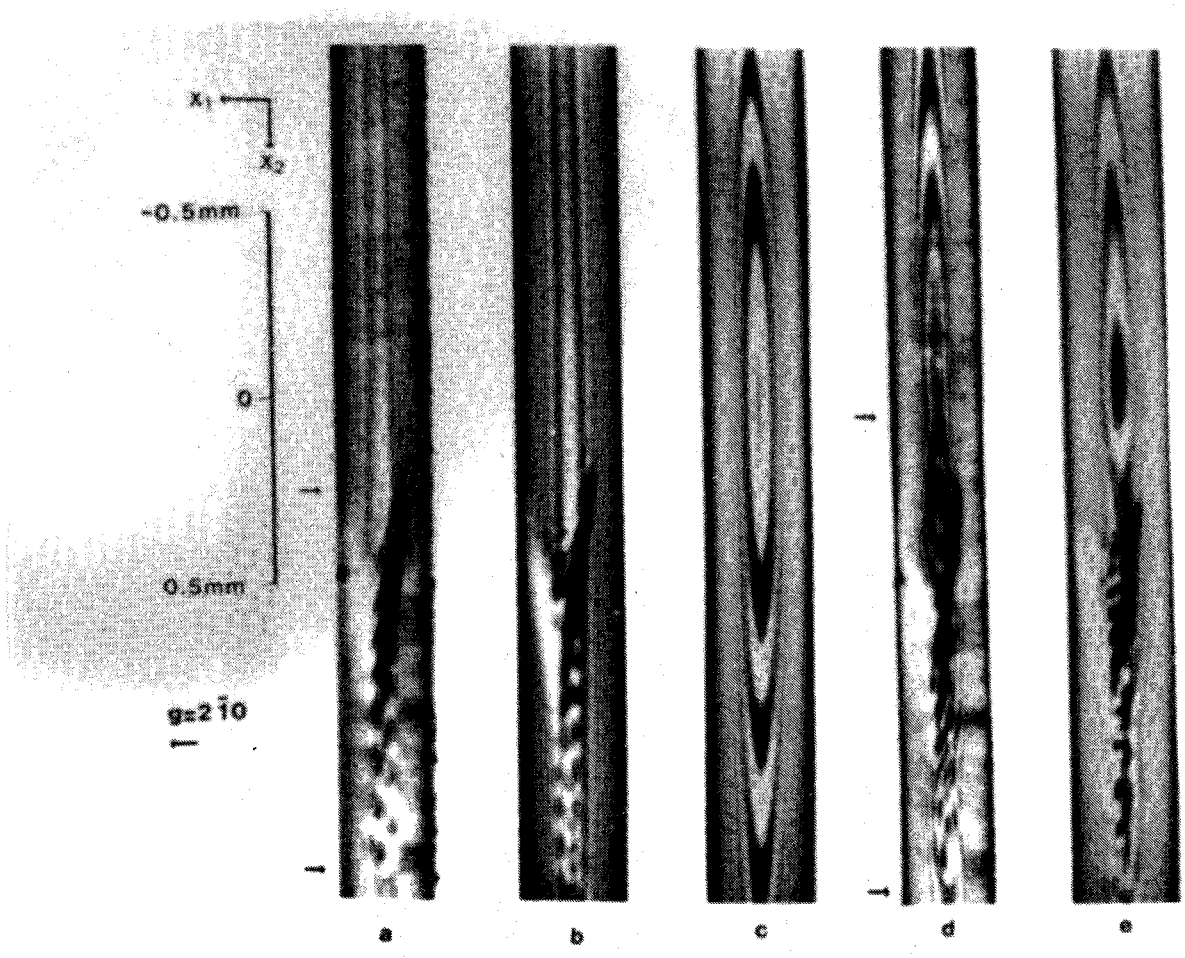


Fig.3 Experimental and simulated stroboscopic X-ray section topographs of a dislocation.

a. Experimental image of the dislocation without acoustic vibration;

b. Simulated image of the dislocation without acoustic vibration;

c. Simulated image of the acoustic vibration without the dislocation where the vibration amplitude u_o for x_2 = 0.0mm is 2.41Å;

d. Experimental image of the dislocation and the acoustic wave where the vibration amplitude u_o is estimated to be 2.41Å;

e. Simulated image by adding the dislocation displacement and the acoustic one where the vibration amplitude u_o is 2.41Å.

deformed in the vicinity of the channel. The channel influence zones are indicated by arrows and it can be seen that these zones are enlarged in function of vibration amplitudes. This means that an interaction between the acoustic wave and the channel should exist and the range of this interaction term can be roughly estimated to be 800μm around the channel for the present case. This value is approximate because of the existence of many different channel types and the ranges of the channel influence may vary more largely than for the case of dislocations.

Conclusion

Stroboscopic experiments performed on quartz resonators containing dislocations and channels have provided interesting informations, namely on the existence of an interaction between the acoustic wave and the crystal defects. It suggests that a non-linear treatment should be carried out to describe a more exact acoustic behaviour in the vicinity of defects. The approximate ranges of this zone are estimated to be 90μm for a dislocation and 800μm for a channel for the present cases. This treatment will be useful to get local acoustic characteristics such as the deviation of the acoustic wave propagation and global acoustic characteristics such as energy dissipation through crystal defects.

Acknowledgment

The authors wish to acknowledge Dr.Epelboin whose assistance was useful for the image simulations. This work was supported by D.R.E.T.

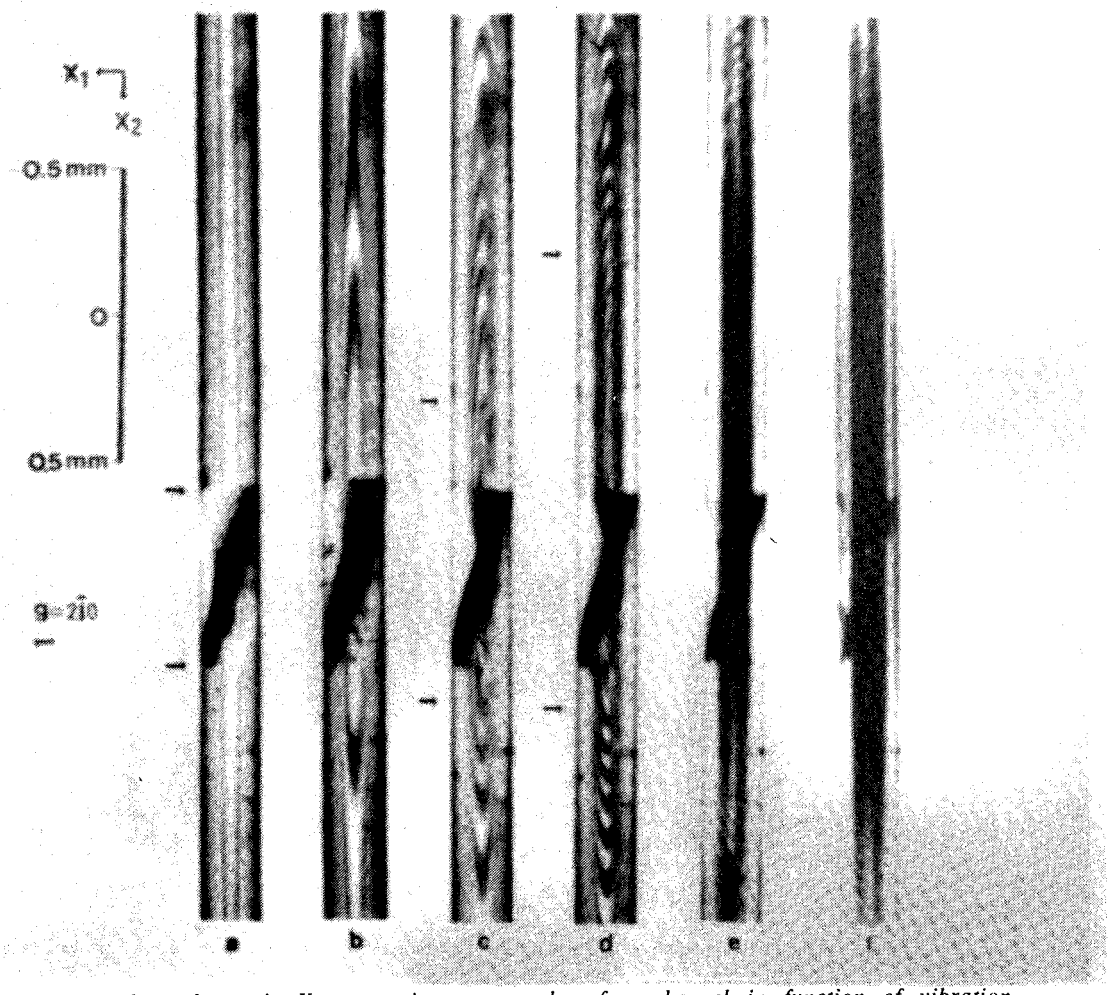


Fig.4 Experimental stroboscopic X-ray section topographs of a channel in function of vibration amplitudes.

a. Image of the channel without acoustic vibration;

b-d. Images of the channel with acoustic vibrations of low amplitude.

e-f. Images of the channel with acoustic vibrations of high amplitude.

(France) through the contract no. 87 34 13 80 04 70 75 01.

References

[1] ZARKA,A., CAPELLE,B., DETAINT,J. and SCHWARTZEL,J. (1987). Proceedings of the 41st Annual Symposium on Frequency Control, p.236-240, 27-29 May 1987, Philadelphia, U.S.A.

[2] ZARKA,A., CAPELLE,B., ZHENG,Y., DETAINT,J. and SCHWARTZEL,J. (1988). Proceedings of the 42nd Annual Symposium on Frequency Control, p.85-92, 1-3 June 1988, Baltimor, U.S.A..

[3] ZHENG,Y., ZARKA,A., CAPELLE,B., DETAINT,J. and SCHWARTZEL,J. (1989). *Pendellösung fringes in stroboscopic X-ray section topography of weakly excited quartz resonators.* Acta Cryst., A**45**, in presse.

[4] HIRTH,J. and LOTHE,J. (1968). *Theory of dislocations*, MacGraw-Hill, New-York, p.411.

[5] AUTHIER,A. and BALIBAR,F. (1970). Acta Cryst., A**26**, 647-654.

[6] MILTAT,J.E.A. and BOWEN,D.K. (1975). J. Appl. Cryst., **8**, 657-669.

[7] TIERSTEN,H.F. and SMYTHE,R.C. (1979). J. Acoust. Soc. Am., **65**, 1455-1460.

[8] STEVENS,D.S. and TIERSTEN,H.F. (1986). J. Acoust. Soc. Am., **79**, 1811-1826.

INFLUENCE OF THE DEFINITION OF MATERIAL CONSTANTS IN THE STUDY OF CRYSTAL PLATES THERMAL BEHAVIOR

B. DULMET, R. BOURQUIN

ECOLE NATIONALE SUPERIEURE DE MECANIQUE ET DES MICROTECHNIQUES
La Bouloie - Route de Gray - 25030 BESANCON CEDEX - FRANCE

ABSTRACT

A consistent description of infinitesimal vibrations of homogeneously heated piezoelectric plates is presented as an application of a macroscopic model previously proposed by Tiersten for thermoelectroelastic interactions. Starting from the natural equilibrium state at some arbitrary reference temperature, the subsequent slow variations of temperature determine a biasing state, the action of which on the actual frequency is easily described since this bias is homogeneous. The field variables are the coordinates of the body in the natural state, thus easily allowing to take into account the resonator geometry in a manner unaffected by temperature variations. The actual frequency is governed by the effective material constants, closely related to the second order partial derivatives of a scalar state function with respect to the deformation and the rotationnaly invariant electric vector. The temperature derivatives of the material elastic constants of quartz are computed up to the third order, from the most widely recognized set of experimental results. The obtained values are used for the computation of the static thermal behavior of some quasi-thickness modes in contoured quartz resonators.

I - INTRODUCTION

The most commonnly quoted values of temperature derivatives of quartz elastic moduli (1) were determined from a complete set of measurements performed on thin plates vibrating in high frequency thickness modes. The treatment of the data was made in the framework of a linear theory and assuming the plate vibrations were pure thickness modes. Since the time of this early analysis, it has become more and more current to use non linear analysis to describe various effects such as the force frequency effect, the amplitude frequency effect, or the thermal transient behavior of quartz resonators. Deriving from the macroscopic model proposed by Tiersten to describe thermoelectroelastic interactions (2), a study of a possibly large quasi static biasing state acting on infinitesimal high frequency vibrations (3) ranks among those non linear analysis. Whenever the bias is not homogeneous, it has been shown that a perturbation method satisfactorily predicts the induced frequency shift (4). The subject of this paper is to present an analysis of inifinitesimal vibrations of homogeneously heated piezoelectric plates. In such a case, the thermal bias is homogeneous and the wave propagation equations and their associated boundary conditions may be written in terms of field variables that can be either the material coordinates of the crystal at the current temperature or its material coordinates at some fixed reference temperature. The initial work in (1) implicitely used material coordinates at the current temperature, while this work uses field variables that are the material coordinates of the crystal at the fixed reference temperature (25°C). Such a choice allows a temperature-independent definition of the geometry of the resonator and its orientation with respect to the crystallographic axes : using the material coordinates at the current temperature would impose an evaluation of the orientation shift resulting from anisotropy of thermal expansion (5, 6). This shift is generally omitted, although its influence on the frequency-temperature response may not be negligible for certain crystal cuts. Its existence is one of the limitations to which the approach followed by Bechmann, Ballato and Lukaszek was bound. Other limitations arise from the vibration model they used and from the linear theory which is not completely consistent with respect to the more general theory of thermoelectroelasticity. Especially, the material coefficients used in the linear description are not derived from a proper scalar state function : such a function should be expanded in terms of the total deformations and of a rotationnally invariant electric vector while energy functions of the linear analysis are just expanded in terms of the dynamic deformations from the initial state (current temperature, no vibrations) and of the electric field, the material constants being *a posteriori* derived with respect to temperature. In a purely thermoelastic case, a previous study (7) computed the consistent temperature derivatives of the fundamental elastic constants, i.e. the constants that appear in the expansion of the scalar state function versus the total strain tensor, from the natural state to the actual one. Those derivatives may be used for the description of non homogeneous thermal bias, but an expansion to high order in terms of temperature is far out of reach because neither the thermal sensitivities of the non linear elastic moduli C_{IJK} nor the fourth order elastic modulu C_{IJKL} are presently known. For this reason, we preferred to build the present description on the use of the effective material constants which are the coefficients of the expansion of the dynamic part of Piola-Kirchhoff stress tensor versus the dynamic part of the total strain and of the rotationnally invariant electric vector. This approach forbids any use of the here-presented results to describe problems in which the bias is not homogeneous, but it reversely allows a high order expansion in terms of temperature. The effective material constants are closely related to the second order partial derivatives of the scalar state function with respect to the independent variables, those derivatives being evaluated around the initial state. The numerical values are obtained from the set of experimental data of (1), and

they are used for the description of the static thermal behavior of some quasi-thickness modes in contoured quartz resonators. It should be noted that the analysis might easily take into account the thermal sensitivities of piezoelectric and dielectric material constants.

II - BASIC EQUATIONS

In this paper $y_k(X_L, t)$ denotes a spatial coordinate of a material point referred to by its material coordinates X_L in a selected reference state. A partial differentiation with respect to a space coordinate is denoted by a comma followed by the index of this coordinate :

$$y_{k,L} = \partial y_k / \partial X_L \quad ; \quad X_{M,j} = \partial X_M / \partial y_j \qquad (1)$$

Only cartesian coordinates system are assumed in this paper where the tensor notation of summing over the repeated indexes is systematically used.

A rotationnaly invariant theory **(8, 2)** ensues from the statement of a thermodynamic potential $\mathcal{X}$ namely electric Gibbs function, dependent on the elastic strain tensor E_{LM}, on the rotationnaly invariant electric vector W_L and on the absolute temperature :

$$\mathcal{X} = \mathcal{X}(E_{LM} ; W_L ; \theta) \qquad (2)$$

where :

$$\left.\begin{aligned} E_{LM} &= (y_{k,L}\, y_{k,M} - \delta_{LM}) \\ W_L &= y_{k,L}\, E^M_k \end{aligned}\right\} \qquad (3)$$

E^M_k being a component of Maxwell electric field. The constitutive relations express the dependent variables, i.e. the mechanical stress τ_{ij}, the electric displacement D_i and the entropy per mass unit η in terms of partial derivatives of $\mathcal{X}$ with respect to the independent variables E_{LM}, W_L and Θ **(2)**. The stress equation of motion and the charge equation of electrostatics may be written in the usual form :

$$\left.\begin{aligned} (\tau_{ij} + T^E_{ij})_{,i} &= \rho \frac{dv_j}{dt} \\ D_{i,i} &= 0 \end{aligned}\right\} \qquad (4)$$

where T^E_{ij} is the electrostatic Maxwell tensor

$$T^E_{ij} = D_i E_j - \frac{1}{2}\, \varepsilon_0 E^M_k E^M_k \delta_{ij} \qquad (5)$$

and ρ denotes the mass density in the actual deformed state. Since all quantities in (4) are functions of the spatial coordinates y_k, the boundary conditions must explicitely be stated at the actual (unknown) position of the material surface. As explained in a previous paper **(6)**, it is largely more convenient, and even necessary, to express the boundary conditions at the material surface in its reference (known) position. To do this, it becomes necessary to replace the above-mentionned dependent variables by corresponding others, which must be functions of the known material coordinates X_L. Thus, we use the first Piola-Kirchhoff tensor $K_{Lj}(X_L)$ which may be expressed by :

$$K_{Lj} = J\, X_{L,i} (\tau_{ij} + T^E_{ij}) \qquad (6)$$

where $J = \det |y_{j,L}| = \rho_0 / \rho$, ρ_0 denoting the mass density in the reference configuration.

Substituting the proper constitutive relation for stress **(2, 3)** and using the chain rule of differenciation $X_{L,i}\, y_{i,M} = \delta_{LM}$ **(9)**, a general expression of K_{Lj} may easily be obtained **(10)** :

$$K_{Lj} = \rho_0 y_{j,M} \frac{\partial \mathcal{X}}{\partial E_{LM}} + J\, X_{L,i} \varepsilon_0 (X_{M,i} X_{N,j} - \frac{1}{2}\, X_{M,k} X_{N,k} \delta_{ij}) W_M W_N \qquad (7)$$

A material electric displacement vector $\mathcal{D}_L$ may be defined in the same manner :

$$\mathcal{D}_L = J\, X_{L,i} D_i \qquad (8)$$

The general expression of $\mathcal{D}_L$ is the following :

$$\mathcal{D}_L = \varepsilon_0 J X_{L,i} X_{M,i} W_M - \rho_0 \frac{\partial \mathcal{X}}{\partial W_L} \qquad (9)$$

With the new quantities defined by (6) and (8), the stress equation of motion and charge equation of electrostatics may be written :

$$\left.\begin{aligned} K_{Lj,L} &= \rho_0 \frac{\partial y_j}{\partial t^2} \\ \mathcal{D}_{L,L} &= 0 \end{aligned}\right\} \qquad (10)$$

while the boundary conditions at the interface between the studied piezoelectric material and a perfectly conductive coating are **(11)**:

$$\left.\begin{aligned} N_L (K'_{Lj} - K_{Lj}) &= 0 \\ N_L \mathcal{D}_L &= \sigma_0 \end{aligned}\right\} \text{ on } S_0 \qquad (11)$$

where S_0 is the material surface in the reference state, N_L denotes a component of the unit normal, outwardly oriented from the piezoelectric material, σ_0 is the charge density at the interface, mapped on the X_L coordinates and K'_{Lj} the Piola-Kirchhoff tensor for the conductor coating.

III - HOMOGENEOUSLY HEATED BODY

This section deals with the propagation of an elastic wave in a piezoelectric medium subject to a static deformation arising from an uniform temperature change. As usual in this kind of problem, we distinguish three states : a natural state or reference state at a fixed temperature θ_0, an intermediate state, without vibration and at a current temperature θ, this state being generally called initial state, although this is a rather confusing name, and a final state which is the actual state of the resonator, when an adiabatic high frequency vibration is superimposed to the initial state. We use the convention appearing in **(3)**, and denoting the natural coordinates by uppercase latin indexes, while greek indexes stand for intermediate coordinates and

lowercase latin indexes stand for final coordinates : for instance X_M, ξ_α and y_j are respectively some material coordinates in the natural, initial and final state.

The mechanical displacement $u_\alpha(\xi_\alpha, t)$ arising from the vibration is defined by :

$$y_j = (\xi_\alpha + u_\alpha)\,\delta_{j\alpha} \qquad (12)$$

where the Kronecker translator $\delta_{j\alpha}$ is introduced to manage the indexes correspondences from one state to another. The total deformation defined in Eq. (3) may be separated into a static deformation $\bar{E}_{LM}$ plus a dynamic strain increment $\tilde{E}_{LM}$:

$$\left.\begin{aligned} E_{LM} &= \bar{E}_{LM} + \tilde{E}_{LM} \\ \bar{E}_{LM} &= \frac{1}{2}(\xi_{\alpha,L}\,\xi_{\alpha,M} - \delta_{LM}) \\ \tilde{E}_{LM} &= \frac{1}{2}(\xi_{\alpha,L}\,u_{\alpha,M} + \xi_{\alpha,M}\,u_{\alpha,L} + u_{\alpha,L}\,u_{\alpha,M}) \end{aligned}\right\} \qquad (13)$$

Eq (13-c) has been obtained using the chain rule of differentiation :

$$y_{j,M} = y_{j,\alpha}\,\xi_{\alpha,M} = \xi_{\alpha,M}\,\delta_{j\alpha} + u_{\alpha,M}\,\delta_{j\alpha} \qquad (14)$$

The dynamic Piola-Kirchhoff tensor $\tilde{K}_{Lj}$ is similarly defined by (3) :

$$K_{Lj} = \bar{K}_{Lj} + \tilde{K}_{Lj}(u_\alpha ; \tilde{W}_L) \qquad (15)$$

where $\bar{K}_{Lj}$ is the static Piola-Kirchhoff tensor and $\tilde{W}_L$ is the increase of the electric vector from the initial to the final state. If the electric field is zero and the temperature fied uniform in the initial state, if the material is not pyroelectric and if the thermal expansion may occur without external constraint, we have :

$$\bar{K}_{Lj} = 0 \;\; ; \;\; \bar{W}_L = 0 \;\; ; \;\; \bar{\mathcal{D}}_L = 0 \qquad (16)$$

In addition, since we are only interested with infinitesimal dynamic fields, we may use the following Taylor's expansions of the partial derivatives of χ :

$$\left.\begin{aligned} \frac{\partial\chi}{\partial E_{LM}} &= \frac{\partial\chi}{\partial E_{LM}}\Big]_1 + \frac{\partial^2\chi}{\partial E_{LM}\partial E_{NP}}\Big]_1\tilde{E}_{NP} + \frac{\partial^2\chi}{\partial E_{LM}\partial W_K}\Big]_1\tilde{W}_K + \dots \\ \frac{\partial\chi}{\partial W_K} &= \frac{\partial\chi}{\partial W_K}\Big]_1 + \frac{\partial^2\chi}{\partial W_K\partial E_{LM}}\Big]_1\tilde{E}_{LM} + \frac{\partial^2\chi}{\partial W_K\partial W_L}\Big]_1\tilde{W}_L + \dots \end{aligned}\right\} \qquad (17)$$

where a subscript 1 neighbouring a bracket indicates that the partial derivative is evaluated in the initial state.

Substituting Eqs (17-a), (14) and (13-a) into (7), and using the chain rule of differentiation in the reverse way, i.e. :

$$X_{M,i} = X_{M,\alpha}\,\xi_{\alpha,i} = X_{M,\alpha}(\delta_{i\alpha} - u_{\alpha,\varepsilon}\,\xi_{\varepsilon,i} + \dots) \qquad (18)$$

we obtain the following expression for $\tilde{K}_{Lj}$, where only linear terms in dynamic quantities have been retained

$$\tilde{K}_{Lj} = \rho_0\delta_{j\alpha}\frac{\partial\chi}{\partial E_{LM}}\Big]_1 u_{\alpha,M} + \rho_0\delta_{j\alpha}\xi_{\alpha,M}\left[\frac{\partial^2\chi}{\partial E_{LM}\partial E_{NP}}\Big]_1\tilde{E}_{NP} + \frac{\partial^2\chi}{\partial E_{LM}\partial W_K}\Big]_1\tilde{W}_K\right] \qquad (19)$$

Because the initial state is completely stress free and according to Eqs (16) and (7), the partial derivatives $\partial\chi/\partial E_{LM})_1$ identically vanish and, with help of Eq (13-c), the linearized expression for the dynamic Piola tensor eventually becomes :

$$\tilde{K}_{Lj} = \rho_0\delta_{j\alpha}\,\xi_{\alpha,M}\,\xi_{\beta,N}\frac{\partial^2\chi}{\partial E_{LM}\partial E_{NP}}\Big]_1 u_{\beta,P} + \rho_0\delta_{j\alpha}\xi_{\alpha,M}\frac{\partial^2\chi}{\partial E_{LM}\partial W_K}\Big]_1\tilde{W}_K + \dots \qquad (20)$$

First defining the dynamic material electric displacement $\tilde{\mathcal{D}}_L = \mathcal{D}_L - \bar{\mathcal{D}}_L$, thereafeter substituting (17-b) into (9), using (18) and (13-c), the corresponding linearized expression for $\tilde{\mathcal{D}}_L$ is obtained :

$$\tilde{\mathcal{D}}_L = \left[\varepsilon_0\bar{J}X_{L,\alpha}X_{M,\alpha} - \rho_0\frac{\partial^2\chi}{\partial W_L\partial W_M}\Big]_1\right]\tilde{W}_M - \rho_0\frac{\partial^2\chi}{\partial W_L\partial E_{MN}}\Big]_1\xi_{\alpha,M}u_{\alpha,N} + \dots \qquad (21)$$

where $\bar{J} = \det|\xi_{\alpha,M}|$. This equation is slightly more general than (20), since it does not necessarily assume the bias being homogeneous and no use of Eqs (16) has been made on obtaining it. We are now able to write Eqs (20) and (21) in the more concise form :

$$\left.\begin{aligned} \tilde{K}_{Lj} &= (G_{L\gamma M\varepsilon}\,u_{\varepsilon,M} - R_{M,\gamma L}\,\tilde{W}_M)\,\delta_{j\gamma} \\ \tilde{\mathcal{D}}_L &= R_{LM}\,\tilde{W}_M + R_{L,\alpha N}\,u_{\alpha,N} \end{aligned}\right\} \qquad (22)$$

The coefficients appearing in these relations are effective material constants which are explicitely dependent on the second order partial derivatives of the scalar state function χ with respect to the strain tensor and the rotationnaly invariant electric vector. We label these derivatives in the following manner :

$$\left.\begin{aligned} \chi_{KL} &= \rho_0\frac{\partial^2\chi}{\partial W_K\partial W_L}\Big]_1 \\ \chi_{K,LM} &= \rho_0\frac{\partial^2\chi}{\partial W_K\partial E_{LM}}\Big]_1 \\ \chi_{KLMN} &= \rho_0\frac{\partial^2\chi}{\partial E_{KL}\partial E_{MN}}\Big]_1 \end{aligned}\right\} \qquad (23)$$

Thus, the effective elastic coefficients $G_{L\gamma M\varepsilon}$, effective piezoelectric constants $R_{L,YN}$ and effective permittivities R_{LM} are given by :

$$\left.\begin{aligned} G_{L\gamma M\varepsilon} &= \xi_{\gamma,K}\,\xi_{\varepsilon,N}\,\chi_{KLMN} \\ R_{M,\gamma L} &= -\xi_{\gamma,K}\,\chi_{MKL} \\ R_{LM} &= \varepsilon_0\,\bar{J}\,\chi_{L,\alpha}\,\chi_{M,\alpha} - \chi_{LM} \end{aligned}\right\} \qquad (24)$$

The stress equation of motion and the charge equation of electrostatics take the following forms :

$$\left.\begin{aligned} \tilde{K}_{Lj,L} &= \rho_0\,\frac{\partial^2 u}{\partial t^2}\,\delta_{j\alpha} \\ \tilde{\mathcal{D}}_{L,L} &= 0 \end{aligned}\right\} \qquad (25)$$

When the studied piezoelectric material is coated with conductive material, the boundary conditions at the separation surface take the form :

$$\left.\begin{aligned} N_L(\tilde{K}'_{Lj} - \tilde{K}_{Lj}) &= 0 \\ N_L\,\tilde{\mathcal{D}}_L &= \tilde{\sigma}_0 \end{aligned}\right\} \text{ on } S_0 \qquad (26)$$

where $\tilde{\sigma}_0$ is the dynamic increment of the surface charge density mapped on S_0. When the Cartesian coordinates system coincides with the crystallographic axes, and when the static strain arises from the stress free thermal expansion, the derivatives $\xi_{\alpha,K}$ appearing in relations (24) are over the entire body and they are given by :

$$\xi_{\alpha,K} = \delta_{\alpha K}\left(1 + \alpha^{(n)}_{KK}(\theta - \theta_0)^{(n)}\right) \qquad (27)$$

in which $\alpha^{(n)}_{KK}$ is the n-th order thermal expansion coefficient.

IV - THERMAL DERIVATIVES OF EFFECTIVE STIFFNESSES OF α-QUARTZ

In the followings, the n-th order thermal coefficient of any quantity Q is defined by :

$$T^{(n)}Q = \frac{1}{n!}\,\frac{\partial^n Q}{Q\,\partial\theta^n} \qquad (28)$$

Since the purpose of this paper is not a determination of material constants from specific new measurements, we restrict the use of the above-presented formalism to the treatment of the very well known set of experimental data obtained by Bechmann, Ballato and Lukaszek (1). We assume that those data may correctly be described by the infinite plate unidimensinal model. To compute the $T^{(n)}G_{L\gamma M\varepsilon}$ from the temperature coefficients of frequency $T^{(n)}f$ published in (1), we neglect any thermal dependency of the piezoelectric coefficients $R_{M,\gamma L}$ and permittivity coefficients R_{LM}, such a dependency remaining within the accuracy of the measurements and, above all, the vibration model. Nevertheless, it should be pointed out that this restriction does not arise from the here-presented formalism itself.

The difference between the present treatment and the initial one in (1) lies in the nature of the determined coefficients, but the calculation technique is altogether the same : first using the results provided by X, Y, AT and BT cuts to directly get the thermal derivatives of four independent coefficients, thereafter determining the remainder of them by a least squares fit of a large set of data provided by doubly rotated cuts. Such a method was also previously used by Tiersten and Sinha (7) to compute the thermal derivatives of the fundamental elastic constants i.e. those appearing in the expansion of the scalar state function versus the components of the total strain tensor. The effective material constants of our interest are just the coefficients of expansion of $\tilde{K}_{Lj}$ in terms of the dynamic fields $u_{\alpha,M}$ and $\tilde{W}_L$. It should be observed that, despite of their different meaning, both sets of constants are closely dependent, through the thermal expansion coefficients and high order elastic constants. However the corresponding dependence relations of thermal derivatives cannot be numerically obtained beyond the first order in terms of temperature, because those relations involve the fourth order elastic moduli C_{IJKL} and the thermal sensitivities of the third-order elastic moduli C_{IJK}, the values of which are not presently known.

It should also be clear that choosing $G_{L\gamma M\varepsilon}$ as unknown, instead of χ_{KLMN} seems rather arbitrary. A practical advantage of this choice lies in the fact it facilitates the comparison with previously obtained results of Ref (6), where the influence of piezoelectricity was completely ignored, and the thermal derivatives of effective stiffnesses $T^{(n)}G_{L\gamma M\varepsilon}$ were directly obtained from the $T^{(n)}C_{ijk\ell}$ provided at Ref (1). The choice of the $G_{L\gamma M\varepsilon}$ rather than χ_{KLMN} introduce a small complication : because of anisotropy of thermal expansion, the G coefficients for the crystalline class 32 of α-quartz, take ten distinct values, paying no attention to the sign, when the temperature differs from the reference temperature (25°C in this paper), whereas the χ coefficients keep only seven distinct values at any value of temperature. Hence, we need to eliminate the dependent coefficients in terms of the undependent ones. To reduce the number of subscripts, we use the classical correspondence law :

$$\left.\begin{aligned} &11 \leftrightarrow 1 \quad 22 \leftrightarrow 2 \quad 33 \leftrightarrow 3 \\ &23 \leftrightarrow 4 \quad 31 \leftrightarrow 5 \quad 12 \leftrightarrow 6 \\ &32 \leftrightarrow 7 \quad 13 \leftrightarrow 8 \quad 21 \leftrightarrow 9 \end{aligned}\right\} \qquad (29)$$

Then, we have chosen G_{11}, G_{13}, G_{14}, G_{33}, G_{44} as independent coefficients, considering that G_{17} depends on G_{14}, and G_{47} and G_{55} depend on G_{44}. The dependence relations are given in appendix. The complete structure of the G matrix in the crystallographic axes is shown on table I. Since the G coefficients directly appear in the wave propagation equations and their associated boundary conditions, using them remove from the user any handling of thermal expansion, while using the derivatives of χ would keep this care on the user (12).

Since the thermal sensitivity of the piezoelectric $R_{L\,.\,YM}$ and dielectric R_{LM} material effective coefficients is ignored in the present analysis, we use the classical conventional notations e_{ij} and ε_{ij} (lower case 1 to 3, uppercase 1 to 6) : their values at the reference temperature are indeed the corresponding values in the linear theory of piezoelectricity. In this way the secular equation for pure thickness modes takes the following shape, expressed in the doubly rotated reference coordinate axes of the thin plate, in any orientation :

$$\begin{bmatrix} G_{99}+\frac{e_{26}^2}{\varepsilon_{22}} & G_{29}+\frac{e_{22}e_{26}}{\varepsilon_{22}} & G_{49}+\frac{e_{24}e_{26}}{\varepsilon_{22}} \\ G_{29}+\frac{e_{22}e_{26}}{\varepsilon_{22}} & G_{22}+\frac{(e_{22})^2}{\varepsilon_{22}} & G_{24}+\frac{e_{22}e_{24}}{\varepsilon_{22}} \\ G_{49}+\frac{e_{24}e_{26}}{\varepsilon_{22}} & G_{24}+\frac{e_{22}e_{24}}{\varepsilon_{22}} & G_{44}+\frac{(e_{24})^2}{\varepsilon_{22}} \end{bmatrix} \begin{bmatrix} u_1^\alpha \\ u_2^\alpha \\ u_3^\alpha \end{bmatrix} = \rho_0 V_\alpha^2 \begin{bmatrix} u_1^\alpha \\ u_2^\alpha \\ u_3^\alpha \end{bmatrix} \qquad (30)$$

where V_α is the celerity of the α-th mode (A, B or C). Since the G coefficients allow the use of the material coordinates at the reference temperature to describe the wave propagation problem and its associated boundary conditions, we do not have to include the temperature derivatives of ρ in this equation, and, because of the above mentionned assumption, only the G coefficients are temperature dependent, thus governing the thermal sensitivity of Christoffel's velocities V_α and eigensolutions u_i^α. In addition, since the orientation angles of the plate are defined at the fixed reference temperature, using the here described formalism avoids any handling of the orientation change of the plate normal in the crystallographic axes, which results from anisotropic thermal expansion. This advantage of the use of the reference coordinates has been previously explained at Ref (5), where the authors use the temperature derivatives of the fundamental constants. An evaluation of the effect of the generally ignored thermal shift of the plate normal on the results provided by using the final coordinates has been given at Ref (6). In that paper, the use of the material effective constants and their thermal derivatives was already suggested, although in a purely thermoelastic case (see also Ref (13)). If the overtone number n, the plate thickness t_0, the electrode plating thickness t'_0 and mass density ρ'_0 at the fixed reference temperature are known, a very accurate expression for the frequency of an infinite plate is given by :

$$\left.\begin{aligned} \omega_n^2 &= \frac{n^2\pi^2}{t_0^2} V_\alpha^2 \left(1 - 8\,k_\alpha^2/n^2\pi^2 - 2\,\hat{R} \right) \\ \hat{R} &= \rho' t'_0/\rho_0 t_0 \\ k_\alpha^2 &= (u_j^\alpha\, e_{2.2j})^2/\rho_0 V_\alpha^2\, \varepsilon_{22} \end{aligned}\right\} \qquad (31)$$

Since we do not know the actual value of $\hat{R}$ for the experimental data of (1), we have to neglect the corresponding corrective term in (31). It is then reasonable to also ignore the corrective term arising from k_α^2, although Bechmann, Ballato and Lukaszek explained in (1) that they used only fundamental thickness modes, thus allowing us to perform the correction provided by the term $1 - 8k_\alpha^2/n^2\pi^2$: this correction would be of little effect, regarding the poor accuracy of the pure thickness modes model. Then all we did is consistent with the use of the very simple formula :

$$f_n^\alpha \approx n\, V_\alpha / 2t_0 \qquad (32)$$

for the n-th overtone of the α-th mode.
Although this formula looks very like the usual one, it involves some major differences :

- t_0 is the thickness at the fixed reference temperature. So, one has not to adjust it when the temperature is varying. The same remark holds for the mass density ρ_0 and for the coefficients of the tensor rotation from the crystallographic axes to the doubly rotated axes,

- according to Eq (24), the G_{IJ} coefficients are defined at constant material electric vector $\tilde{W}$. At the reference temperature, they coincide with the C_{IJ} of linear Piezoelectricity, largely used in vibration models, whereas the tabulated values of Ref (1) must be understood for elastic stiffnesses at constant normal electric displacement.

The experimental values of the temperature coefficients of frequency used in the present paper, obtained from (1), are reproduced at table II. To get the $T^{(n)}G_{IJ}$, we numerically computed the following quantities for each pure mode of this table :

$$\left.\begin{aligned} M_{IJ} &= (\partial f/\partial G_{IJ})(G_{IJ}/f) \\ N_{IJKL} &= (\partial^2 f/\partial G_{IJ}\partial G_{KL})(G_{IJ}G_{KL}/f) \end{aligned}\right\} \qquad (33)$$

without sum on the repeated couples of indexes.

The partial derivatives are computed with help of a three points interpolation. The G_{IJ} increment for this interpolation was set to $10^{-5} \times C_{66}$ at 25°C. According to the relatively linear behavior of the solution of Christoffel's problem, we did not judge necessary to get the cross derivatives up to the third order. Then, the temperature coefficients of the frequency are related to $T^{(n)}G_{IJ}$ by the following equations :

$$\left.\begin{aligned} T^{(1)}f &= M_{IJ}.T^{(1)}G_{IJ} \\ T^{(2)}f &= M_{IJ}.T^{(2)}G_{IJ} + \frac{1}{2} N_{IJ\,KL}.T^{(1)}G_{IJ}.T^{(1)}G_{KL} \\ T^{(3)}f &= M_{IJ}.T^{(3)}G_{IJ} + N_{IJ\,KL}.T^{(2)}G_{IJ}.T^{(1)}G_{KL} \end{aligned}\right\} \qquad (34)$$

where the repeated couples of indexes for the undependent coefficients, i.e., (11), (13), (14), (33), (44), (66) are summed. The calculation of M and

N for each mode takes a much longer time than the solution of the linear systems itself. As above mentionned in this section, the $T^{n}G_{13}$ and $T^{n}G_{33}$ were computed by a least square fit, after the obtention of the other undependent coefficients from the particular cuts X, Y, AT, BT. We provide at Table III the obtained values for the $T^{(n)}G_{IJ}$. The $T^{(n)}G_{12}$ coefficients may be easily deduced from :

$$G_{12} = G_{11} - 2\,G_{66} \qquad (35)$$

From here-presented values, the following sets of temperature derivatives may be obtained :

- the first order derivatives of the fundamental elastic constants, according to Eqs (2-2)(2-3) of **(7)**,
- all the derivatives of the χ_{KLMN} coefficients, according to Eqs (24) and (27) of this paper,
- all the derivatives of the effective constants referred to the coordinates of the intermediate state, according to Eq (28) of **(6)**, in which $A_{ijk\ell}$ holds for an effective coefficient referred to the coordinates of the natural state, while $B_{ijk\ell}$ is referred to the coordinate of the initial state.

The last set of coefficients is altogether identical to the usual ones used in linear-only analysis, provided the orientation change due to anisotropy of thermal expansion is correctly taken into account. The effective coefficients referred to the coordinates of the initial state also appear in Eq (77) of **(3)**. In the case studied in the present paper, those coefficients reduce to :

$$H_{\alpha\beta\gamma\varepsilon} = \frac{\rho_0}{J}\,\xi_{\alpha,K}\,\xi_{\beta,L}\,\xi_{\gamma,M}\,\xi_{\varepsilon,N}\left(\frac{\partial^2\chi}{\partial E_{KL}\,\partial E_{MN}}\right)_1 \qquad (36)$$

Using those coefficients should not be recommended since to keep a satisfactory agreement with the general laws of thermoelectroelasticity, they involve a very cumbersome handling of thermal expansion in the wave motion equations and the associated boundary conditions. Furthermore, this handling becomes a real problem in case of resonators exhibiting a non parallelepipedic form, like contoured resonators.

V - STATIC THERMAL BEHAVIOR OF CONTOURED RESONATORS

The trapped modes in contoured resonators have been satisfactorily predicted in an asymptotic model due to Tiersten and Stevens **(14)**. This model gives an analytical expression of the eigenfrequencies :

$$\omega_{nmp} = \omega_n^2\left[1 + \frac{1}{n\pi}\left(\frac{t_0}{R_0}\right)^{1/2}\left(\sqrt{\frac{M'_n}{C\alpha}}\,(2m+1) + \sqrt{\frac{P'_n}{C\alpha}}\,(2p+1)\right)\right] \qquad (37)$$

where ω_n, t_0 and n have been defined in the previous section of this paper, C_α is the α-th eigensolution of Christoffel's problem, Eq (30), multiplied by the corrective factor $(1 - 8k_\infty^2/n^2\pi^2 - 2\hat{R})$ of Eq (31-a), R_0 denotes the radius of contour of the plate at the reference temperature (25°C), m and p are integer numbers counting the number of nodal lines in two orthogonal directions of the plane of the plate : they are zero for the so-called metrologic modes used for high quality and stability applications. M'_n and P'_n are closely related to the curvatures of the proper branches of the dispersion curves at small plate wavenumbers. M'_n and P'_n express in terms of special elastic coefficients C^T_{IJKL} through Eqs (74) and (100) of **(14)**. To keep consistency with the present analysis, those coefficients may be derived from the $G_{L\gamma M\varepsilon}$ coefficients in the rotated plate axes through the following linear transformation :

$$\left.\begin{aligned} C^T_{LPRM} &= Q_{Pj}\,Q_{R\ell}\,G_{L\gamma M\varepsilon}\,\delta_{j\gamma}\,\delta_{\ell\varepsilon} \\ Q_{ij} &= \delta_{i\alpha}\,u^{\alpha}_{j} \end{aligned}\right\} \qquad (38)$$

where u^{α}_{j} is a normalized eigensolution of eigenproblem of Eq (30). The linear transformation law is different from Eq (23) of Ref **(14)**, because the G coefficients have less symmetry properties than the usual C referred to in Ref **(14)**.

Figs (1) and (2) present some frequency temperature curves computed with the results of the above presented analysis, and respectively corresponding to C modes in contoured AT and SC cuts. The curves have been obtained for the limit case of an electrode plating involving no effect on the frequency ($\hat{R} = 0$). This case is the same as the one of a resonator with non adherent electrodes when the air gap tends to zero. From our experience on actual BVA resonators, the here-presented frequency temperature curves correctly predict the turn-over points provided a correction of a few minutes should been made on the Θ orientation angle. However, the predicted frequency difference between the upper and the lower turn-over points may be slightly exaggerated. Other predicted results using the vibration model of Ref **(15)**, and not available on time for this printing, are expected to confirm this limitation of accuracy for here obtained $T^{(n)}G_{IJ}$ values. Incidently, the large standard errors obtained for $T^{(n)}G_{13}$ and $T^{(n)}G_{33}$ predict this limitation of accuracy. Thus, expected new measurements followed by a treatment consistent with thermoelectroelasticity equations and systematically using a more accurate vibration model than the unidimensionnal model seem to be the obliged but fruitful way to improve the description accuracy of the static thermal behavior of high frequency quartz resonator.

Appendix

We give here the dependence relations of $T^{(n)}G_{17}$, $T^{(n)}G_{47}$ and $T^{(n)}G_{55}$ versus $T^{(n)}G_{14}$ and $T^{(n)}G_{44}$, in the crystallographic axes :

$$\left.\begin{aligned} T^{(1)}G_{17} &= T^{(1)}G_{14} + \alpha^{(1)}_{11} - \alpha^{(1)}_{33} \\ T^{(1)}G_{47} &= T^{(1)}G_{44} + \alpha^{(1)}_{11} - \alpha^{(1)}_{33} \\ T^{(1)}G_{55} &= T^{(1)}G_{44} + 2(\alpha^{(1)}_{11} - \alpha^{(1)}_{33}) \end{aligned}\right\} \qquad (A\text{-}1)$$

$$\left.\begin{aligned} T^{(2)}G_{17} &= T^{(2)}G_{14} + \alpha^{(1)}_{11}T^{(1)}G_{14} - \alpha^{(1)}_{33}T^{(1)}G_{17} + \alpha^{(2)}_{11} - \alpha^{(2)}_{33} \\ T^{(2)}G_{47} &= T^{(2)}G_{44} + \alpha^{(1)}_{11}T^{(1)}G_{44} - \alpha^{(1)}_{33}T^{(1)}G_{47} + \alpha^{(2)}_{11} - \alpha^{(2)}_{33} \\ T^{(2)}G_{55} &= T^{(2)}G_{44} + 2\alpha^{(1)}_{11}T^{(1)}G_{44} - 2\alpha^{(1)}_{33}T^{(1)}G_{55} + 2(\alpha^{(2)}_{11} - \alpha^{(2)}_{33}) \\ &\quad + (\alpha^{(1)}_{11})^2 - (\alpha^{(1)}_{33})^2 \end{aligned}\right\}$$

(A-2)

$$T^{(3)}G_{17} = T^{(3)}G_{14} + \alpha_{11}^{(1)}T^{(2)}G_{14} - \alpha_{33}^{(1)}T^{(2)}G_{17}$$

$$+ \alpha_{11}^{(2)}T^{(1)}G_{14} - \alpha_{33}^{(2)}T^{(1)}G_{17} + \alpha_{11}^{(3)} - \alpha_{33}^{(3)}$$

$$T^{(3)}G_{47} = T^{(3)}G_{44} + \alpha_{11}^{(1)}T^{(2)}G_{44} - \alpha_{33}^{(1)}T^{(2)}G_{47}$$

$$+ \alpha_{11}^{(2)}T^{(1)}G_{44} - \alpha_{33}^{(2)}T^{(1)}G_{47} + \alpha_{11}^{(3)} - \alpha_{33}^{(3)}$$

$$T^{(3)}G_{55} = T^{(3)}G_{44} + 2\alpha_{11}^{(1)}T^{(2)}G_{44} - 2\alpha_{33}^{(1)}T^{(2)}G_{55}$$

$$[2\alpha_{11}^{(2)}-(\alpha_{11}^{(1)})^2]T^{(1)}G_{44} - [2\alpha_{33}^{(2)}-(\alpha_{33}^{(1)})^2]T^{(1)}G_{55}$$

$$+ 2[\alpha_{11}^{(3)}-\alpha_{33}^{(3)} + \alpha_{11}^{(1)}\alpha_{11}^{(2)} - \alpha_{33}^{(1)}\alpha_{33}^{(2)}] \qquad (A\text{-}3)$$

NOTES AND REFERENCES

(1) R. BECHMANN, A. BALLATO, T.J. LUKASZEK : "High Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of quartz", P.I.R.E. (1962) p 1812.

(2) H.F. TIERSTEN : "On the Non Linear Equations of Thermo Electro Elasticity", Int. J. Engng Sci., Vol 9, pp 587-604, Pergamon Press (1971).

(3) J.C. BAUMHAUER, H.F. TIERSTEN : "Non Linear Electroelastic Equations for Small Fields Superposed on a Bias", J. Ac. Soc. Am. 54, p 1017 (1973).

(4) H.F. TIERSTEN : "Perturbation Theory for Linear Electroelastic Equations for Small Fields superposed on a Bias", J. Ac. Soc. Am. 64 (3), p 832 (1978).

(5) B.K. SINHA, H.F. TIERSTEN : "On the Temperature Dependence of the Velocity of Surface Waves in Quartz", J. Appl. Phys. 51 (9), p 4659 (1980).

(6) B. DULMET, R. BOURQUIN : "Méthodes de détermination des courbes fréquence température d'un résonateur à quartz vibrant en mode d'épaisseur", Rev. Phys. Appl. 18, p 619 (1983).

(7) B.K. SINHA, H.F. TIERSTEN : "First Temperature Derivatives of the Fundamental Elastic Constants of Quartz", J. Appl. Phys. 50 (4), p 2732 (1979).

(8) R.A. TOUPIN : "The Elastic Dielectric", J. Ration. Mech. Anal. 5, 849 (1956).

(9) R.N. THURSTON : "Waves in Solids", Handbuch der Physik VI a/4 Springer Verlag, Berlin (1974).

(10) H.F. TIERSTEN : "Lectures on thermoelectroelasticity" given at ENSMM, Mach 1988.

(11) H.F. TIERSTEN : "Non Linear Electroelastic Equations cubic in the small field variables", J. Ac. Soc. Am. 57 (3), p 660 (1975).

(12) But it should be known that assuming $R_{L,M}$ and R_{LM} unaffected by temperature does lead to other values for $T^{(n)}G_{L\ell M\ell}$ than assuming χ_{KL} and χ_{KLM} unaffected by temperature.

(13) B. DULMET, F. FICHET : "Couplings of thickness vibrations in Contoured Resonators and their Effect on frequency Spectrum and frequency Temperature Behavior" Ultrasonics Symposium, IEEE, p 388 (1984).

(14) D.S. STEVENS, H.F. TIERSTEN : "An Analysis of doubly rotated quartz resonators utilizing essentially thickness modes with transverse variation", J. Ac. Soc. Am. 79 (6), p 1811 (1986).

(15) B. DULMET : "Application d'une méthode de perturbation à l'étude de résonateurs à quartz présentant des modes d'épaisseur à énergie piégée", Rev. Phys. Appl. 19 (1984) p 839.

	1	2	3	4	5	6	7	8	9
1	G_{11}	G_{12}	G_{13}	G_{14}	0	0	G_{17}	0	0
2		G_{11}	G_{13}	$-G_{14}$	0	0	$-G_{17}$	0	0
3			G_{33}	0	0	0	0	0	0
4				G_{44}	0	0	G_{47}	0	0
5					G_{55}	G_{17}	0	G_{47}	G_{17}
6		symmetry				G_{66}	0	G_{14}	G_{66}
7							G_{55}	0	0
8		$G_{IJ} = G_{JI}$						G_{44}	G_{14}
9									G_{66}

TABLE I

φ	θ	mode	T^1f 10^{-6}	T^2f 10^{-9}	T^3f 10^{-12}
30.0	00.00	A	-20.50	-53.20	-36.60
30.0	10.00	A	-29.30	-67.00	-58.30
30.0	20.00	A	-42.00	-86.00	-92.70
30.0	30.00	A	-54.80	-105.0	-119.0
00.0	-49.2166	B	0.00	-40.00	-128.0
05.0	-47.00	B	-0.90	-41.10	-118.0
05.0	-48.00	B	-1.50	-43.10	-122.6
10.0	-38.00	B	1.10	-39.00	-91.30
10.0	-40.00	B	-2.40	-41.90	-115.4
15.0	-34.50	B	-7.40	-28.10	-39.70
15.0	-35.00	B	-7.45	-33.50	-58.80
30.0	20.00	B	-9.30	-21.80	-34.40
30.0	30.00	B	-19.10	-24.50	-28.90
00.0	00.00	C	92.50	57.50	5.80
00.0	35.25	C	0.00	0.40	109.50
10.0	-32.00	C	0.26	-9.80	-32.40
10.0	-33.00	C	-0.87	-7.81	-21.50
12.5	-33.00	C	0.80	-7.60	-19.60
12.5	-33.50	C	-0.40	-7.40	-14.00
15.0	-34.50	C	1.30	-7.30	-6.60
15.0	-35.00	C	-0.20	-7.02	-2.60
20.0	34.3333	C	-0.06	-8.90	52.00
30.0	34.00	C	0.75	-12.96	17.40
30.0	36.00	C	-4.55	-14.30	20.60

TABLE II

T^nG_{IJ}	(10^{-6})	(10^{-9})	(10^{-12})
IJ	n = 1	n = 2	n = 3
11	-41.4	-106.9	-71.6
13	-755.7	-1131.5	-106.9
σT^nG_{13}	36.7	89.8	221.7
14	109.7	-44.3	-593.9
33	-215.2	-199.4	79.7
σT^nG_{33}	15.5	37.8	93.3
44	-170.8	-212.3	-215.7
66	188.5	125.9	23.5
12	$[C_{11}T^nG_{11} - 2\ C_{66}T^nG_{66}] / C_{12}$		
17	115.9	-40.1	-594.3
47	-164.6	-209.8	-218.1
55	-158.3	-207.3	-220.4

Table III

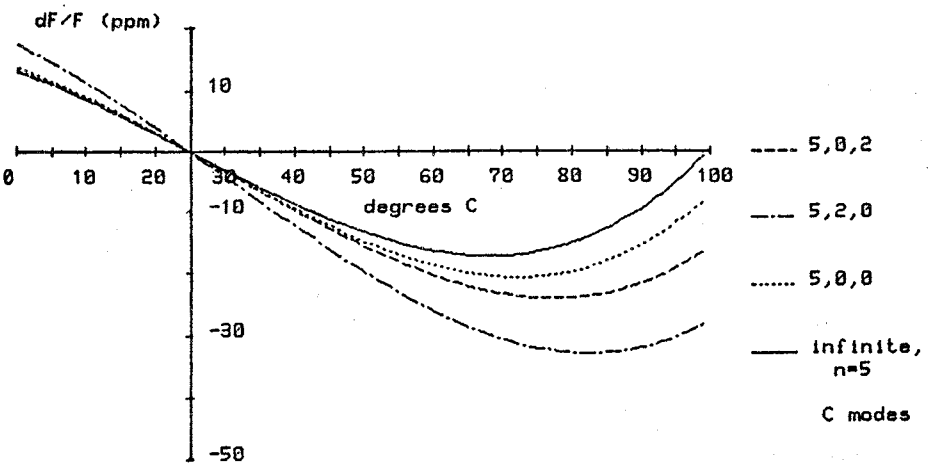


Fig.1 Phi =0d0' Theta = 35d21' thickness =1.679mm Rc= 150 mm.

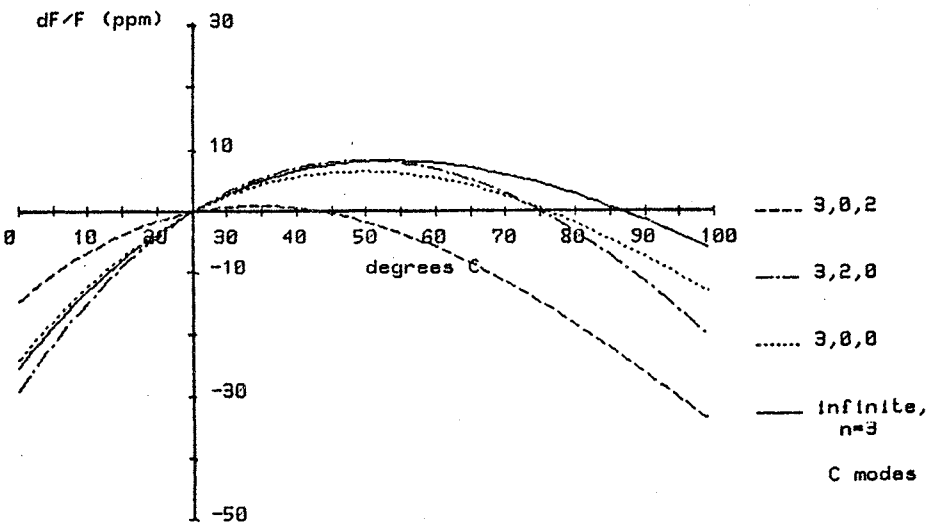


Fig.2 Phi =-22d45' Theta = 33d55' thickness =1.089mm Rc= 300 mm.

Loran-C Signal Synthesizer

Enrico Rubiola

Politecnico di Torino
Dipartimento di Elettronica

Abstract

The Loran-C system is still in use for time scales comparison and continuous frequency monitoring. Some improvements seems to be achieved by increasing the knowledge of the internal delay of the receiving equipments and its behaviour in various situations.

In this paper a synthesizer, designed for timing applications, is presented. The main feature is the high time stability, obtained with digital techniques. The complete pattern of a chain can be simulated setting time of arrival and amplitude individually for each station.

Criteria of design, bloch diagrams and some experimental results are given.

1 Introduction

Despite of the growing importance of the GPS system, Loran-C is still in use for time scales comparison and continuous monitoring. Loran-C offers some advantages for its higher availability as compared with other long range systems; an off-air period of a few hours per year can be assumed [1,2,3]. Moreover, continuous monitoring can be performed without switching the reference source, as happens using nonstationary satellites, like the GPS ones. The Loran-C system can be also used as a frequency reference driving a time scale [4].

Finally, in the hypothesis of of a full diffusion of satellite-based methods, Loran-C could be the best backup system.

The goal of this paper is to present a Loran-C synthesizer designed for time and frequency measurements. The main feature is the high time stability, that is obtained with broad band circuitry, whose total delay is intrinsically small.

This instrument is a tool of investigation in some experimental activities, such as:

- test of metrological and navigation receivers,
- measurement of the internal delay – and its stability – of metrological receivers,
- measurement of the effects of coherent interferences; in this case the simulator is to be used in conjunction with a continuous wave synthesizer, driven by the same frequency standard,
- metrological use of navigation receivers, using a self-calibration approach [5].

The instrument has the capability of generating the complete pattern of a chain, handling amplitude and delay separately for each station. In order to get a quite realistic situation when testing navigation receivers, noise can be added and the time delays can be swept, simulating a moving installation of the equipment.

A good description of the Loran-C system, covering many aspects and recent developments, can be found in [6]. All the details concerning the transmitted signal are widely described in [7].

2 Synthesis of the Loran-C pulse

The approach followed consists in generating the whole Loran-C pulse, comprehending both envelope and carrier, with a digital-to-analog converter; the instantaneous amplitude are mapped in a look-up table (read only memory) scanned at the sampling frequency. This method is adequate for generating high accuracy Loran-C pulses for the following reasons: first, independent synthesis of carrier and envelope, as done in some commercial or laboratory equipments, such as [8], does not allow to control their phase relationship as well as can be done with simultaneous synthesis; second, if the instrumental delay is small enough, variations or drifts are proportionally small, and consequently high phase stability can be easily achieved. The meaning of this concept will be more clear in the following sections.

If time and amplitude resolutions of the sampling performed by the converter are sufficient, there is no need of filtering the output of the pulse generator; only the suppression of the high frequency disturbances is useful. The final cleaning of the signal will be performed inside the receiver, always equipped with a filter, whose bandwidth is in the range from 20 to 40 kHz.

Fast digital-to-analog converters (DAC) are commercially available in a wide spread of accuracy and settling time in the range from 8 to 12 bits and 20 to 200 ns. Consequently a sampling frequency of 10 MHz seems to be a convenient practical value: the resolution of 100 ns is immediately achievable; being the sampling frequency two order of magnitude higher of the signal one, the quantization noise is negligible in the Loran-C band.

These considerations lead directly to the block diagram of the pulse generator, shown in fig. 1.

The normalized Loran-C pulse is described by

$$x_t = (k\,t)^2\, e^{-2(k\,t-1)} \sin \omega t \qquad \text{for } t \geq 0 \quad (1)$$

where $\omega = 2\pi 10^5$ rad/s and $1/k = 65\mu$s; the corresponding waveform is shown in fig. 2. The time measurements, according to [7], should be performed on the zero crossing between the 3-th and 4-th cycle of the carrier, at $t = 30\mu$s. The accuracy of the envelope shape, during the rising edge, is very important because it allows the cycle selection performed by the automatic receivers. The falling edge of the envelope is not used by the receivers because contamination of ionospheric wave is possible, with unpredictable results. In the transmitting stations the falling edge is damped with a time-law not defined; the duration of the whole pulse is usually 200–300 μs.

In order to find the minimum resolution N of the converter, expressed in bit, consider that the peak-to-peak values of (1) are in the ± 1 interval; the maximum excursion is $a_m = 2$. The first half-cycle ($t = 2.5\mu$s) is the smallest in the rising edge of the envelope; denoting its peak amplitude with $a_1 \simeq 0.01$, the minimum N, including the sign bit, is expressed by:

$$N = T\left(\log_2 \frac{a_m}{a_1}\right) = 8 \qquad (2)$$

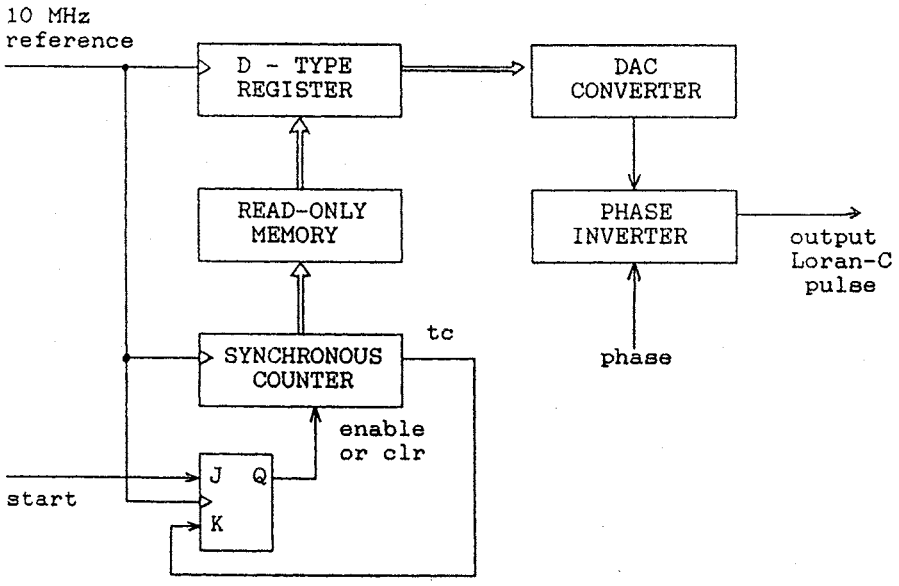


Figure 1: block diagram of the pulse generator.

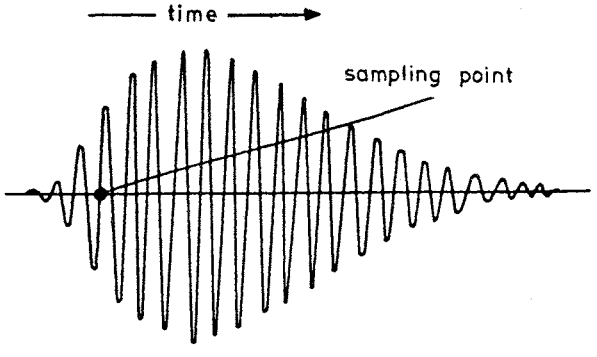


Figure 2: waveform of the Loran-C pulse

where T(.) means the integer greater or equal to its argument. In order to get the highest accuracy also in the first part of the pulse, a 12 bit converter is used.

An amplitude resolution of less than 8 bit could produce a distortion of the envelope, or the first half-cycle is skipped at all, depending on the algorithm of approximation; various criteria can be followed in this last function:

1. the nearest integer
2. the integer m with the nearest energy $|m|^2$
3. a resolution increasing obtained keeping a track of the roundoff errors.

The approach 1 and 2 shows some differences if a poor resolution — around 8 bit — is used; with an higher resolution the sampled values are the same. The approach 3 causes higher noise level, due to the presence of harmonics below the sampling frequency.

In order to keep the duration within 300 μs, a linear function is subtracted from the envelope of the formula 1 since $t = 250\mu$s.

3 Description of the instrument

The whole instrument is built around the pulse generator, following the block diagram of fig 3.

The kernel of the timing system is the main counter (at the center of the diagram) working at 10 MHz. The comparator, taking the inputs from the main counter and the time register, provides a coincidence signal with a resolution of 100 ns, used in many ways, depending on the content of the control register:

- start of a Loran-C pulse,

- clear of the main counter, when GRI is completed,
- activation of auxiliary TTL outputs, such as a pulse per Loran-C pulse, a pulse per GRI etc.

The coincidence signal sends also a service request to the microprocessor, with the highest priority. During the service response, the processor sets all the parameters of the next pulse to be generated: time, amplitude, phase and auxiliary signals.

Counting the service requests, the processor updates a time scale used for the automatic modification of the time differences, when the simulation of the speed is activated. When above mentioned services are not active, the processor provides all the functions of keyboard and display management.

The programmable amplifier includes a large bandwidth (500 kHz) two real poles filter for the suppression of the highest noise frequencies.

4 Considerations about the stability

In order to understand the meaning of the stability of the instrument, let us consider an ideal Loran-C signal driven by the same frequency reference. The signal available at the analog output is phase shifted, with respect to the ideal signal, of some amount of time due to the internal circuitry of the pulse generator (fig. 1). But the ideal signal is not available; when testing a receiver, the auxiliary signals, such as the TTL pulse corresponding to each Loran-C pulse, should be used as reference in all the measurements. A phase shift of these signals is another source of errors and instabilities, coming from the timing and control logic (fig 3). The considerations that follow involve both stability and accuracy.

All the timing and control logic is built with synchronous counters and D-type registers, and consequently the signals are synchronized to the frequency reference. All internal delays are obtained as a sum of an integer number of clock periods, plus the delay from clock to output of the last counter or register along the path considered.

The number of circuitry levels from the main frequency reference — internal quartz oscillator or external input — is kept minimum. From experiences and the data sheet of the electronic parts (TTL Schottky) typical switching times are about 3–5 ns; in laboratory conditions the stability is guaranteed within a small fraction of one nanosecond, that is almost meaningless dealing with Loran-C signals.

The delay D from the start signal of the pulse generator to the analog output of the synthesizer is the most critical parameter concerning accuracy and stability. It can be written as a sum:

$$D = D_1 + D_2 + D_3 + n \cdot 10\mu s \qquad (3)$$

where:

D_1 is a fixed value determined by the standard frequency,

D_2 represents the equivalent phase shift from the digital input to the analog output of the converter; for this applications it can not be directly derived from the characteristics of the component. What is known for the converter adopted (DAC 812) is that the whole delay is in the range 30–55 ns, depending on the amplitude of the difference from the previous value, and the temperature,

D_3 is the delay of the wide band filter, including the programmable amplifier and the phase inverter. In order to prevent possible variations, all the analog parts are low thermal sensitivity ones,

and the last term, $n \cdot 10\mu s$, has an obvious meaning: it depends on the n-th zero crossing involved in the measurement; n is integer or half-integer.

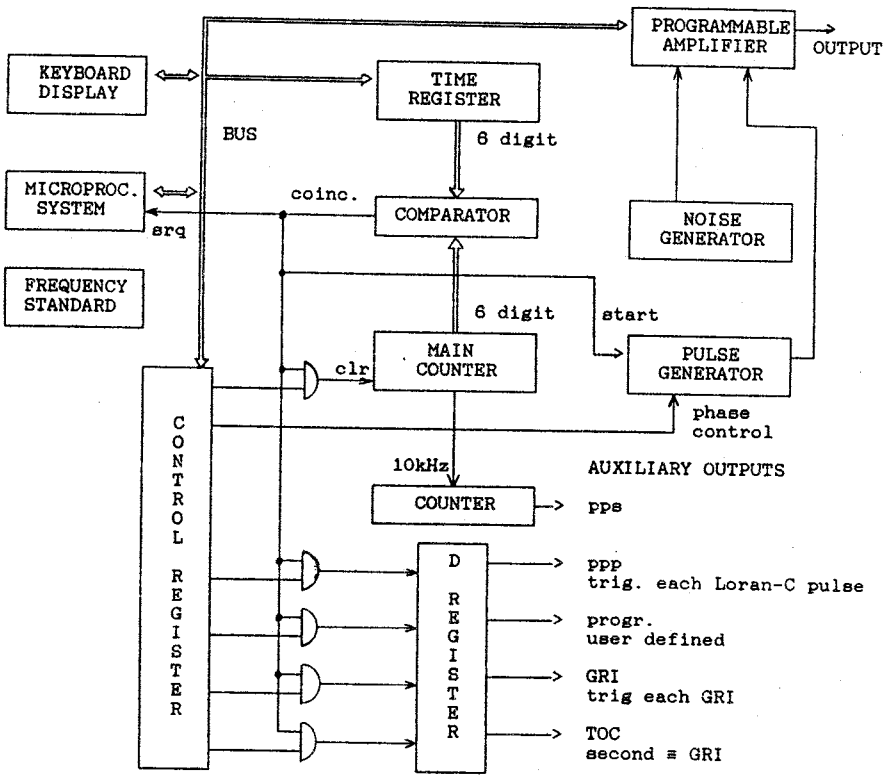


Figure 3: block diagram of the whole instrument

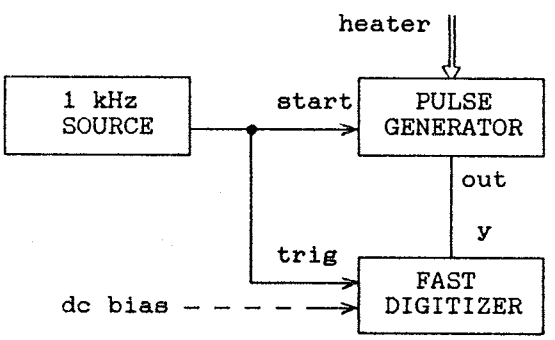


Figure 4: test of the pulse generator

The variations of the pulse generator were tested in the conditions showed in fig. 4. The digitizer (Tektronix 7612 used at a sampling frequency of 20 MHz) has a resolution of 8 bit, but the amplitude resolution has been set about equal to the DAC one by overloading the y input; when the peak amplitudes where under test, a dc bias was added in order to keep the signal in the range of the instrument.

Heating the pulse generator from the room temperature to 60 °C, in the rising edge of the envelope, but excluding the first cycle, no variation can be detected within a *zero uncertainty* of 2 ns; the same value is expected in the falling edge.

In the same conditions, the variations of peak amplitudes can not be detected within 2 less-significant-bit of the converter.

5 Final remarks

The stability obtained is a consequence of the method of synthesis.

The digital synthesis is flexible and the pulse shape can be easily modified; substituting the look-up rom with a random access memory dynamically downloaded by the microprocessor, it is possible the simulation of some other effects, such as the envelope-to-cycle distortion, useful in testing navigation

receivers.

The method is also suitable for a more simple practical realization. This was done, combining the pulse generator of fig. 1 with a subset of the control logic, in a very small (one eurocard) equipment for the simulation of a single Loran-C transmitter. In this second instrument, used in conjunction with a navigation receiver for the continuous monitoring of some transmitters, the signal is generated by an 8 bit converter working at 2.5 MHz, without any filtering; consequently, despite of the use of very low cost parts, the stability is just a little bit better than in the other one.

An improvement of the method is under study: an higher time resolution, a fraction of the sampling period, can be achieved by switching the memory mapped shape. The ultimate resolution for the Loran-C waveform is not well known, but, with a pure sinusoidal signal, a value of less than 0.001 rad was obtained.

References

[1] Ruhnow W.B. "Long Range Radio Navaid Signal Reliability" *Navigation* vol. 29 n. 2 summer 1982, pp. 152–159.

[2] Leschiutta S., Senese V., Cordara F. "Reliability of Mediterranean Sea Loran-C Chain, Period 1977–1983" in Proc. of *Wilde Goose Association European Symposium*, Sylt 1985.

[3] Leschiutta S., Cordara F. Song Jin An "Reliability of Mediterranean Sea Loran-C Chain, Period 1977–1985" in Proc. of *IEEE Position Location & Navigation Symposium — Navigation into the 21st Century* Las Vegas 1986.

[4] Kalliomäki K. "National Time and Frequency Standards of Finland" in *Radio Science* vol. 14 n. 4, july-august 1979, pp. 677–679.

[5] Rubiola E. "Metrological Use of Loran-C Navigation Receivers" paper submitted to *IEEE Transaction on Instrumentation and Measurement*, 1989.

[6] Frank R.L. "Current Developments in Loran-C", in Proc. of IEEE, vol. 10, October 1983.

[7] "Specification of the Transmitted Loran-C Signal" published by U.S. Coast Guard, Washington D.C. July 1981.

[8] He Zhi-long "A Method of Generating the Standard Loran-C Pulse and Envelop Signal" *Publications of Shaanxi Observatory*, vol.6 n.2 pp.14-19 Dec. 1983.

DIGITAL FREQUENCY SYNCHRONISATION WITH LF TRANSMITTERS

Peter SCHUMACHER

Observatoire Cantonal de Neuchâtel, Switzerland

SUMMARY

There are several LF time signal transmitters in Europe with frequencies at 75 kHz, 77.5 kHz, 60 kHz and 50 kHz. The frequencies of these transmitters are stabilized by Cesium beam atomic frequency standards with a long term stability of better than $1 \cdot 10^{-12}$. For short time periods (1 to 10 seconds) the frequency stability of the received carrier is limited to about $1 \cdot 10^{-8}$ to $1 \cdot 10^{-9}$ due to variations in the propagation delay of the LF radiowaves.

In this paper we present a new receiver, developed at the Observatory of Neuchâtel, which contains a BVA X-tal oscillator phaselocked to the received carrier. The phaselock loop is non-linear and digital, with a time constant which varies typically between 10^2 and 10^6 seconds. The digital loop has been realized with a single chip processor and a simple DA and AD conversion technique.

The receiver construction is described and the principle characteristics of the digital phaselock loop are discussed.

1. INTRODUCTION

In Europe there are several time signal transmitters that are synchronized by Cesium beam frequency standards:

Table I: Time signal transmitters in Europe

HBG	75 kHz	Switzerland
DCF	77.5 kHz	Germany
MSF	60 kHz	England
OMA	50 kHz	Tschekoslowakia
FJ	162 kHz	France

The carrier frequencies of these transmitters are very stable and can be used to synchronize local quartz crystal oscillators at distances of up to 1000 km from the transmitter. For short measurement time periods the variations of the propagation characteristics are causing some frequency fluctuations at the receiving site. The frequency stability of these transmitters has been published in the literature [1], [2] and is represented in Fig. 1.

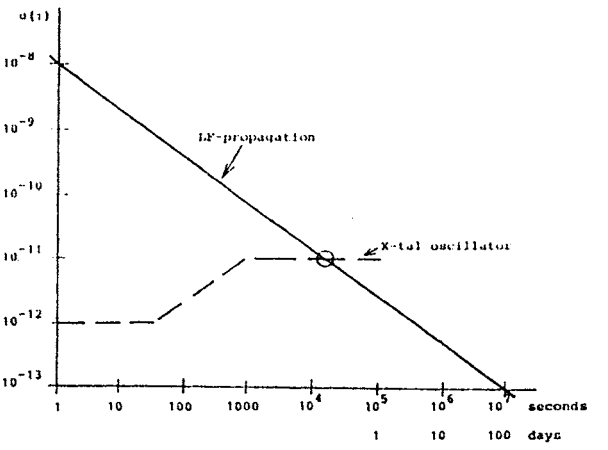


Fig. 1: Frequency stability σ(τ) as a function of measurement time:
—— LF transmitter
----- BVA X-tal oscillator

Fig. 1 shows that the frequency stability at the receiving site is improving with the averaging time τ, from a value of 10^{-8} for τ = 1 sec to a value of 10^{-12} for τ = 1 day. However a measurement of the last value is quite extraordinary. The presented values are for a distance of 100 km from the transmitter.

The dashed line in Fig. 1 represents the frequency stability of a BVA X-tal oscillator. For short time periods the values are from the specification sheet of the manufacturer; for time periods longer than 1000 sec we consider frequency fluctuations associated to temperature fluctuations of 2°C in a normal environment [3]. BVA X-tal oscillators having better short term stability performance are also available.

With a radio frequency receiver the crystal oscillator can be phaselocked to the incoming standard frequency signal. For periods shorter than the phaselock loop time constant, the output frequency is determined by the quality of the X-tal oscillator. For periods longer than the loop time constant, the frequency is following the transmitted signal characteristics. There is an optimum time constant at the crossover point of the 2 lines in Fig. 1. Its value is about 10^4 seconds. Phaselock loops with so slow time constants can only be done with digital components. We present therefore in this paper the design of a digital control loop with long time constants.

2. LINEAR PHASELOCK LOOP

The linear phaselock loop is physically and mathematically well known [4], [5]. The basic elements of the loop are a phase detector, a loop filter and a voltage controlled (crystal) oscillator (Fig. 2).

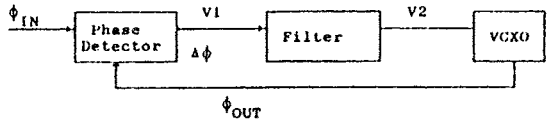


Fig. 2: Schematic of phaselock loop.

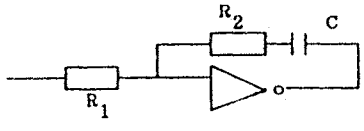


Fig. 3: Active filter

For a second order loop with a first order active filter, we get the following equations:

$$\omega_n = \left(\frac{K_0 K_D}{T_1}\right)^{1/2}$$

$$\zeta = \frac{1}{2} T_2 \cdot \left(\frac{K_0 K_D}{T_1}\right)^{1/2}$$

where ω_n is the natural frequency of the loop and ζ is the damping factor; ζ is dimensionless. K_0 is the oscillator constant (1/Volt), K_D is the phase detector constant (Volt/μsec).

The lock-in range of the loop is defined by the product of ω_n and ζ:

$$\omega_L = 2\omega_n\zeta = K_0 K_D \cdot T_2/T_1$$

The behaviour of a linear phaselock loop can be determined by the response of the error transfer function $H_E(S)$, if an input frequency step $\Delta\omega$ is applied to the input signal. The result is a set of functions, that is typical for a second order differential equation.

A graph of these functions is represented in Fig. 4.

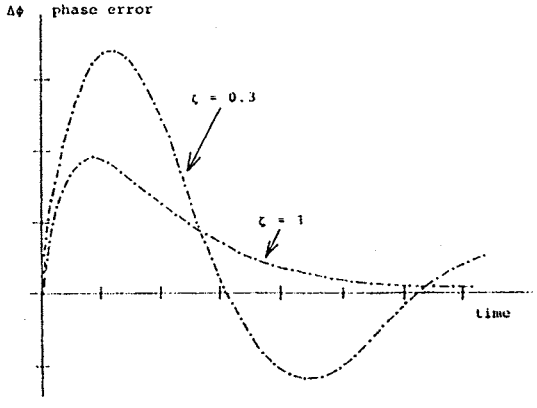


Fig. 4: Loop response to an input frequency step $\Delta\omega$ (linear loop).

We distinguish the typical solutions for damped oscillations ($\zeta < 1$) and critical damping ($\zeta = 1$). The time constant of the loop T_L is the inverse of the lock-in range

$$T_L = \frac{1}{\omega_L} = \frac{1}{2\omega_n \zeta}$$

T_L is one half of the value of the decay time in the exponential term $e^{-\omega_n \zeta t}$. Note that for a first order loop, decay time and loop time constant are equal.

3. DIGITAL NONLINEAR PHASELOCK LOOP

Practical phaselock loops are always nonlinear: Phase detectors have only a limited linear range, typically within half a cycle of the input frequency; active filters have a limited voltage output level; the oscillator control voltage has a limited range and within that range the variation of the oscillator constant K_0 is nonlinear up to a factor of 2 or 3. However, one can suppose that the loop is linear, as long as the phase and frequency deviations from the locked condition are small.

A linear digital phaselock loop is similar to its analog equivalent (Fig. 5).

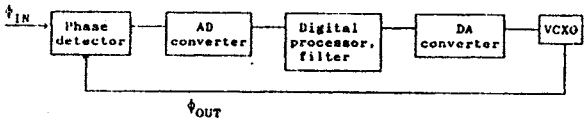


Fig. 5: Schematic of digital phaselock loop.

It includes essentially an analog-digital converter (AD), a digital processor and filter, a digital-analog converter (DA), as well as a phase detector and a voltage controlled oscillator. Compared with an analog loop, the loop filter is replaced with a loop filter algorithm. An integration constant is done by repeated additions, an analog amplification factor is done by digital multiplication. If the resolution of the converters is sufficiently high, for instance: 8 bit for the AD-converter and 16 bit for the DA-converter, then the loop design follows the same rules as in the analog loop, since the phase and frequency steps in the loop are quasi-continuous.

For the nonlinear loop that we are considering here, phase and frequency steps are great; we have therefore a loop where the step width is essential. The basic hardware schematic is shown in Fig. 6.

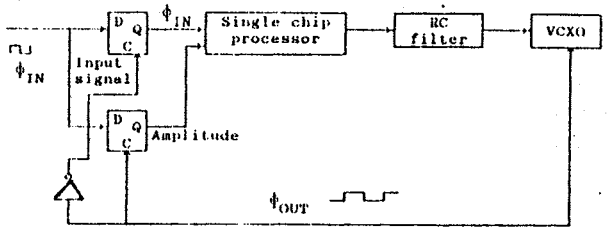


Fig.6: Hardware of nonlinear loop.

The phase detector and the AD-converter consist simply of a D-type flipflop. Another D-type flipflop detects the amplitude of the signal. The DA-converter is a RC-time constant (1 sec) which is filtering the phase-modulated output signal of the single chip processor. The VCXO is a BVA crystal oscillator. Everything else is done by software.

Our nonlinear phaselock loop has therefore the following features: It is a simple loop with very few hardware components. For the filter algorithm we have chosen the most simple algorithm possible that is equivalent to a second order loop. The phaselock loop is very well working in practice and it is interesting from a mathematical point of view.

The digital processing is schematically represented in Fig. 7.

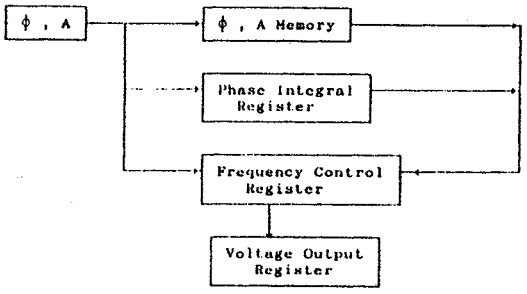


Fig. 7: Schematic of digital filter.

The output of the D-type flipflop phase detector is read with a rate of 64 Hz. The acceleration rate factor (a_f) that varies in steps of 1 to 10^5 is multiplied with the sign of the input phase value (positive or negative). The result is added to the frequency control register (FC). The unsigned acceleration factor is added to the phase integral register (PJ). Both registers have a length of 32 bits. If there has been a phase transition compared with the previous phase sample in the (ϕ, A)-memory and if the amplitude detection is positive, then the PJ register is divided by 2, multiplied with the inverse sign of the transition, added to the FC-register and cleared. This is called a halfstep operation.

The first 16 bits of the FC register are transferred to the voltage output register (VO). With the aid of a control loop of the internal timer signal (1 μsec - steps) a phasemodulated 64 Hz signal is created, where the duty cycle is proportional to the value of the VO-register. The real time operation of the loop is graphically represented in Fig. 8.

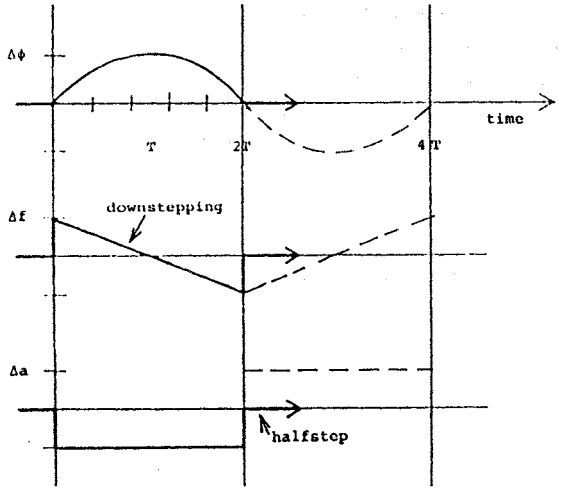


Fig. 8: Phase, frequency and acceleration in function of time (nonlinear loop).

We suppose that the loop is initially locked. At the time 0 we apply a frequency step Δf at the input. The phase error $\Delta\phi$ gets positive and consequently the frequency is downstepping with a rate a: a is the frequency acceleration rate ($\Delta f/f/sec$), which is equivalent to the gainfactor in a linear loop. The phase error $\Delta\phi$ in function of time is parabolic.

At the time T the original frequency step has been downcounted to 0 and the phase has attained its peak value $\Delta\phi_{max}$. At the time 2T the phase has come back to 0 and the original frequency error Δf has changed the sign. Since the parabola is symmetric to the initial condition at the time 0, the loop would oscillate with another parabolic segment. However, if we introduce at this moment a halfstep operation, then the loop returns immediately back to its initial condition before time 0, that is $\Delta\phi = 0$, $\Delta f = 0$. This is similar to the case $\zeta = 1$ in the linear loop.

The mathematics of the loop is contained in the following equations:

1) $\Delta\phi = \Delta f \cdot t - 1/2\ a \cdot t^2$ ($\Delta\phi$ in seconds)

2) $\Delta f = a \cdot t$ (Δf dimensionless).

The acceleration rate is equal to the complete frequency control domain of the crystal oscillator ($2 \cdot 10^{-8}/5$ Volt), multiplied by the acceleration rate factor (a_f : 1 to 10^5), the FC register factor ($1/2^{32}$) and the addition rate (64 Hz).

The time constant of a nonlinear loop can be defined by the time required to cancel an initial phase or frequency perturbation (see Fig. 8). Its value can easily be derived from equations 1), 2).

$$T_{\Delta\phi} = \sqrt{\frac{2}{a}} \quad \sqrt{\Delta\phi} \qquad \text{for } \Delta\phi_{INIT} = \Delta\phi, \quad \Delta f_{INIT} = 0$$

$$T_{\Delta f} = \frac{1}{a} \quad \Delta f \qquad \text{for } \Delta\phi_{INIT} = 0, \quad \Delta f_{INIT} = \Delta f$$

This result shows that the phase time constant $T_{\Delta f}$ is proportional to the root of the phase step and the inverse root of the acceleration, where the frequency time constant is proportional to the frequency step. This is very different from a linear phaselock loop, where the time constant is independent of the input phase or frequency deviations. If one defines the loop bandwidth $B_{\Delta\phi}$ as the inverse of $T_{\Delta f}$:

$$B_{\Delta\phi} = \frac{1}{T_{\Delta\phi}} = \sqrt{\frac{a}{2}} \cdot \frac{1}{\sqrt{\Delta\phi}} \quad ,$$

then one gets the concept of a phasenoise dependent loop bandwidth, as it is found also under high noise conditions in linear phaselock loops.

Another interesting quantity is the maximal phase deviation ϕ_{max} for an initial step Δf and the maximal frequency deviation Δf_{max} for an initial step $\Delta\phi$:

$$\Delta\phi_{max} = \frac{1}{2} a \cdot T_{\Delta f}^{\,2} = \frac{1}{2a} \cdot \Delta f^2$$

$$\Delta f_{max} = a \cdot T_{\Delta\phi} = \sqrt{2a \cdot \Delta\phi} \quad .$$

Numerical values of the time constants have been calculated for different values of a and are represented in Table 2.

TABLE 2: Numerical values of nonlinear loop time constants for an initial phase step of 0.5 μsec and an initial frequency step of $1 \cdot 10^{-10}$.

Position	Acceleration rate a (Δf/f/sec)	$T_{\Delta\phi}$ [Δφ = 0.5 μsec] (sec)	$T_{\Delta f}$ [Δf = 10^{-10}] (sec)	$2 \cdot \Delta\phi_{max}$ [Δf = 10^{-10}] (sec)	Δf_{max} [Δφ = 0.5μsec] no dim.
1	10^{-15}	$3 \cdot 10^{4}$	10^{5}	10^{-5}	$3 \cdot 10^{-11}$
2	10^{-14}	10^{4}	10^{4}	10^{-6}	10^{-10}
3	10^{-13}	3000	1000	10^{-7}	$3 \cdot 10^{-10}$
4	10^{-12}	1000	100	10^{-8}	10^{-9}
5	10^{-11}	300	10	10^{-9}	$3 \cdot 10^{-9}$
6	10^{-10}	100	1	10^{-10}	10^{-8}

4. MEASUREMENTS

A complete phaselock system has been realized and implemented on two electronic boards. The schematic is presented in Fig. 9.

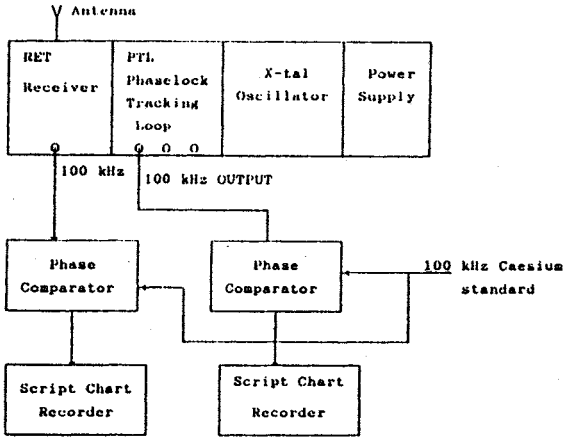


Fig. 9: Phaselock loop with measurement test set.

We distinguish a phaselock receiver (RET), a phaselock tracking loop (PTL), a BVA X-tal oscillator and a power supply with NiCd-backup.

The RET phaselock receiver is a new receiver, that is very similar to the receiver, that has been designed some years ago [6]. The specifications are:

Antenna: 10 cm-ferrite, 75 kHz
Input sensitivity: better than 1 μV
Loop time constant: 1 second
Output signals: carrier amplitude,
coherent 100 kHz signal
Dimensions: 100 x 160 mm

The PTL phaselock tracking unit has the following specifications:

Reference input: 5 MHz or 100 kHz
Receiver input: 100 kHz, square wave; carrier amplitude
Output: 100 kHz, 1 MHz, 10 MHz synchronous with the receiver input
Time constant: (0.5 μs phase step)
12 position switch with steps from 1 to $3 \cdot 10^5$ seconds
Frequency stability: better than $1 \cdot 10^{-10}$ for $\tau < 10^4$ seconds
better than $1 \cdot 10^{-11}$ for $\tau < 10^5$ seconds

If the system is equiped with a BVA X-tal oscillator, it can be considered as a receiver frequency standard with a very precise output frequency. It can be used for calibrations of other crystal oscillators, frequency and time interval counters, or as a basic low cost frequency standard.

Different test measurements have been performed. The basic measurement setup is represented in Fig. 9. It consists of two phase comparators and two script chart recorders. The phase output of the comparators is registered for the input and output signals of the PTL phase tracking loop; both signals are compared with a Cesium beam frequency standard. A typical record is represented in Fig. 10. It is easy to see that the output signal of the PTL has much smaller variations than the filtered output signal. Fast ionospheric perturbations, visible on Fig. 10, have disappeared. All perturbations are smaller than 10^{-10} and for a 1 day average the values are better than 10^{-11}. This is for a time constant of 10^{-4} seconds at a distance of 80 km from the transmitter. This record shows a qualitative and quantitative agreement with the $\sigma(\tau)$-plot of Fig. 1.

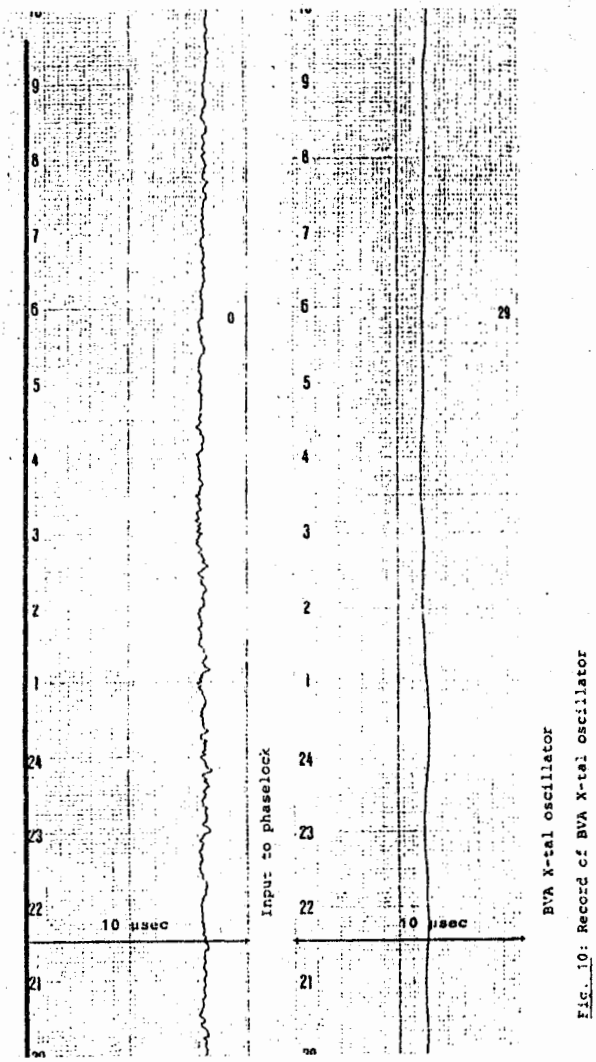


Fig. 10: Record of BVA X-tal oscillator

In another measurement setup the BVA X-tal oscillator has been replaced by a low cost temperature controlled oscillator in a dual in line case. The output is registered in Fig. 11. Note the difference of ionospheric variations between daytime and night-time. It is easy to see that during daytime the oscillator performance is visible. At night time the oscillator and the receiver have approximately the same noise. Note also a parabolic peak at the hour 12:45. The time constant is approximately 300 seconds.

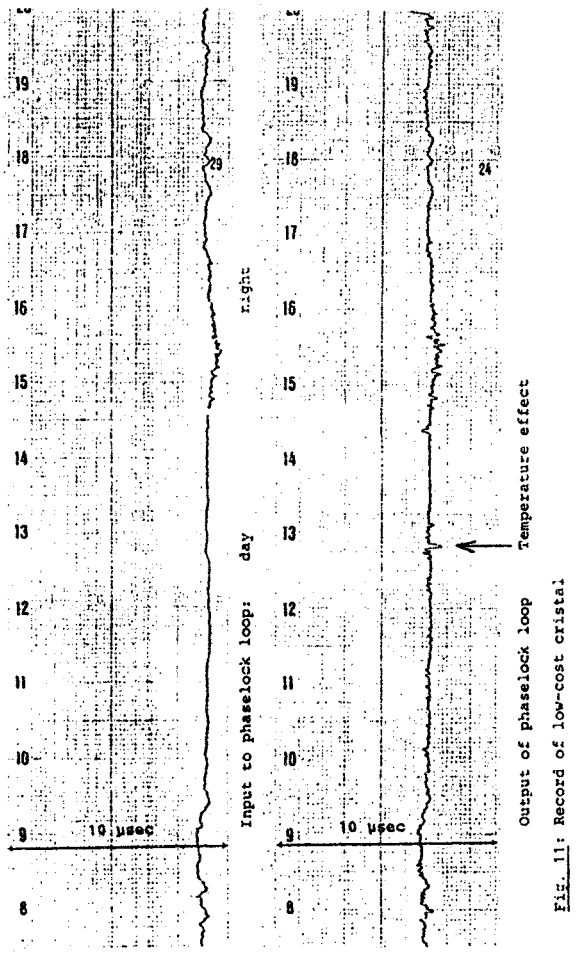


Fig. 11: Record of low-cost crystal

The total varicap domain of the 2 oscillators is a factor of 1000 different: ($2 \cdot 10^{-8}$ and $2 \cdot 10^{-5}$ respectively). This changes automatically the loop time constant accordingly to the calculations in Table 2. For a linear loop, it would have been necessary to adapt also the ζ-factor which depends on the oscillator constant. For our nonlinear loop, however, this is not necessary; the loop is selfadapting to different oscillator constants.

Another test that has been performed was the replacement of the receiver output signal by a 100 kHz signal that has been derived from the Cesium beam standard. The optimal time constant can easily be found by measuring the output with different acceleration rate factors. For time constants below the optimum, the oscillator is tracking exactly the input signal. For time constants above the optimum, thermal and other fluctuations become apparent (see Fig. 10). This is in agreement with the introductory discussion of Fig. 1.

5. CONCLUSION

In this paper we have demonstrated a practical application of a digital nonlinear phaselock loop. The nonlinear case has been compared as far as possible with the linear one. A receiver-frequency standard based on the presented mathematical calculations has been realized. The measurements are in agreement with the expected performance.

REFERENCES

[1] J. Bonanomi et P. Kartaschoff: Le nouveau service horaire HBG sur 75 kHz, Société Suisse de Chronométrie, 1966.

[2] P. Hetzel und U. Hübner: Der Zeitsignal- und Normalfrequenzsender DCF 77, Congrès International de Chronométrie, 1974.

[3] E.P. Graf and U.R. Peier: BVA Quartz Crystal Resonator and Oscillator Production, paper presented at the 37th Annual Frequency Control Symposium in Philadelphia, USA, 1983.

[4] Floyd M. Gardner: Phaselock Techniques, John Wiley and Sons, New York, 1966.

[5] W.C. Lindsey and C. Ming Chie: A Survey of Digital Phase-Locked Loops, Proceedings of the IEEE, vol. 69, no. 4, April 1981.

[6] P. Schumacher: Oscillateurs-étalons asservis par l'émetteur HBG, Société Suisse de Chronométrie, 1976.

Service quality of a long distance telecommunication network in distributing timing signals (frequency)

Tapio Kulmala, lic.tech.

P & T Telecom Research Centre
Finland

Abstract

The long term timing quality of a long distance transmission network is investigated by comparing the phase of a 300 kHz test signal with that of a Cesium frequency standard, monitored by a Loran-C-receiver. The transmission route, Helsinki-Sodankylä (1100 km), consists of coaxial cable systems (700 km) and microwave radio links (400 km). The results show strong diurnal delay variations, typically within 0.4 μs, thus demonstrating the ducting effect of microwaves in tropospheric inversion layers. The slow annual delay variation is within 4 μs, and can be explained by earth temperature variation. The capability of the finnish trunk network in transferring precise timing signals needed by modern digital networks is predicted.

1. Introduction

In telecommunications, incorporating both the conventional FDM and modern digital transmission systems, the local oscillators must be synchronized to produce a correct frequency within a given uncertainty. The frequency accuracy requirement is most stringent at the highest hierarchical level of the national digital networks, where the long term accuracy of 1×10^{-11} is required for reference clocks, according to CCITT Recommendation G.811. However, in synchronizing the network, in accordance with a certain scheme [1], the frequency accuracy will be degraded due to the temporal variations of delay between the network nodes. The slow delay variation is sometimes called "wander" /1/.

The meteorological background of the delay variations of cable and microwave radio links is reasonably well-understood /2, 3, 4/, but only a limited number of experimental results concerning the long term wander of real long distance telecommunication networks are available at the present /5/. In order to facilitate the evaluation of the timing accuracy performance of future digital networks, P & T of Finland started an experimental study in 1984. This report presents essential results and preliminary conclusions of the study which is still in progress. The results are applied to two types of timing networks:

a) general purpose reference frequency distribution network and

b) digital synchronization network.

1) Hierarchical Master-Slave synchronization scheme is used in Finland.

2. Transmission network and measurement methods

The transmission route under study is presented in figure 1. The broadband network consists of coaxial cable systems (Helsinki-Muhos, 700 km) and microwave radio links (Muhos-Sodankylä, 400 km). The "test signal" of this study is a 300 kHz sinusoidal "pilot" which was originally intended to control "manually" the frequencies of local carrier oscillators of FDM-transmission systems. The source of the pilot frequency is a Cesium-oscillator in Helsinki.

The pilot passes the 2700 speech channel cable systems without any modulation or demodulation processes, only the bandpass filtering is needed to separate the 300 kHz signal from the transmission band. In the case of the 1800 and 900 speech channel radio links conventional FM-techniques and 6-8 GHz microwave bands are used. The FDM-cable network Helsinki-Muhos includes 12 telecommunication stations (nodes) and 132 repeater amplifiers. The FM-radio network Muhos-Sodankylä consists of 8 radio hops which means 5 terminal stations and 4 repeater stations. The average distance of one radio hop is about 50 km.

The delay variation between Helsinki and Sodankylä is measured by comparing the phase of the received 300 kHz pilot to that of a Cesium-oscillator (Cs2, in figure 1.). The Cs2-oscillator is a property of EISCAT-project. The distance between the Cs2-oscillator and the telecommunication station is about 7 km. Hence, it is necessary to transmit the Cs-frequency (10 kHz) to the phase comparator $\varphi 2$ via a symmetrical pair of an overhead cable. Two other pairs and a phase comparator $\varphi 3$ are used to eliminate the effect of the cable delay variation caused by temperature changes (ΔT). In fact, the phase comparator $\varphi 3$ will act as an atmospheric temperature indicator. To accomplish a reliable long term phase reference for pilot measurements the Cesium-oscillators Cs1 and Cs2 are monitored by using Loran-C-receivers 1. and 2. In order to indicate the interruptions on the pilot frequency distribution the voltage level of the pilot is registered by a multi-channel chart recorder together with the phase comparison results.

For investigating the short term frequency instability of the pilot a counter based Allan-variance measurement

system is utilized. For relating the observed delay variations to the relevant weather data temperature and humidity values at different altitudes. collected by radiosondes of the Meteorological Institute, are worthwhile.

3. Results

In considering delay variations of transmission networks, three different types of variation may be distinguished /3/:

A. diurnal
B. seasonal (including annual) and
C. random (sometimes called "short term") variation.

A. Diurnal delay variations

In this case the observed diurnal delay variations could be divided into the following categories a-d:

a) periodic day-night -variation, which correlates clearly with the atmospheric temperature variation:
- the delay is shortest in the daytime, while the temperature is at maximum
- typical from March to October, most common in summertime
- the magnitude of the diurnal delay variation typically ≤ 0.4 µs, from peak to peak
- example: figure 2.

b) nonperiodic diurnal delay variation, which doesn't correlate clearly with temperature:
- quite rapid delay changes may occur, especially early in the morning (from 7.00 to 9.00 am) or, sometimes, late in the evening (from 9.00 to 12.00 pm)
- irregularly from November to February; i.e. typical wintertime phenomenon
- the observed maximum rate of delay change is about -0.5 µs/2 h; the quite rapid decrease in delay is normally followed by a slow increase
- example: figure 3.

c) constant delay and temperature, or only neglible variations
- this sort of "quiet" periods with the duration of a few of days or, sometimes, a couple of weeks may occur throughout the year.

d) no measurable or only very small delay variations with noticeable diurnal temperature variations.
- periods with the duration of a few of days may occur at times throughout the year.

B. Seasonal delay variations

An example of a very long term delay variation is presented in figure 4. where the results of the period from November 1986 to November 1987 are shown. The delay variation ("pilot-Cs1") has been calculated from the observed phase differences between the 300 kHz pilot and the Cs2-oscillator by correcting for the effect of the mutual phase drift of the oscillators Cs1 and Cs2. This phase drift correction was made by using the Loran-C-system as a common phase reference and the sample interval of one week.

The annual delay variation is within 4 µs, from peak to peak. Also shorter term seasonal variations are observable. The total transmission delay seems to be at minimum in wintertime. Similar results have been obtained from the period 1984-1986.

C. Random delay variations

The Allan-variance of fractional frequency fluctuations ($y=\Delta f/f$) is a practical statistical measure for the short term instability of the pilot phase and frequency /6/. The square roots of the variance (σ_y) at sample intervals $\tau=10^{-3}...10^{5}$ s are shown in figure 5. It seems that the short term instability of the received pilot is noticeably contributed by white phase noise. At long sample intervals $\tau>10^{4}$ s the instability of the pilot frequency approaches that of the commercial Cesium-oscillators (Cs1, Cs2). In order to avoid the influence of diurnal delay variations on the observed σ_y-values ($\tau>10^{3}$ s) the measurements were carried out during a "quiet" period, i.e. diurnal variations were neglible in that period.

D. Availability of pilot

During the period of 1984-1987 the experiments suffered from a number of unexpected interruptions. Most of the interruptions of long duration (days ... months) were caused by the network development work or failures in the measurement system. The rest of the interruptions results in the pilot avalability of 0.94 (94 %) and "MTBF" of three days, which are quite pessimistic values.for a present-day transmission network. The poor availability is certainly influenced by the maintenance work along the network and by the fact that the pilot is distributed over the fixed route without any active redundancy.

The phase comparison results of the pilot contain small temporary phase jumps (< 0.1 µs), at most a few of tens per year. These phase jumps are most probably due to the switch-over -operation involved with the diversity reception techniques of the radio links.

4. Interpretation of results

This chapter discusses the possible causes of the observed delay variations. The emphasis is on the diurnal and seasonal variations and the meterological factors responsible for these variations.

The subterranean coaxial cable systems are well-shielded against atmospheric temperature changes. Hence, only very slow cable and equipment temperature changes, of the order of 0.1°C per day, are possible. At the normal cable installation depth of 0.5-0.8 m the annual earth temperature variation in Finland is within 15°C, including the stabilizing effect of the snow cover /7/. One can conclude that only negligible diurnal delay variations are expected due to the coaxial cable systems. The results of the preliminary measurements (Oulu, 1982) support this conclusion. On the other hand, the annual delay variation can be noticeable, as shown in figure 4.

The radio refractivity [2] (N) of the troposphere and its height dependence or "profile" (N(h)) can vary considerably in a day /2/. A special weather condition that can result in an anomalous propagation of radio waves is the occurrence of so-called inversion layers. The inversion means that a relatively cold and moist air layer is situated below a warm and dry air mass. Within this kind of tropospheric structure the vertical refractivity gradient (dN/dh) can obtain very large negative values compared to the standard atmospheric conditions, where dN/dh = -40 N units/km. In this context the concept of duct is often used, especially when dN/dh < -157 N units/km within the inversion layer.

Under those superrefractive conditions the ducting of radio waves may occur, especially at microwave frequencies. Hence, an inversion layer or duct can act as a waveguide with relatively low loss for appropriate propagation modes /8/. In Finland, and in countries with the similar climate, two types of tropospheric inversion layers are common:

a) nocturnal radiation inversion (summertime)
- is developed as a result of heat and moisture exchange between the ground and the lower troposphere at night, when air is calm and the sky is clear
- a typical ground based inversion layer or duct with a height of at most a few of hundreds of meters
- especially in summer, the solar heating or surface wind will destroy the inversion layer in the morning.

b) subsidence inversion (wintertime)
- in high-pressure areas the dry air coming from a high level in the atmosphere is heated by adiabatic compression
- produces stable inversion layers of relatively warm air above a cooler, moist air mass
- may form a so-called elevated duct; at night the lower boundary of the inversion may reach the earth's surface as a result of accompanying nocturnal radiation (a)
- in wintertime the inversion can stay, at least partly, for many days.

The propagation delay of the radio path can be presented by formulas

$$\tau = \frac{L}{c} = \frac{1}{c}\int_0^S n(S)dS = \frac{\bar{n}S}{c} \qquad (1)$$

where the integral (L) is the electrical path length of the radio ray propagating over the geometrical path S, $\bar{n}$ is the averange of n over path S and c is the velocity of light in vacuum.

---

2) Refractivity $N = (n-1)10^6$, where n is refractive index (n 1.0003 near the surface of the earth)

$$N = 77.6\, p/T + 3.73 \times 10^5\, e/T^2$$

where p is total atmospheric pressure (mb), T is temperature (K) and e is partial pressure of water vapor (mb).

The diurnal variation $\Delta\bar{n}/\bar{n}$ is estimated to be less than $3 \cdot 10^{-5}$ /2/. On the other hand, the observed diurnal delay variation $\Delta\tau/\tau$ may be of the order of $3 \cdot 10^{-4}$. Thus, one can conclude that about 90 % of the maximum diurnal delay variation is due to the variation in the geometrical path length $\Delta S/S$. This is very interesting result, because in many reports /4, 5/ discussing delay variations on microwave line-of-sight radio links the major cause of the delay variation has been assumed to be the variation in $\bar{n}$, and hence in the velocity of the plane wave ($v=c/\bar{n}$).The diurnal variation in geometrical path length may be about 15 m per 50 km of radio hop. This exceptional large variation can't be explained by the conventional tropospheric refraction /2/. The author's conclusion is that microwaves will be trapped in ducts formed by the inversion layers described above. This result is in agreement with a number of field strength measurements carried out in investigating the so-called beyond the horizon propagation at VHF-, UHF- and microwave bands /2, 9/.

5. Application of results to timing networks

Two types of timing networks are discussed briefly in this chapter:

a) general purpose reference frequency distribution networks for calibrating (remote) frequency measurement equipment, comparing frequency standards and disseminating time

b) synchronization networks for digital transmission, as mentioned in introduction (chapter 1).

A. General purpose reference frequency networks

In the case of the example network under study the uncertainty of distributed frequency due to the diurnal and seasonal delay variations is shown in figure 6.

These "worst case" uncertainty values have been obtained by time derivation of phase comparison records. As a reference, a typical frequency uncertainty (two sigma) of Loran-C receptions is presented /3/. The short term or random uncertainties are shown in figure 5. The overall frequency accuracy of the pilot network is acceptable, however, the limiting effect due to the strong diurnal delay variations is noticeable.

B. Digital synchronization networks

The long term delay variations observed on analog FDM-networks are present also on digital networks utilizing the same transmission media. The timing accuracy degradation caused by delay variations or "wander" can be evaluated as a function of observation time by using figure 6. It is interesting to compare the results with the frequency accuracy requirement [3] for the

---

3) The frequency accuracy requirement has been derived directly from the permissible time interval error (TIE) by assuming a constant frequency error in a given observation time.

national reference clocks. The comparison reveals that in our experiment the safety margin between CCITT Recommendation G.811 and the timing accuracy limit set by the network is less than decade.

6. Conclusion

The results obtained for the cable-radio network combination show strong diurnal delay variations, typically within 0.4 μs. In summertime these day-night -variations are quite periodic and seem to correlate reasonably well with the atmospheric temperature. In wintertime the variation is more anomalous and at times, especially early in the morning, the delay may decrease quite rapidly, even at the rate of 0.5 μs/2 h. The author's conclusion is that the strong diurnal delay variations can be explained by the ducting of microwaves in tropospheric inversion layers. Anyway, the major part (90%) of the variation in delay is due to the variation in the geometrical path length.

This result is interesting, because in many previous reports the delay variation has been assumed to be mainly due to the variation in the average refractive index ($\bar{n}$) integrated over the propagation path. Thus, the measured diurnal delay variation is about one decade larger than was expected. The slow annual delay variation is within 4 μs and can be explained by the annual earth temperature variation of the order of 15°C. The frequency uncertainty caused by the diurnal and seasonal delay variations will be $\pm$ 7x10^{-11}...$\pm$ 1x10^{-12}, when the observation time is 15 minutes ... one month or more.One important application of the results obtained is to predict the capability of the finnish trunk network in transferring precise timing signals needed by the digital network synchronization. The transmission network under study fulfils the timing accuracy requirement for international digital links according to CCITT Recommendation G.811 for reference clocks. However, diurnal delay variations caused by very long microwave radio links can restrict the "worst case" timing accuracy to the level corresponding about $\pm$ 10^{-10} uncertainty in frequency.

The analysis of the radiosonde data is still in progress and the results might give more detailed information of the role of inversion layers on the ducting of microwaves. Also statistical distributions of diurnal delay variation magnitudes will be studied more thoroughly than in this "summarized" report.

Acknowledgment

The author wishes to thank the following persons and organizations for their help:

Professor K. Kalliomäki at the University of Oulu
Mr. T. Laakso and M. Postila at EISCAT/Sodankylä
Dr. E. Kyrö at the Meteorological Institute
Mr. V. Lehtoranta at Broadcasting Company
Mr. R. Ranttila, H. Keinotie, R. Laine, E. Markkanen and R.Ripatti at P & T of Finland.

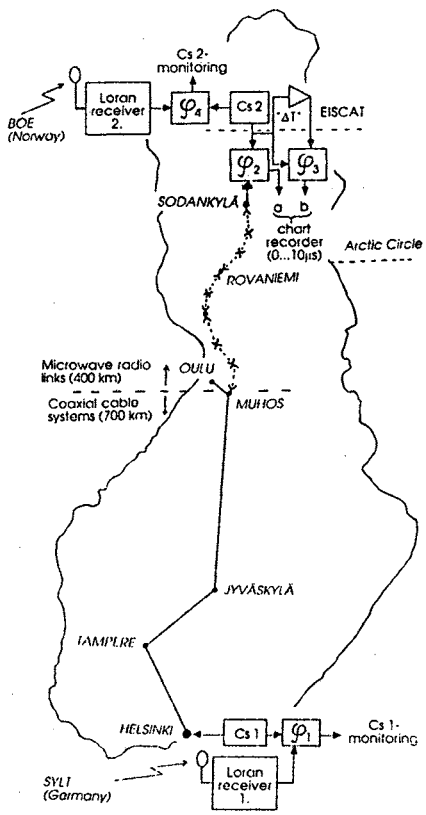


Figure 1: Transmission route and phase comparison measurement systems.

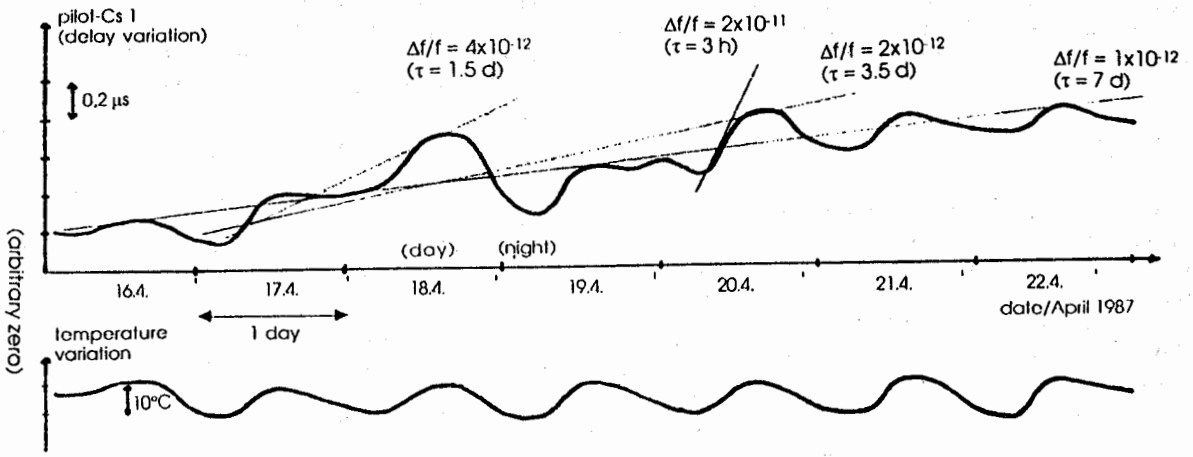


Figure 2: An example of diurnal delay and temperature variation; a period of one week, 16-22 April 1987. A slow seasonal delay variation is also observable. Frequency uncertainties are evaluated at different observation times (τ).

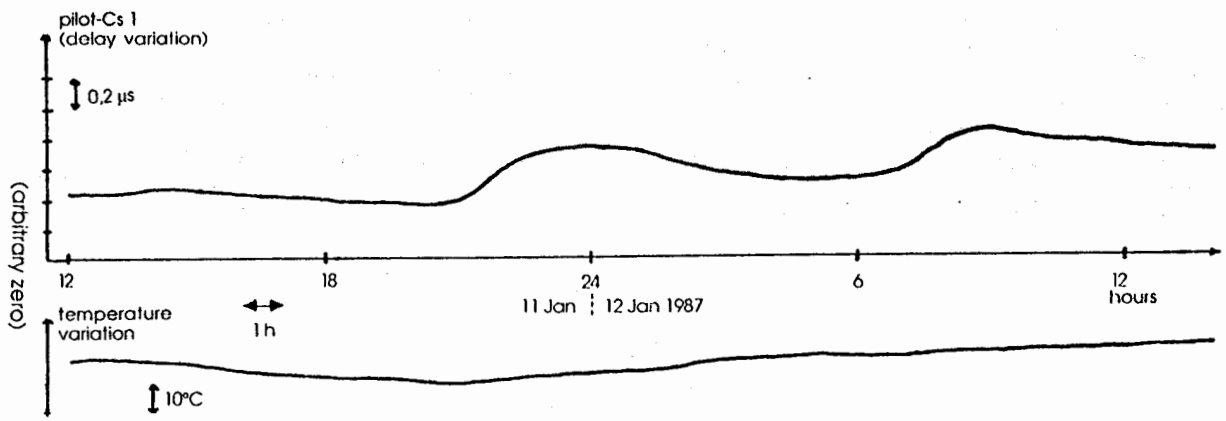


Figure 3: An example of diurnal delay and temperature variation: a period of one day, 11-12 January 1987.

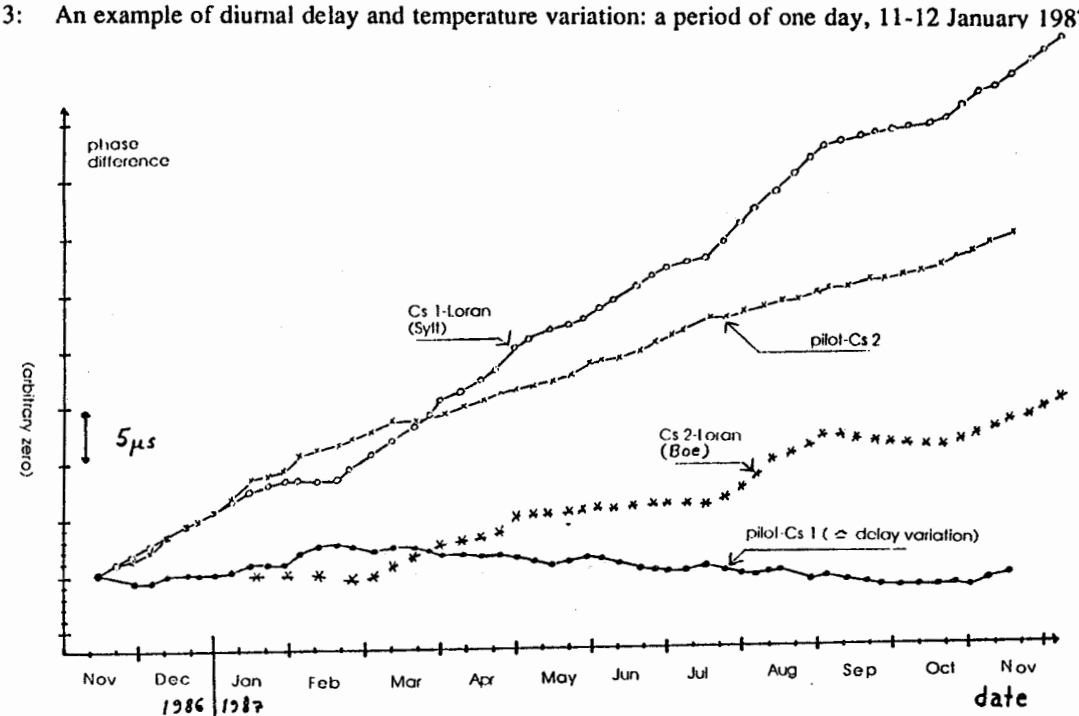


Figure 4: An example of annual delay variation; November 1986 - November 1987. Primary phase comparison results are also shown.

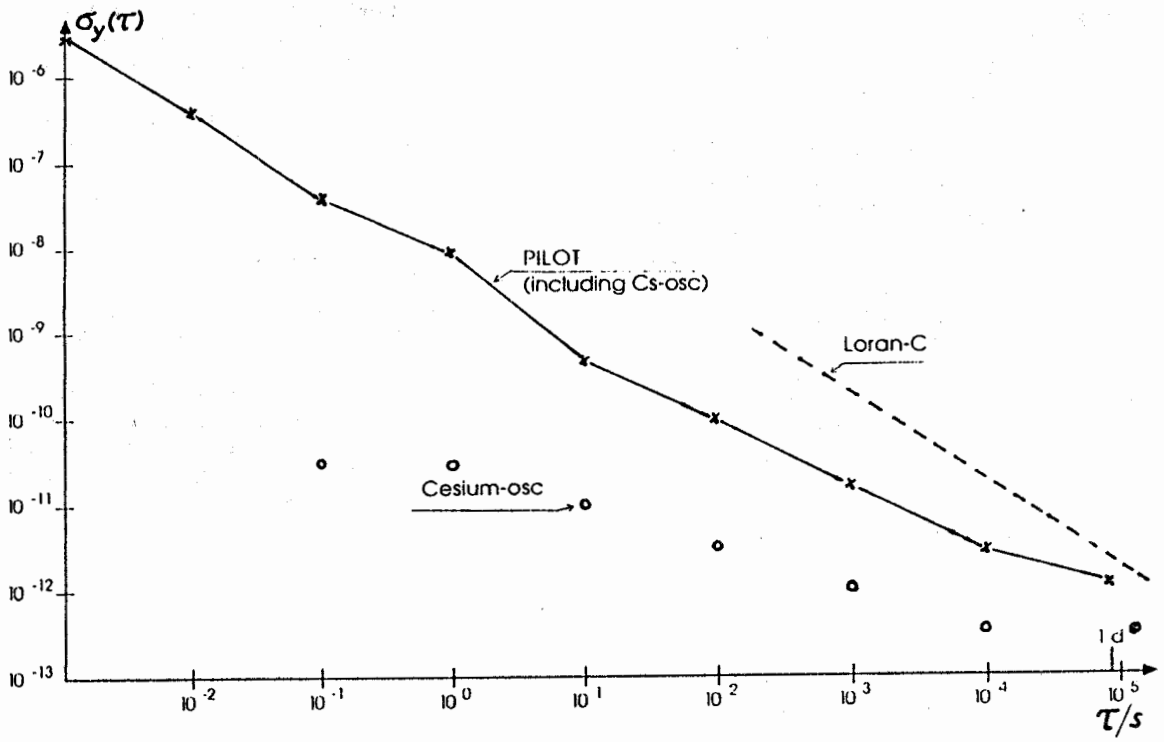


Figure 5: Random or short term delay variation: Square root of Allan variance, $\delta y(\tau)$. Typical values of commercial Cesium-oscillators and Loran-C -receptions are presented as a reference.

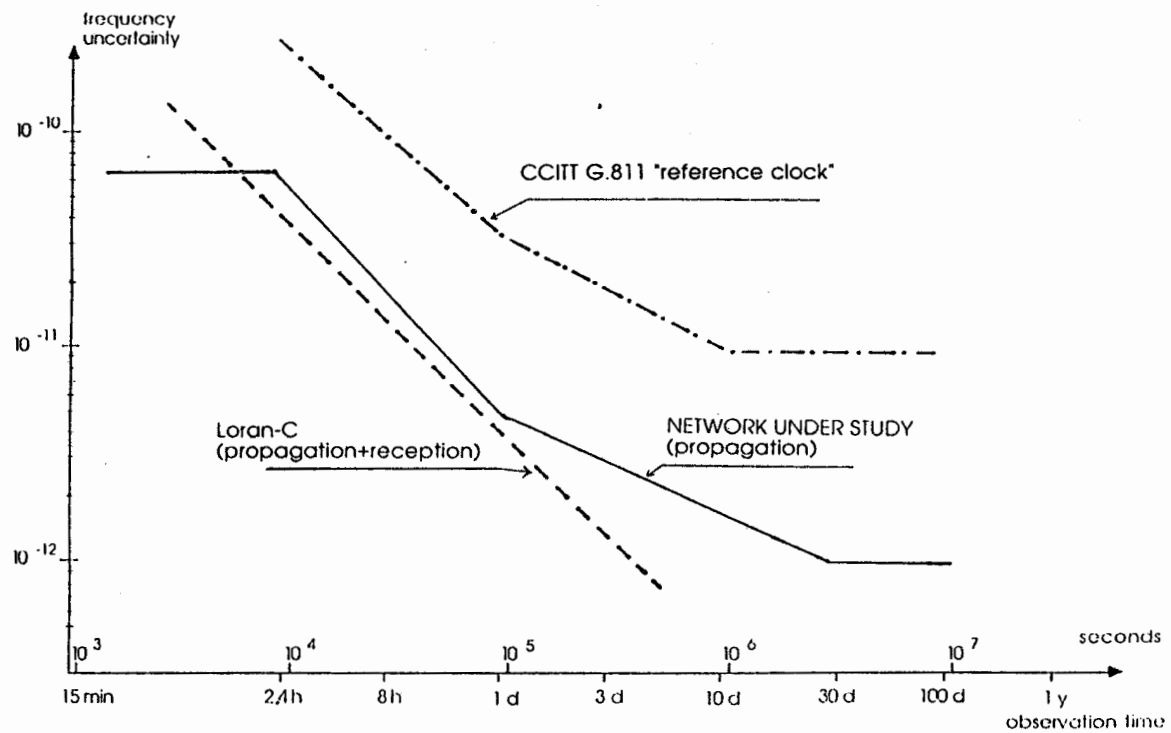


Figure 6: Estimated frequency uncertainties as a function of observation time: network under study ("pilot), Loran-C and CCITT G.811 "reference clock" (recommendation).

References

1. Calvert R.D.: Timing requirements of future ISDNs, proc. of EFTF 88, March 1988, Neuchatel, Switzerland, pp 71-98

2. Bean B.R., Dutton E.J.: Radio Meteorology, NBS Monograph no 92, US Government Printing Office, Washington DC, 1966

3. Winkler G.: Path delay, Its Variations and Some Implications for the Field Use of Precise Frequency Standards, proc. IEEE, vol 60, no 5, May 1972, pp 522-529

4. Davies K. (editor): Phase and Frequency Instabilities in Electromagnetic Wave Propagation, AGARD Conference Proc No 33, Circa Publications, inc, New York, 1970

5. Gray D.A.: Transit-Time Variations in Line-of-Sight Tropospheric Propagation Paths, The Bell System Technical Journal, vol 49, July-August 1970, pp 1059-1068

6. Kartaschoff P.: Frequency and Time, Academic Press, London, 1978

7. Results of soil temperature measurements in Finland 1961-1970, Finnish Meteorological Institute, Helsinki, 1979

8. Wait J.R.: Electromagnetic waves in stratified media, Pergamon Press, Oxford, 1962

9. Hovi M.: Propagation characteristics of radio waves at VHF- and UHF-bands in Finland on the basis of field strength measurement, Master's Thesis, Helsinki University of Technology, 1982, (in Finnish)

SYNTHETISEUR DIGITAL DE FREQUENCE POUR STANDARDS ATOMIQUES DE FREQUENCE

J. VIENNET

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

I. INTRODUCTION

L'asservissement de fréquence d'accord de la cavité d'un maser à hydrogène et l'asservissement d'un pilote à quartz sur la transition atomique du césium nécessitent tous deux de pointer avec la meilleure précision possible le sommet d'une raie de résonance [1,2]. Résonance à une fréquence voisine de 1420 MHz avec une bande passante de l'ordre de 20 KHz pour la cavité du maser à hydrogène et à une fréquence voisine de 9192 MHz avec une bande passante de l'ordre de 500 Hz pour l'horloge à césium.

D'un point de vue purement électronique les deux asservissements relèvent du même principe : exciter la résonance avec un signal d'interrogation modulé en fréquence de façon à élaborer un signal d'erreur permettant de corriger l'élément à asservir.

Nous allons dans la suite présenter une façon de générer numériquement les signaux d'interrogation nécessaires et discuter de leurs caractéristiques.

La synthèse directe de ces signaux n'est pas réalisable dans les gammes de fréquences mises en jeu. Il est nécessaire d'opérer par transposition. Dans le cas du maser nous avons choisi d'effectuer la synthèse sur un signal de fréquence moyenne égale à 405 750 Hz et de transposer cette fréquence par mélange avec un signal de fréquence 1420 MHz obtenu par multiplication de fréquence à partir du pilote à 5 MHz. Dans le cas de l'horloge à césium la synthèse est effectuée autour de la fréquence 131 770 Hz. La transposition sera ensuite effectuée par étapes successives à l'aide de fréquences dérivées du pilote à 5 MHz. La chaine de synthèse est schématisée sur la figure 1.

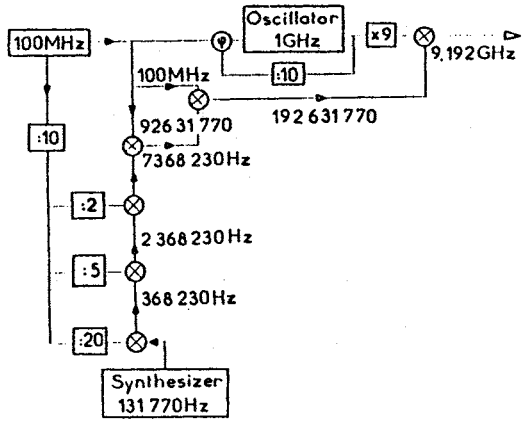


Fig. 1

L'étude de la génération du signal d'interrogation se scinde donc en trois parties qui vont être abordées successivement :

a) La synthèse numérique dans le domaine des basses fréquences de la fréquence moyenne, ou porteuse, du signal d'interrogation.

b) La modulation carrée de fréquence de cette fréquence porteuse.

c) La transposition du signal basse fréquence obtenu dans le domaine des hautes fréquences.

Le principe du synthétiseur à accumulation de phase est représenté sur la figure 2. Il se compose essentiellement d'un registre mémoire de R bits connecté à la sortie d'un additionneur binaire de R bits également. L'additionneur effectue la somme d'une constante K, exprimée en base 2, choisie par l'utilisateur avec la valeur binaire de sortie du registre mémoire. Le registre mémoire est actualisé à chaque période d'horloge $T_c = \frac{1}{F_c}$. La valeur contenue dans le registre s'incrémente donc de la constante K à chaque impulsion d'horloge. L'opération effectuée est décrite par la relation de récurrence :

$$S_n = S_{n-1} + K \text{ (module } 2^R) \qquad (1)$$

où S_i représente le contenu de l'accumulation après la $i^{ème}$ impulsion d'horloge.

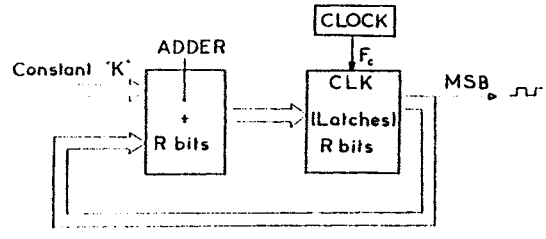


Fig. 2

Le bit de poids fort du registre initialement à zéro, revient périodiquement à zéro à chaque débordement du registre. Le signal rectangulaire obtenu sur ce bit constitue le signal utile de sortie du synthétiseur. A titre d'exemple la figure 3 montre le contenu A du registre accumulateur en fonction du temps et le signal S(t) obtenu sur le bit de poids fort, dans le cas particulièrement simple où R = 4 et k = 5.

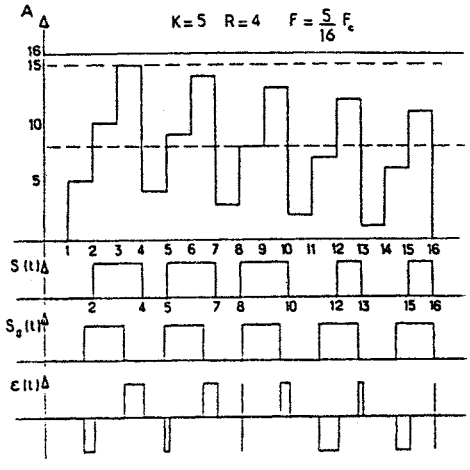


Fig. 3

II. 1. ANALYSE DU FONCTIONNEMENT DU SYNTHETISEUR

Le registre accumulateur est constitué de R bits repérés par leurs poids respectifs variant de 2^0 à 2^{R-1}. Le premier bit de poids faible non nul de la constante K définit la taille effective du registre accumulateur. En effet les zéros éventuels de poids faibles n'interviennent pas dans le processus d'accumulation et restent à zéro. La constante K est donc par principe toujours impaire. La figure 4a représente le registre accumulateur lorsque, après le premier coup d'horloge, il contient la valeur K. La figure 4b représente ce même registre après 2^R périodes d'horloge.

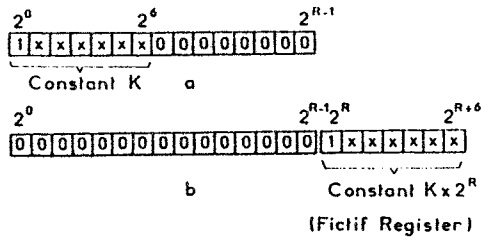


Fig. 4

Initialement le registre accumulateur est à zéro. La constante K s'y ajoute à chaque front d'horloge. Après 2^R périodes d'horloge, l'accumulateur revient à l'état initial et un cycle identique recommence. Le contenu de l'accumulateur est donc une fonction périodique du temps de période :

$$T_M = 2^R T_c \qquad (2)$$

et donc de fréquence :

$$F_M = \frac{F_c}{2^R} \qquad (3)$$

II. 2. CARACTERISTIQUES DU SIGNAL DE SORTIE

II.2.1 Etude temporelle

Il est clair qu'il faut en moyenne $\frac{2^R}{K}$ périodes d'horloge pour remplir l'accumulateur et l'amener au débordement. La fréquence fondamentale moyenne du signal de sortie est donc donnée par :

$$F = \frac{K}{2^R} F_c \qquad (4)$$

Dans les cas particuliers où la quantité $\frac{2^R}{K}$ est entière cette expression donne effectivement la fréquence instantanée du fondamental du signal de sortie. Dans le cas général la quantité $\frac{2^R}{K}$ n'est pas entière et l'expression n'est valable qu'en moyenne sur une période $T_M = 2^R T_c$. Dans les deux cas la résolution de fréquence est donnée par une variation de K de plus ou moins une unité. Elle est égale à :

$$\Delta F = \frac{F_c}{2^R} \qquad (5)$$

Le signal de sortie n'est donc pas dans le cas général un signal carré. On peut considérer qu'il est la somme d'un signal carré dont la fréquence fondamentale est donnée par (4) et d'un signal impulsionnel perturbateur responsable du "jitter" de phase. La figure 3 montre sur le cas particulier déjà considéré où R = 4 et K = 5 que le signal de sortie S(t) est la somme du signal carré $S_o(t)$ de fréquence $F = F_H \times \frac{5}{16}$ et du signal impulsionnel perturbateur $\varepsilon(t)$. Les impulsions perturbatrices du signal $\varepsilon(t)$ ont toujours une largeur inférieure à une période d'horloge. Leur influence est donc faible si la fréquence d'horloge est grande vis à vis de la fréquence F du signal de sortie.

II.2.2 Etude fréquentielle

Le spectre du signal s(t) est constitué par la superposition du spectre du signal carré désiré et du spectre du signal perturbateur. Le signal perturbateur dépend de la constante K choisie donc de la fréquence de sortie désirée. Il n'est donc pas possible de calculer son spectre dans le cas général. Il est cependant possible de préciser ses caractéristiques essentielles.

Le signal perturbateur est périodique de période $T_M = 2^R T_c$. Sa fréquence fondamentale est égale à $\frac{F_c}{2^R}$ et les raies de son spectre sont régulièrement espacées de la même quantité. La fréquence harmonique d'ordre K correspond à la fréquence du signal utile de sortie.

Le signal perturbateur est constitué d'impulsions de largeur τ variable entre 0 et T_c, régulièrement réparties sur une période T_M, la deuxième demi période de $T_M/2$ à T_M se déduisant de la première par une translation d'une durée égale à $T_M/2$ et par un changement de signe. Les composantes de Fourier C_n du spectre d'une paire d'impulsions périodiques de largeur τ, d'amplitude h, de signes opposés et centrées aux instants t_o et $t_o + T_M/2$ sont données par l'expression :

$$C_n = 2h \frac{\tau}{T_M} \frac{\sin \pi n \frac{\tau}{T_M}}{\pi n \frac{\tau}{T_M}} \times \sin \frac{\pi n}{2} \times e^{-j\frac{\pi n}{2}} \qquad (6)$$

Cette expression montre que les composantes spectrales correspondant aux valeurs paires de n sont nulles et que le spectre est très étendu. La première valeur de n pour laquelle la première composante spectrale impaire est égale à zéro est donnée par :

$$n_o = \frac{T_M}{\tau} \qquad (7)$$

Puisque $\tau < T_c$ et $\frac{T_M}{T_c} = 2^R$ il vient : $n_o > 2^R$. Comme $2^R >> K$, la valeur du sinus cardinal de la relation (6) est très voisine de 1 dans une large plage de fréquence autour de la fréquence de sortie désirée.

Chaque raie du spectre du signal perturbateur est donc constituée par la somme vectorielle d'un grand nombre de composantes de module moyen de l'ordre de $\frac{h}{2^{R-1}}$ et de phases équitablement réparties sur l'intervalle 0, 2π. L'amplitude de la raie résultante est dans le cas générale très inférieure à $\frac{h}{2^{R-1}}$.

L'intérêt d'avoir R grand et donc d'utiliser une fréquence F_c aussi grande que possible apparait clairement ici.

II. 3. REALISATION

Nous avons réalisé une maquette comportant un ensemble registre additionneur sur 32 bits. Notre choix s'est fixé sur la technologie MECL 10 KH de Motorola qui offre un bon compromis rapidité-souplesse d'emploi. L'additionneur comporte 8 circuits type MC 10 H 181 associés à deux générateurs de retenue anticipée du type MC 10 H 179. Il permet d'effectuer une opération d'addition sur 32 bits en moins de 20 ns ce qui autorise une fréquence d'horloge de 50 MHz. Le registre mémoire est réalisé à l'aide de bascules type D intégrées par 6 dans le circuit MC 10 H 186.

La constante K est appliquée à l'additionneur à l'aide de registres adressés par un microprocesseur TMS 70 C 82 qui gère un ensemble clavier afficheurs LCD et qui assure diverses fonctions de contrôles et de commandes sur l'horloge liée au synthétiseur. Le schéma de principe de la réalisation est représenté sur la figure 5.

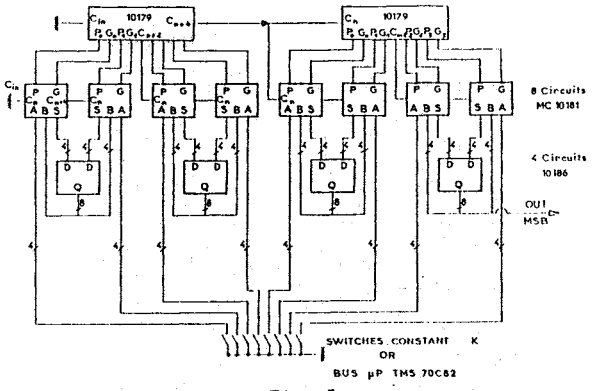


Fig. 5

II.3. 1 Performances

La figure 6 montre un spectre caractéristique du signal obtenu en sortie du synthétiseur pour des fréquences voisines de 405 KHz. On remarque que le signal est toujours à plus de 70 dB au-dessus du plancher de bruit, ce qui est convenable pour les applications envisagées. Cette figure montre que pour certaines fréquences, il peut exister quelques raies parasites à plus 60 dB sous le signal. Ces raies généralement éloignées de la fréquence désirée peuvent être éliminées sans trop de difficultés.

L'ensemble du circuit est câblé sur une carte 10 x 16 cm environ. La consommation relativement élevée, puisqu'il s'agit de circuits rapides, ne dépasse cependant pas 10 W.

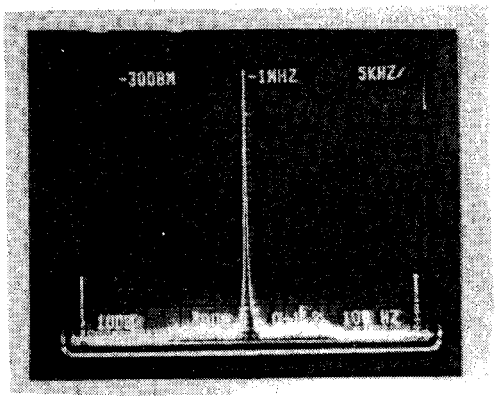


Fig. 6

III. GENERATION NUMERIQUE DU SIGNAL D'INTERROGATION MODULE

III.1. PRINCIPE

Considérons (figure 7) un signal s(t) périodique de pétiode T, constitué par n_1 périodes d'un signal sinusoïdal de fréquence F_1 suivies d'un signal similaire de n_2 périodes de fréquence F_2.

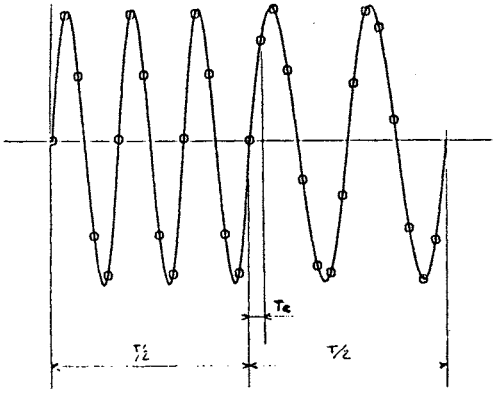


Fig. 7

Les deux demi-périodes étant de valeurs égales, nous avons :

$$\frac{n_1}{F_1} = \frac{n_2}{F_2} = \frac{T}{2}$$

La fréquence moyenne du signal est égale à :

$$F_o = \frac{F_1 + F_2}{2}$$

et on a :

$$F_1 = F_o + \Delta F$$
$$F_2 = F_o - \Delta F$$

ΔF étant l'excursion de fréquence de part et d'autre de la porteuse F_o.

Il est clair que le signal s(t) est un signal modulé en fréquence par un signal carré tel que celui que nous désirons pour interroger la résonance. L'indice de modulation m est égal à :

$$m = \Delta F \times T$$

Les relations liant n_1 et n_2 aux paramètres caractéristiques du signal s(t) sont donc :

$$n_1 + n_2 = F_o T$$

$$n_1 - n_2 = m$$

Ces relations permettent de déterminer facilement les nombres de périodes n_1 et n_2 en fonction des caractéristiques du signal s(t) désiré.

Pour générer le signal s(t), il suffit de l'échantillonner sur une période T, de placer les N échantillons successifs dans une table inscrite en mémoire PROM, et de lire cette table de façon séquentielle à l'aide d'un compteur commandé par un signal de fréquence $F_e = F_o \times \frac{N}{n_1+n_2}$. Les échantillons ainsi obtenus sont appliqués à un convertisseur digital-analogique qui restitue le signal s*(t) correspondant au signal s(t) échantillonné bloqué.

III.2. CARACTERISTIQUES SU SIGNAL OBTENU

Le signal obtenu s*(t) ne reproduit pas exactement le signal s(t) désiré à cause de la quantification opérée sur la valeur de chaque échantillon pour la mise en mémoire. L'effet de quantification est de superposer au signal original s(t) un signal d'erreur $\varepsilon(t)$ de telle sorte que

$$s^*(t) = s(t) + \varepsilon(t)$$

La figure 8 représente ces trois signaux dans le cas d'une quantification par arrondi sur 8 niveaux (3 bits).

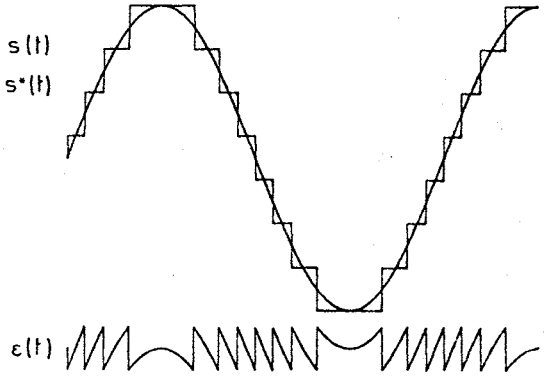


Fig. 8

Le signal $\varepsilon(t)$ est comme le signal s(t), périodique de T. Son spectre est donc un spectre de raies distantes entre elles de $\Delta f = \frac{1}{T}$ et d'amplitudes proportionnelles au pas de quantification. Il est bien évident que ces raies vont se superposer aux raies du spectre du signal s(t). Leur contribution est cependant très faible si le pas de quantification est petit et si la fréquence d'échantillonnage F_e est grande devant la fréquence f_o. En effet, l'examen du signal $\varepsilon(t)$ montre sans ambiguité possible que la majeure partie de l'énergie du signal $\varepsilon(t)$ se trouve au voisinage de la fréquence F_e.

III.3. REALISATION PRATIQUE

Le schéma de principe de la maquette réalisée est représenté figure 9. La maquette comporte une mémoire PROM de 256 mots de 16 bits, composée de deux boitiers 63 S 281, pour le signal d'interrogation de la cavité du maser à hydrogène, ou une mémoire PROM de 4096 mots de 16 bits, composée de deux boitiers 63 S 3281, pour le signal d'interrogation de la résonance du césium.

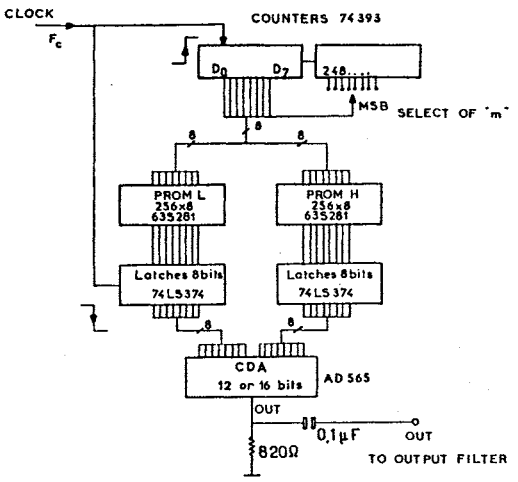
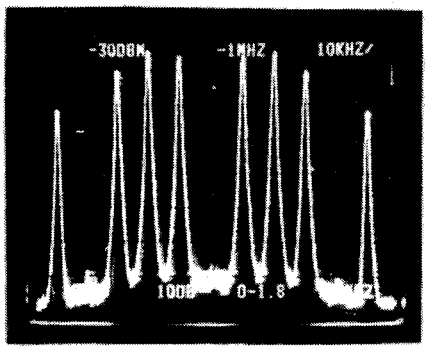


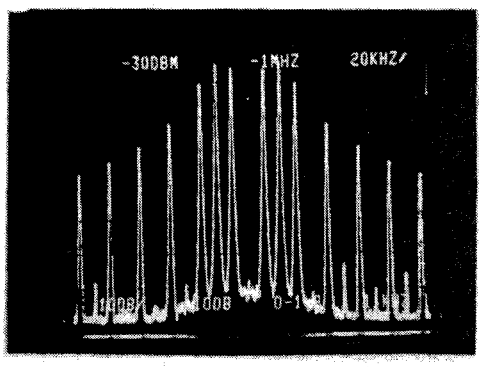
Fig. 9

Cette mémoire est adressée par deux compteurs binaires 8 bits type 74393. Les bits de poids faibles de la mémoire sont connectés aux bits de poids faibles du compteur. Le bit de poids fort est commutable sur une des huit bascules du compteur poids fort de façon à pouvoir choisir l'indice de modulation "m" suivant les puissances de deux. Les données de sortie des PROM sont verrouillées dans un ensemble de deux circuits 74 LS 374. Les compteurs sont actualisés sur les fronts montants de l'horloge alors que les circuits de verrouillage sont actualisés sur les fronts descendants. On dispose de cette manière d'une demi-période d'horloge pour rendre les données valides à la sortie des mémoires avant de les prendre en compte. Le convertisseur digital analogique est câblé à la sortie des circuits de verrouillage.

Les performances globales du système dépendent beaucoup des qualités du circuit convertisseur C.D.A. utilisé. Les spectres représentés figure 10 montrent le résultat obtenu pour une fréquence de sortie de 405 KHz environ, un indice de modulation égal à 2 et une excursion de fréquence de l'ordre de 37 KHz, avec un C.D.A. de 12 bits type AD 565 de la firme Analog Devices. La figure 10a montre le spectre obtenu sur une plage relativement large et la figure 10b détaille la partie centrale du même spectre. On remarque que le plancher de bruit est à environ 75 dB sous le niveau des raies principales et qu'il existe quelques raies parasites, dont le niveau est inférieur à - 60 dB, relativement écartées de la fréquence centrale. La raie centrale, si elle existe, est noyée dans le plancher de bruit et en tous cas inférieure à - 65 dB.



a



b

Fig. 10

BIBLIOGRAPHIE

1. C. Audoin, Fast Cavity Auto-Tuning Systems for Hydrogen Maser. Rev. Phys. Appl. 16, Mars 1981, pp. 125-130
2. N. Yahyabey, R. Barillet, J. Viennet et C. Audoin. Accord automatique de la cavité résonnante d'horloges à hydrogène. Actes du 1er Forum Européen "Temps-Fréquence", Besançon 18-20 Mars 1987, pp. 334-338
3. J. Tierney, C. Rader, B. Gold. A Digital Frequency Synthesizer. IEEE Transactions on Audio and Electroacoustic. Vol. AU-19, n° 1, March 1971
4. H.W. Cooper. Why Complicate Frequency Synthesis ? Electronic Design 15, July 19, 1974
5. B.E. Bjerede, G.D. Fischer. A New Phase Accumulator Approach to Frequency Synthesis. NAECON'76 RECORD, pp. 928-932
6. E.C. Kisenwether, W.C. Troxell. Performances Analysis of the Numerically Controlled Oscillator. Proc. of 40th Annual Frequency Control Symposium, 1986, pp. 373-378

Remerciements

Ce travail a été réalisé grâce au soutien financier de la DRET dans le cadre de la convention n° 87.34.051. 00.570.75.01.

ORIGINS OF DISLOCATIONS IN SINGLE CRYSTAL CULTURED QUARTZ

Gary R. Johnson, Robert A. Irvine, Jonathan W. Foise

Sawyer Research Products, Inc.
Eastlake, Ohio 44094 U.S.A.

Abstract

Previous work to identify, characterize and reduce etch channels in single crystal cultured quartz is reviewed. Experiments are reviewed confirming that etch channels result from selective etching near edge dislocations. The dominant source of the dislocations in the grown crystal is determined to be inclusions at the seed crystal interface. Using x-ray topography, optical and scanning electron microscopy, the composition of the interface is seen to consist of several inclusion species which may coexist with bands of dislocations identified as misfit dislocations by Burgers vector invisibility criteria. The analysis of crystals grown in a high purity environment on specially selected seeds is discussed.

Introduction

Device Considerations

Applied as a frequency control or timing device, components fabricated from single crystal cultured quartz are widely used in military, industrial and consumer electronic equipment. New designs continually offer higher performance, miniaturization and lower costs. To answer the need for device improvements, quartz component manufacturers implemented processes requiring the exposure of quartz to strong etchants. In some cases, this was an effort to adapt the batch photolithographic techniques of the semiconductor industry to quartz resonator production, offering lower costs and realization of innovative designs. In other instances, the design makers were trying to reduce the subsurface damage caused by previous lapping operations to improve performance. On still other occasions, the fabricators were trying to achieve very thin sections for high frequency operation. Each of these efforts was constrained by the formation of micron sized holes (often referred to as etch channels) in cultured substrates on exposure to etchants.

Figure 1 is typical of the surface condition following prolonged etching in ammonium bifluoride (ABF).

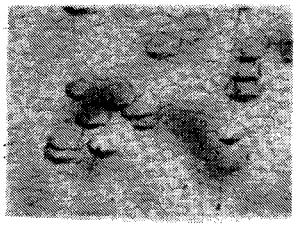


FIGURE 1

Note the appearance of pits with well defined inverted apices which sometimes serve as the starting point for long narrow channels. Higher magnification by scanning electron microscopy (SEM) is shown in Figures 2 and 3.

FIGURE 2

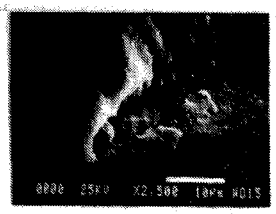


FIGURE 3

Etch channels limit device makers by:

1. Creating electrical faults upon application of conductive electrodes.

2. Acting as stress concentrators to reduce the ultimate strength of the device.

3. Trapping contamination resulting in increased aging.

The first widespread application of photolithographic processing techniques occurred in the manufacture of tuning fork resonators. An example of a device with size about 1 mm x 5 mm is shown in Figure 4.

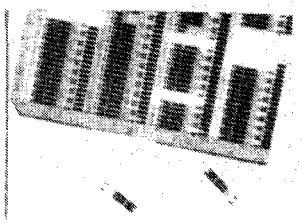


FIGURE 4

Thicksten (1983) catalogued various etch related defects encountered in fabricating tuning forks. Figure 5 illustrates the presence of etch pits and channels in photolithographic processed parts.

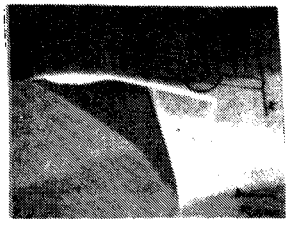


FIGURE 5

In some cases, electrical discontinuities are caused by voids in the metalization attributable to etch defects. Additionally, the shock resistance of the device is severely impaired by the presence of stress concentrators. As devices undergo further miniaturization and more complex shapes are required, the need to reduce the concentrations of pits and channels increases.

The direct effect on component performance of dislocations without an associated channel is not well documented. However, Zheng (to be published in 1989) provides experimental evidence of some effect in discussion of unconfined vibrational modes around dislocations, observed with stroboscopic x-ray diffraction methods.

Defect Characterization and Origin

The well known method of etching to detect the presence of dislocations in metals and crystals was applied to cultured quartz early in its development and remains a standard characterization technique. The formation of etch pits in the presence of dislocations is predicted by free energy considerations. Cabrera, Levine and Plaskett (1954) describe the conditions necessary to open a dislocation into an etch pit when the crystal is placed in an undersaturated medium. The basic consideration is an energy balance between the dissolution of a volume of crystal into the undersaturated medium and the release of potential energy stored as strain near the dislocation balanced against formation of additional surface area free energy. Two of these terms are nearly the same throughout the bulk of the crystal. However, the increased strain energy near the dislocation results in much higher etch rates in those areas.

The qualitative similarity between patterns formed by Lang topographs and etching is readily seen in Figure 6.

A) X-RAY TOPOGRAPH B) PHOTOGRAPH AFTER ABF ETCHING

FIGURE 6

However, in quartz it was noticed that in addition to the etch pits which existed in other materials, small areas, beginning at the apices of some of the pits, etched much faster than the surface, creating channels extending well below the surface. Arnold (1957) observed that very large (over 250 microns) inclusions related to the formation of etch channels. At about the same time Augustine (1957) and Augustine, Hale and Berry (1957) were conducting similar etching experiments and reporting similar observations.

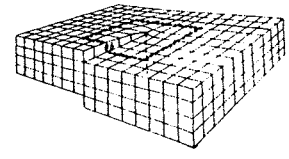
The composition of channels in natural and cultured quartz was investigated in Buerger (1954) and Cohen (1960). The possibility of stuffing quartz with beta-eucryptite and beta- spodumene to explain the presence of large quantities of certain impurities was mentioned.

Nielsen and Foster (1960) conducted further etching experiments with similar results. They noted the presence of fluid inclusions spanning the thickness [0003] of the seed plate and a one-to-one correspondence between these inclusions and etch channels in the grown crystal. They further discussed possible causes of these channels, mentioning dislocations, nonuniform incorporation of impurities and a combination of these two possibilities. They ultimately decided that they could come to no conclusion.

Further important etching studies were carried out by Tsinzerling and Mieonua (1963) and Hanyu (1964). This work concluded that the faces of the surface pits matched naturally occurring faces of quartz and attributed the pits to surface intercepts of dislocations.

Spencer and Haruta (1966) applied the transmission x-ray diffraction topographic methods developed by Lang to single crystal cultured quartz. This study established that these linear defects had some characteristics of dislocations and if they were dislocations, they were most likely screw or mixed types.

Line drawings illustrating the general idea of screw and edge dislocations are shown in Figures 7 and 8.



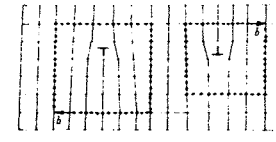


FIGURE 7 SCREW DISLOCATION FIGURE 8 EDGE DISLOCATION

On a unit cell scale, the screw dislocation appears as an offset creating a spiral staircase effect. An edge dislocation is visualized as an extra plane of atoms. The Burgers vector is labeled b.

If g is the reciprocal lattice vector of a set of planes and b is the Burgers vector of the dislocation, the contrast effect of a dislocation image will disappear when $g \cdot b = 0$. Conversely, the contrast effect will maximize when $g \cdot b = 1$. This implies that the image at dislocation lines exhibits the maximum when the reflecting planes are oriented perpendicular to the Burgers vector and vanish when they are parallel to b. In the case of a pure edge dislocation, b and the dislocation line both lie in the same plane, the axis of the edge is obtained from the criteria that $g \times L = 0$, where L is the direction of the dislocation line.

An extensive study of dislocations by Lang and Miuscov (1967) confirmed the link between etch pits and dislocations. Specifically, they concluded that with rare exception, pits formed at dislocations and that no pits occurred not at dislocations. They further decided that based on contrast effects from various planes that only about 15% had C axis component [0003] associated with screw dislocations. Presumably the remaining 85% exhibited strong edge dislocation characteristics.

Lang and Miuscov further observe that dislocations are randomly oriented early in the growth cycle, but that dislocations soon organize into a polygonal cell boundary network (Figure 6 illustrates), the final stage of which appears as growth hillocks on the rough final surface with orientation approximately (0001). Fault surfaces were observed in the topographs which were interpreted as indicative of impurity segregation in the cell walls.

Further confirmation of the relationship between etch pits, etch channels and dislocations was presented by McLaren, Osborne and Saunders (1971). Once again it was concluded that most of the dislocations are of the edge variety, some mixed and rarely finding a pure screw dislocation. As in other studies, the dislocations were confined to a narrow (10°-25°) cone, centered about [0003]. The topographs produced illustrated dislocations originating in the seed or from the seed surface, propagating in the general direction of growth. A suggested model for a pure edge dislocation is presented in Figure 9.

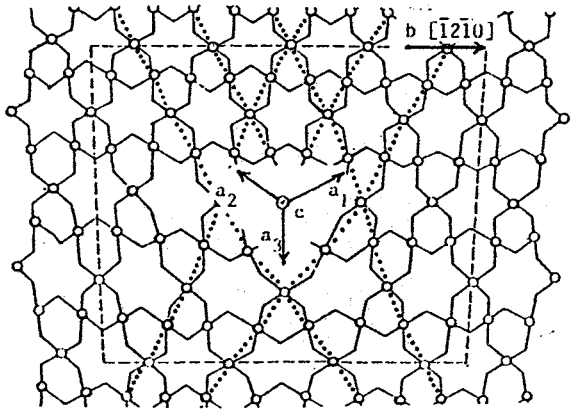


FIGURE 9

An important feature of this model is a hollow core with a series of unbonded oxygens. It was hypothesized that this was an obvious place for the gathering of impurities, especially hydrogen and substitutional aluminum (Al^{+3}).

Iwasaki (1977) draws the same association between edge dislocations, etch pits and etch tunnels as previous researchers. Another structural model is proposed with the key feature of unbonded oxygen surrounding a hollow core as presented by McLaren et al. (1971). However, the problem of transporting etchant to the reaction interface down the micron sized channel is discussed. The mechanism posed is to assume vapor etching by the outgassing of an acid with higher mobility along the channel than in the case of liquid.

Barns et al. (1978) provided additional confirmation that etch channels and pits are related to dislocations. They also observed certain shallow flat bottom pits which they associated with flaws created during grinding or lapping. However, the majority of their report concerning pits described sharply defined apices and etch channels; the same as mentioned by earlier observers. This research failed to uniquely determine the Burgers vector, thus leaving open the question of the character of the dislocation.

Etch channels were separated into two categories:

1. Those propagating from fluid inclusions in the seed and

2. Those emanating from particulate inclusions near the seed.

Homma and Iwata (1973) provided another link among observations with etching, Lang topography and polygonal cell structure. They also applied ion probe methods to samples taken within tens of microns of the as-grown Z-surface and detected a large concentration of aluminum and iron in a 10 micron area of high dislocation density between cell walls. The boundary of high dislocation density observed by Lang and Miuscov (1967) was not seen in this work.

Saha, Annamalai and Bandyopadhyay (1979) conclude that about 80 to 85% of the dislocations in cultured quartz are pure edge dislocations constrained to a cone of about 25° around the C axis. These dislocations congregate along cell wall boundaries which are also a region of high impurity segregation. Additionally, the remainder of the dislocations are predominantly of screw character, each of them associated with a single terraced cobble apex.

The various relationships shown by the cumulative work is briefly illustrated in Figure 10.

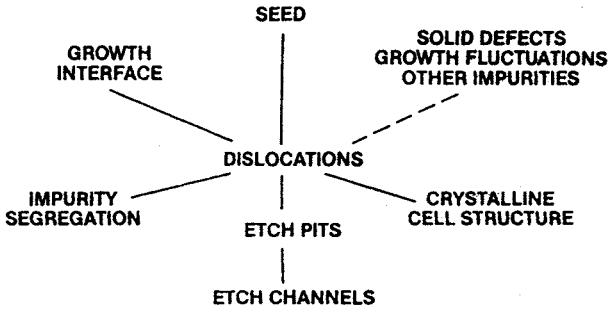


FIGURE 10

Efforts to Reduce Dislocations

The main thrust of the Barns et al. (1978) work was to develop a reproducible technique for the growth of low dislocation density cultured quartz. Based on their analysis of inclusions, showing that iron was a critical component in inclusions which created dislocations, combined with knowledge that dislocations present in the seed would propagate, they found their best results using low dislocation seeds in a noble metal liner to isolate the growth of the crystals from steel autoclaves. Standard Brazilian nutrient was employed as were typical growing conditions. Analysis indicated that all dislocations formed etch pits. Though no analysis was reported, they indicated the dislocations decorated with iron formed channels.

They reported low dislocation material and concluded that it was possible to create dislocation free material under normal growth conditions in noble metal lined vessels if nearly perfect seeds are used. They further stated that across a broad growth rate variation there was no relationship to dislocation density. Crystals grown on (01$\bar{1}$1) seed plates were found more prone to dislocations than those grown on (0001) seed plates in this study.

The problem of obtaining suitable seeds was addressed by Zarka, Lin and Buisson (1981). Previous research consistently illustrated propagation of dislocations nearly parallel to the direction of growth. Zarka et al. chose seeds from the greater X-region with growth direction [2$\bar{1}$$\bar{1}$0] to minimize dislocations in the [0003] direction of intended growth. This lowered the dislocations propagating from the seed, although the topographs show that considerable dislocations nucleated from inclusions in close proximity to the seed.

In a series of reports prepared by Armington et al. (1981, 1982, 1984, 1985, and 1986) the topics of lowering impurities, especially aluminum, and reduction in dislocation density are simultaneously addressed. The summary of the result of this activity is shown below in Figure 11.

Experimental Variable	Effect on Dislocations
Growth Environment	none
Noble liner	
Unlined	
Mineralizer	
Seed Source	strong
Low defect natural	
Greater X cultured	
"Normal" cultured	
Low strain	
Supply Source	none*
High purity cultured	
High purity glass	
High purity sand	
Seed Preparation	none
Acid etch	
Hydrothermal etch	
Swept	
Growth Rate	none*

*Showed strong effect on aluminum uptake.

FIGURE 11

In contrast to prior investigations, the 1985 report indicated that most dislocations originated within seeds supplemented by a small number nucleating at inclusions. The 1986 report indicated that typical etch channel densities of under 50 cm^{-2} were possible using greater X seeds and pure Z cultured quartz nutrient. Suitable results were realized across a variety of vessel surface materials, mineralizers, growth rates, and seed preparation methods.

Croxall et al. (1982) created quartz with dislocation density of under 10 cm^{-2}. The method was based on employing a noble metal vessel surface, high purity nutrient and growtn solution, low dislocation seeds, and standard growing conditions appropriate to sodium hydroxide. They reported the appearance of dislocations in the seed crystal interface, but reported no dislocation generation at inclusions.

Post Culturing Processing

Although the complete elimination of dislocations is desirable, most devices can tolerate the presence of dislocation densities of 10^3 cm^{-2} or greater and many photolithographic uses are feasible with densities of 10^2 cm^{-2}. However, in the case of certain devices with relatively large dimensions and a low tolerance for dislocations (especially etch channels), processes designed to reduce the formation of etch pits and channels in the presence of dislocations have been developed. Electrolytic sweeping as a method to inhibit etch channel formation has been reported by Vig et al. (1977), Martin et al. (1983), Gaultieri (1985), and Armington and Balascio (1985). Large reductions are consistently reported with Balascio and Armington (1986) reporting a decrease in mated half bars from 328 cm^{-2} to 36 cm^{-2}. They further reported an incidence of a reduction from 24 cm^{-2} to 1 cm^{-2} in crystal grown on a greater X seed.

Another technique reported by Bernot (1985) concerns the addition of various fluoride salts to the normally used acid etchants. This work showed that the addition of KF along with unnamed secondary additives to either HF acid or ABF greatly reduced the occurance of etch pits and etch channels. Both of these results were reported using unswept commercially available quartz.

The motivation for developing these further steps was the desire to fabircate a precise altimeter requiring an etch channel free active area and a chemically polished surface. This combination is a severe requirement for presently available material and requires swept high quality material and special etchant to achieve successful fabrication. This device is shown in Figure 12 (U.S. patent 4,479,070).

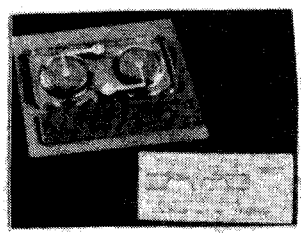


FIGURE 12

Identify and Catalogue Dislocation Species

Lattice defects in Y and Z cut cultured quartz crystals were investigated by x-ray topography. Burgers vector was determined and structural inferences were made by examination of the contrast effect viewed in topographic images reflected from a variety of planes.

Topographs on a Y cut crystal grown at an average growth rate of about 0.4 mm day^{-1} under standard commercial growth conditions for the sodium carbonate process were prepared. The results, using several values of g, are shown in Figures 13 to 19.

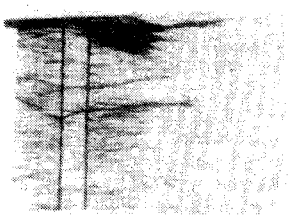


FIGURE 13 $[\bar{1}2\bar{1}0]$

FIGURE 14 $[01\bar{1}0]$

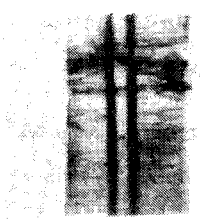


FIGURE 15 [0003]

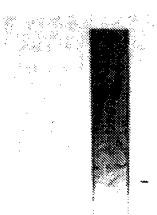


FIGURE 16 $[\bar{1}011]$

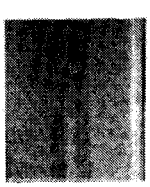


FIGURE 17 $[10\bar{1}1]$

FIGURE 18 $[10\bar{1}2]$

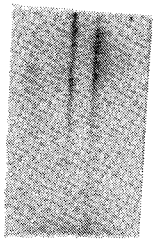
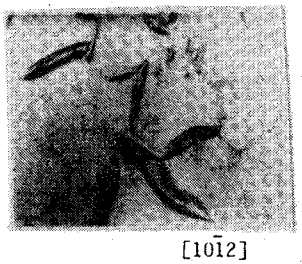


FIGURE 19 $[10\bar{1}0]$

Examining the figures, it is seen that the maximum contrast was obtained for $[\bar{1}2\bar{1}0]$ and the minimum for $[10\bar{1}0]$. Consequently b of the dislocation was $[\bar{1}2\bar{1}0]$. Carrying out gxl shows that the direction is [0003], parallel to the C-axis. Inspection of the topographs shows that most of the dislocation lines follow this course within a cone of about 10°. The topographs further reveal that the dislocation lines either begin in the seed or at the seed crystal interface. The interfaces were consistently characterized by dark lines parallel to the X axis $[10\bar{1}0]$. The dislocation density at the interfaces was high and individual resolution was not possible. The nature of the dislocation structure emerged, however, when the reflections approached the $[10\bar{1}0]$ direction which minimized the contrast of most of the edge dislocations. Figures 17 and 18 exhibit strong contrasts, indicating a dense dislocation structure which remains in high contrast at all reflections without definite Burgers vector. Presumably misfit dislocations bridge any lattice misorientation or mismatch between the seed and crystal. Consequently, the direction of Burgers vector varies depending on local conditions. It is likely this is a highly strained region, potentially decorated by impurity atoms.

The dislocation cell structure discussed by Lang and Miuscov (1967) was also observed in this study. The observed diffraction contrast effects associated with the wall structure suggested that the cell walls are depository sites for impurities.

Another observation of this study, indicating a strong possibility of the impurity segregation at the cell walls, concerns the presence of fault surfaces. Figure 20 is a Z cut with $g = [10\bar{1}2]$.



$[10\bar{1}2]$

FIGURE 20

These diffraction effects reveal large local strain accumulation. The combination of internal surfaces and local strain intensification are factors conducive to impurity segregation.

Shadowgraph images are an alternative method for detecting line defects in quartz. This is a common technique in the optics field to contrast slight variations in the refractive index of transparent materials. The low cost and convenience are an attractive alternative to the costly (and destructive) etching and x-ray topographic tests. However, the results are qualitative, while the other methods can yield quantitative determinations, Figure 21 illustrates the experimental setup. Figure 22 shows typical results.

SHADOWGRAPH IMAGING TECHNIQUE

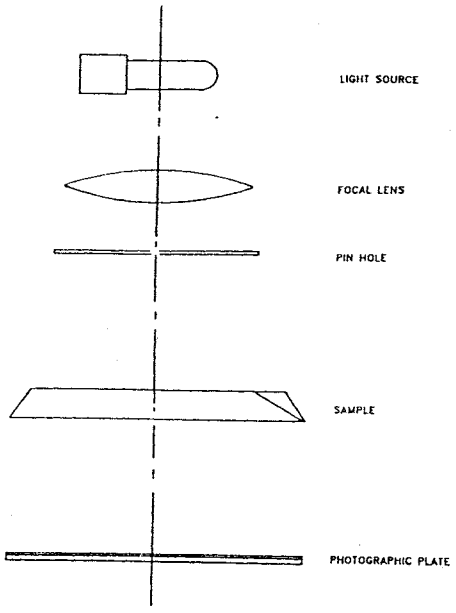


FIGURE 21

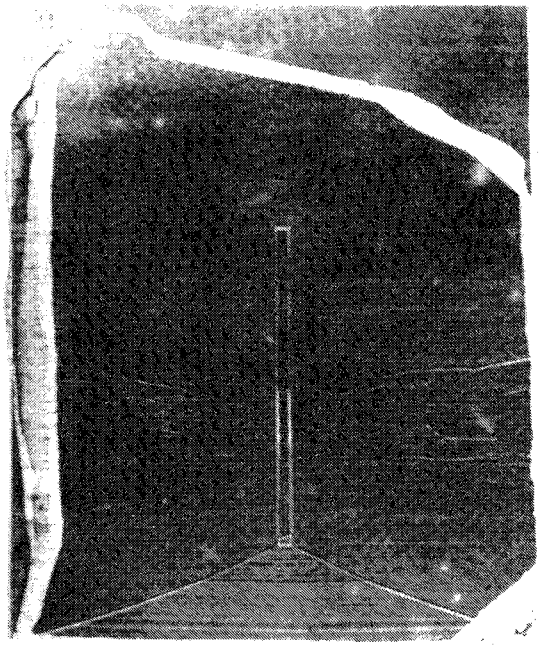


FIGURE 22

Bulk Inclusion Study

In 1986, experiments were conducted under funding of the National Science Foundation to determine the relationship, if any, between dislocations and inclusions (Johnson and Irvine, 1987). The first step was to correlate bulk inclusion densities with etch channels. Figure 23 shows the results of the initial plot generated from a data base of similar run types grown under nominal production conditions.

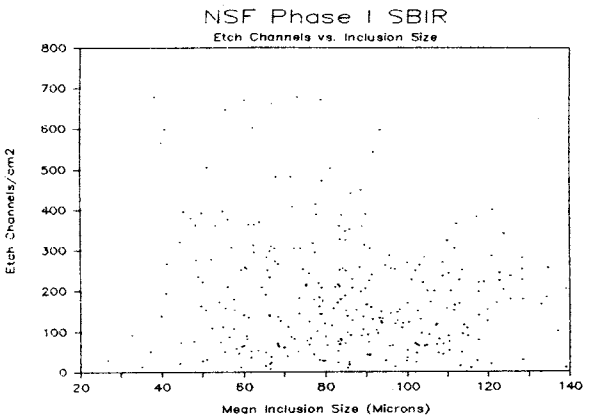


FIGURE 23

Identify and Catlogue Inclusion Species

Since little effect was attributable to bulk inclusions, a microscopic study of inclusions was initiated, with special emphasis on the seed crystal interface.

Obvservations were made starting at the centerline of the seed (between the Z faces) and working outward in the direction of growth. Figure 24 summarizes the results.

SPECIES	LOCATION	SIZE (UM)	FREQUENCY
BUBBLE	SEED		
PLATE	BOUNDARY	15-30 X 1 (Z)	HIGH
BOX	BOUNDARY	3 X 5 X 8	HIGH
DIRT	BOUNDARY-BULK	<3	HIGH
PIN CUSHION	BOUNDARY-BULK	250-750	LOW
CLUMP	BULK	20-250	$<10\ cm^{-2}$
PATCH	BULK	20-250	LOW
LOG JAM	BULK	50-150	LOW

FIGURE 24

Figures 25 through 36 illustrate the various inclusion species.

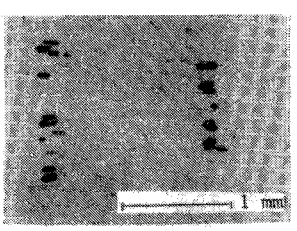


FIGURE 25 BUBBLE

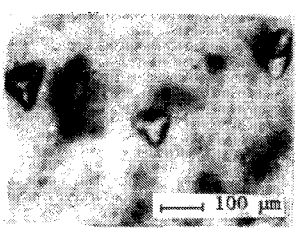


FIGURE 26 PLATE

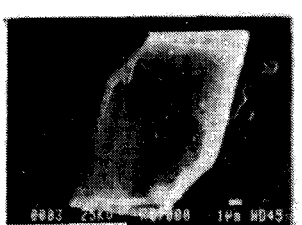


FIGURE 27 PLATE SEM

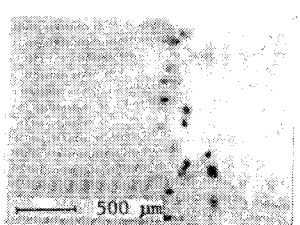


FIGURE 28 BOX

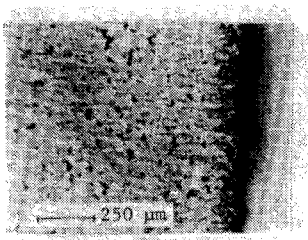


FIGURE 29 DIRT

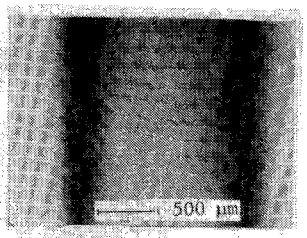


FIGURE 30 SEED VEIL

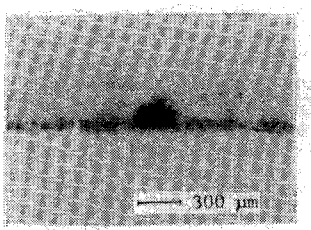


FIGURE 31 PIN CUSHION

FIGURE 32 P.C. SEM

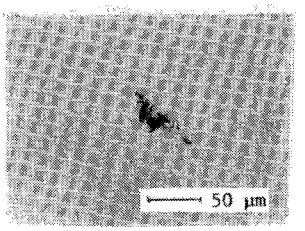


FIGURE 33 CLUMP

FIGURE 34 CLUMP SEM

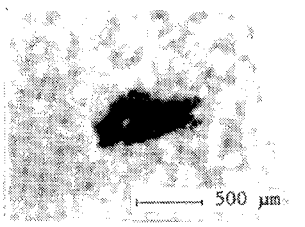


FIGURE 35 PATCH

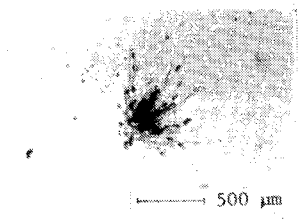


FIGURE 36 LOG JAM

The scale which forms on the interior surface of the steel autoclave wall was examined. The results of an SEM photograph are shown in Figure 37.

FIGURE 37

This has long been suspected as a source of inclusions in cultured quartz. However, as a result of this investigation, it is concluded that based on dissimilar morphologies, wall scale is not an important source of inclusions in cultured quartz.

Inclusion Structure and Composition

The most common inclusion species are the plate, box and clump. For the purposes of composition analysis, the plate and box are referred to as boundary inclusions and the clump as bulk inclusions.

A sample of a normal crystal was pulverized and a density separation performed to segragate the more dense inclusion particles. Using optical microscopy, it was determined the grains were biaxial, rectangular with large length to width ratio, exhibited subparallel extinction and showed cleavage parallel to the long axis. These observations are consistent with the common silicate mineral structures amphiobole and pyroxene.

SEM-EDAX analysis consistently revealed stoichiometry consistent with acmite ($NaFe(SiO_3)_2$) for bulk inclusions. In the case of boundary inclusions, the elements or ionic species varied. For example Mg^{+2}, Al^{+3} and Fe^{+2} (as opposed to the Fe^{+3} of the bulk inclusions) were detected. However, the basic model of acmite, or a similar pyroxene, fit the data. Importantly, the ratio of Si to the other elements is consistent with pyroxene compositions (Brantley, 1989). It is proposed that inclusions in cultured quartz are acmite or similar pyroxenes formed by a reaction of elements present in the growth solution.

Recent Efforts to Reduce Dislocations

Under funding of the Department of Defense and the National Science Foundation, various experiments were undertaken to reduce the dislocation density. The process was modified through two series of changes. The first involved initial efforts to decrease dissolved and particulate impurities in the growth solution. The second stage added further particulate filtration.

The results were compared with an initial data base of 66 runs of a standard crystal type grown to an average Z-dimension of 26 mm on a Z-plate seed with X of 37 mm at a growth rate of about 0.35 mm/day under conditions appropriate for carbonate growth The seed population was from normal production. The results are shown in Figure 38.

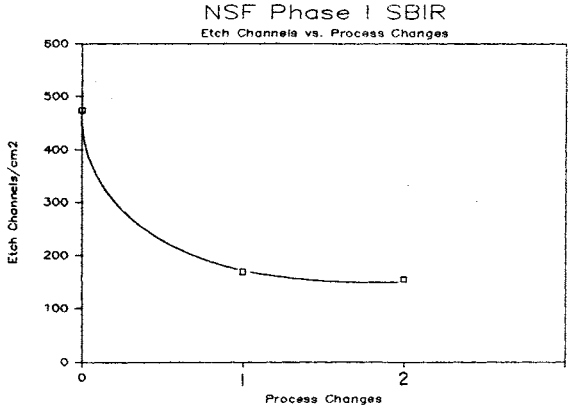


FIGURE 38

Selected seeds can yield dramatically lower etch channel densities (often fewer than 100 cm^{-2} and sometimes under 50 cm^{-2}), though on average, the results are not much different than shown in the above figure. In those crystals with abnormally low dislocation densities, the seed veil is diminished. This leads to the conclusion that for now the boundary inclusions are the dominating source of dislocations.

Figure 39 shows the result of using a greater-X seed coupled with few boundary inclusions. Note also the nucleation of dislocation lines from a bulk inclusion. With other sources suppressed, this phenomenon is observed.

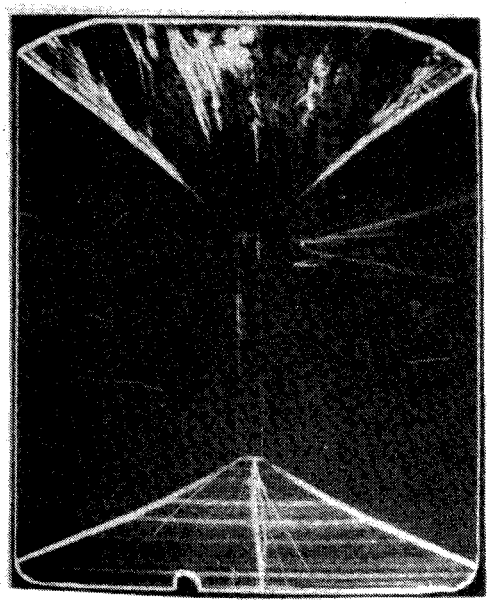


FIGURE 39

To lower the constituents of acmite in the system, a pressure balanced silver liner was used to eliminate interaction of the solution with the steel autoclave and its acmite passivation layer. Figure 40 shows this liner.

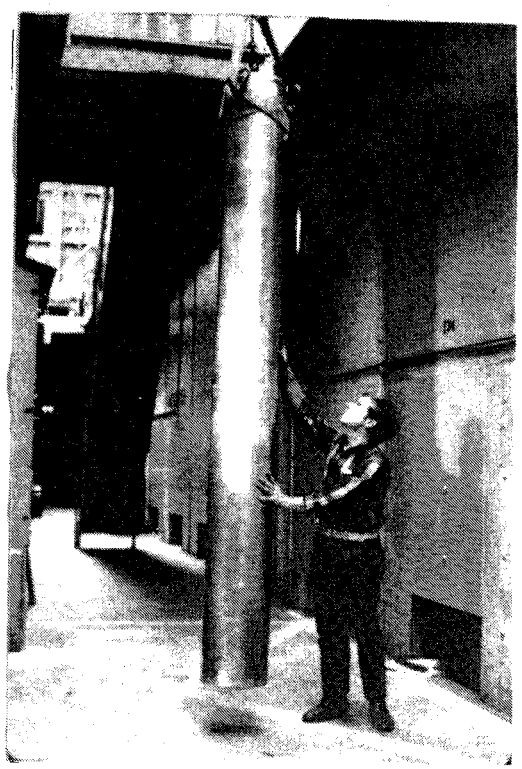


FIGURE 40

A photomicrograph of the seed boundary from a crystal produced in the silver liner is shown in Figure 41.

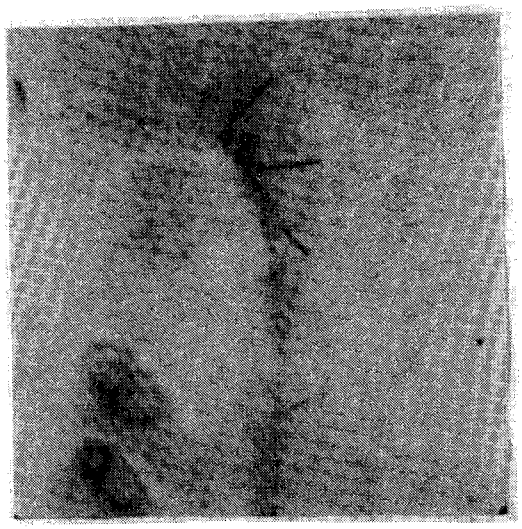


FIGURE 41

The inclusions on the seed are of different morphology than in unlined autoclave batches. Also, they were indistinguishable from the surrounding quartz during SEM analysis. Thus, they may be quartz inclusions which spontaneously nucleated on the seed surface prior to the onset of controlled growth.

A typical inclusion observed in the bulk of the crystal is seen in Figure 42. Similar to crystals grown in unlined autoclaves, the bulk inclusions are clumps of smaller inclusion species, though the morpohlogy differed considerably from crystals grown in unlined vessels.

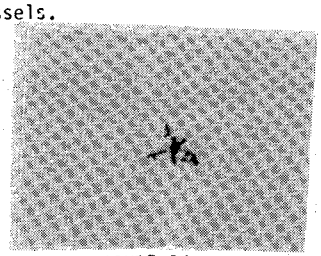


FIGURE 42

Figure 43 is an x-ray topograph of a sample from the silver run. The most interesting observations are:

1. The dislocation density of the seed is low.
2. The dislocation density of the crystal is not low.
3. No band of misfit dislocations is seen.

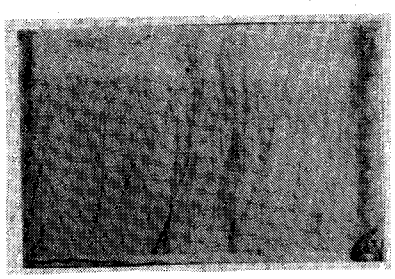


FIGURE 43

Discussion

Completed research indicates that substantial improvements are possible in reducing dislocations in cultured quartz. Although hydrogen content, growth rate, growth conditions, presence of point defects, etc. must contribute to the formation and propagation of edge dislocations resulting in etch pits and channels during etching, present evidence suggests that these are not the most critical issues. The two most important factors in producing commercial quantities of low dislocation material are the availability of low dislocation seeds and an improved understanding leading to eventual control of the seed crystal interface, especially the effect of inclusions in that area.

These inclusions seem to be acmite or similar pyroxenes formed in the growth solution. They are often, but not always, associated with a band of misfit locations. From this study, inclusions in the seed crystal interface appear to be the leading cause of bulk dislocations.

In the event this is correct, it becomes necessary to bring the formation of these inclusions under improved control. Past efforts have focused on overall process purity, including growth in noble metal lined vessels. Our research indicates, however, that even in these high purity systems, inclusion species can form. The complex chemistry of the system, coupled with the wide range of temperature and pressures through which the system passes, indicates that a large variety of compounds could nucleate and remain stable. Continuing efforts are focused on reduction of inclusions, especially at the seed crystal interface.

CONCLUSIONS

1. Dislocations in the Z-region of cultured quartz associated with pits and channels formed during etching are of the edge type directed approximately parallel to the direction of growth with Burgers vector $[\bar{1}2\bar{1}0]$.

2. The dominant sources of dislocations in cultured quartz are inclusions at the seed boundary and latent seed defects. Other observations about the seed/crystal interface are:

   A. Bubbles form during the onset of growth at surface terminations of seed dislocations and typically nucleate a single dislocation line.

   B. A combination of plates, dirt and boxes forms a debris plane at the onset of growth. Multiple dislocations nucleate from these particles.

   C. A band of misfit dislocations may coexist with the debris plane. The presence of misfit dislocations does not contribute to the formation of edge dislocations in the grown crystal.

3. Acmite and related pyroxenes are the most important inclusion forms and compositions in cultured quartz. The dominant source of inclusions is reaction of constituents of the system.

Acknowledgements

This work was partially supported by:

National Science Foundation through Small Business Innovation Research Program, Phase I, (Grant Number ECS-8460755) and Phase II, (Grant Number ISI-8521276).

Joint funding of the Department of Defense and Lawrence Livermore National Laboratory, administered by LABCOM, contract EAAL01-86-C-0012.

The authors also acknowledge the efforts of Mr. Kelley E. Scott of Sawyer Crystal Systems in Conroe, Texas in preparing samples and for his many observations and discussions leading to a better understanding of the seed veil.

We acknowledge further:

Dr. Sigmund Weissmann of Rutgers University for preparation and analysis of the x-ray topographs shown in Figures 13-19.

Dr. Evelyn Willibald-Riha of Siemens, Gmbh, Research Center in Munich, West Germany for preparing the x-ray topograph in Figures 6 and 39.

Honeywell, Sperry Flight Systems Division, for permission to use Figure 12.

Prof. S.L. Brantley of The Pennsylvania State University for examination and discussion of the structure and stoichiometry of inclusions.

The authors wish to thank the Program Committee of the European Frequency and Time Forum especially Dr. J.J. Gagnepain of CNRS-LPMO, Besancon, France for extending an invitation to present this paper.

Mr. K. Hamaguchi of Toshiba Ceramics, Inc., Tokyo, Japan, for discussions regarding shadowgraph methods.

References

Arnold, G.W., Proceedings of the 11th Frequency Control Symposium, 1957, p. 112.

Augustine, F., Hale, D.R., and Berry, J., 5th Progress Report, U.S. Army Contract DA33-039-SC-72415, p. 15.

Armington, A.F., Larkin, J.J., O'Connor, J.J. and Horrigan, J.A., Proceedings of the 35th Frequency Control Symposium, 1981, p. 297.

Armington, A.F., Larkin, J.J., O'Connor, J.J. and Horrigan, J.A., Proceedings of the 36th Frequency Control Symposium, 1982, p. 55.

Armington, A.F., and Balascio, J.F., Proceedings of the 38th Frequency Control Symposium, 1984, p. 3.

Armington, A.F., and Balascio, J.R., Proceedings of the 39th Frequency Control Symposium, 1985, p. 230.

Balascio, J.F., and Armington, A.F., Proceedings of the 40th Frequency Control Symposium, 1986, p. 70.

Barns, R.L., Freeland, P.E., Kolb, E.D., Laudise, R.A., and Patel, J.R., Journal Crystal Growth, 43, 1978, p. 676.

Bernot, A.J., Proceedings of the 39th Frequency Control Symposium, 1985, p. 271.

Brantley, S.L., The Pennsylvania State University, personal correspondence, February 1989.

Buerger, M.J., Journal American Minerals, 39, 1954, p. 600.

Cabrera, N., Levine, M.M. and Plaskett, J.S., Phys. Rev. 96, 1954, p. 1133.

Cohen, A.J., Journal Phys. Chem. Solids, 13, 1960, p. 321.

Croxall, D.F., Cristie, I.R.A., Holt, J.M., Isherwood, B.J., and Todd, A.G., Proceedings of the 36th Frequency Control Symposium, 1982, p. 62.

Gaultieri, J.G., Proceedings of the 39th Frequency Control Symposium, 1985, p. 247.

Hanyu, T., Journal Applied Phys. Soc. Japan, 19, 1964, p. 1489.

Homma, S. and Iwata, M., Journal Crystal Growth, 19, 1973, p. 125.

Iwasaki, F., Journal Crystal Growth, 39, 1977, p. 291.

Johnson, G.R., Irvine, R.A., 41st Frequency Control Symposium, 1987, p. 175.

Lang, A.R. and Miuscov, V.F., Journal Applied Physics, 38, 1967, p. 2477.

Martin, J.J., Bossoli, R.B., Halliburton, L.E., Subramanian, B., and West, J.D., Proceedings of the 37th Frequency Control Symposium, 1983, p. 164.

McLaren, A.C., Osborne, C.F. and Saunders, L.A., Phys. Stat. Sol. (a) 4, 1971, p. 235.

Nielsen, J.W. and Foster, F.G., Am. Mineral., 45, 1960, p. 299.

Saha, P., Annamalai, N., Baudyopadhyay, T., Mater Sc. Bull (India), 1, 1979, p. 15.

Spencer, W.J. and Haruta, J., Applied Physics, 37, 1966, p. 549.

Thicksten, H., Proceedings of the 5th Quartz Crystal Conference, 1983, p. 234.

Tsinzerling, E.V. and Mieonua, Z.A., Kristallog 8, 1963, p. 117.

Vig, J.R., LeBus, J.W. and Filler, R.L., Proceedings of the 31st Frequency Control Symposium, 1977, p. 131.

Weissmann, S., Unpublished Proprietary Report to Sawyer Research Products, 1985.

Zarka, A., Lin, L. and Buisson, M., Journal Crystal Growth, 54, 1981, p. 394.

Zheng, et al, to be presented at the 43rd Frequency Control Symposium, 1989.

TENSORIAL CHARACTERIZATION OF THE GEOMETRICAL FEATURES FOR ETCHED QUARTZ PLATES

C.R. TELLIER and T.G. LEBLOIS

Laboratoire de Chronométrie, Electronique et Piézoélectricité
ECOLE NATIONALE SUPERIEURE DE MECANIQUE ET DES MICROTECHNIQUES
Route de Gray - La Bouloie - 25030 BESANCON CEDEX - FRANCE -

ABSTRACT

The anisotropic etching of quartz crystals is explained in terms of an orientation-dependent dissolution slowness vector $\vec{L}(\phi, \theta)$. The magnitude, $L(\phi, \theta)$ of the slowness vector is founded to be expressed in terms of dissolution tensors. The number of dissolution constants is reduced by considering the degree of symmetry of quartz crystals. The calculations performed up to the tensor of rank 6 allow us to propose a generalized equation for the representative surface of the dissolution slowness. Once this equation is established we can derive the dissolution shapes firstly, of an initially circular section and secondly of a surface profile. To illustrate the first possibility the successive dissolution shapes of a starting circular section are presented in the special case of an asymmetric polar path. Then for singly rotated quartz plates the unknown coefficients appearing in the equation, $L_{YZ}(\theta)$, coefficients of the polar path are determined from experiments on the etch rate. We observe a very satisfactory agreement between the theoretical shapes of dissolution profiles as given by a numerical simulation and the experimental Z' traces obtained after repeated etchings.

1. INTRODUCTION

The mechanical lapping of quartz crystals has the disavantage to create a disturbed surface layer (1-3) then some authors (1, 4-9) have suggested that the chemical etching can constitute an alternative procedure to prepare quartz plates with clean and smooth surfaces. Some experiments on AT-cut plates (4, 8, 10) have shown that it is effectively possible to obtain quartz plates with satisfactory surface features. However a lot of recent experiments (10-18) on the chemical etching of differently oriented quartz plates have given evidence for an anisotropic etching governed by the crystal orientation. In particular small variations in the angles of cut can induce marked changes in the shape of the dissolution figures which develop on etched quartz surfaces as observed for example for the cuts close to the AT-cut (12).

For an orientation-dependent dissolution process it is convenient to interpret the changes in the dissolution shapes with orientation in terms of the kinematic theory of etching (19, 20, 21) originally given by Lighthill and Whitman (22) and further extended by Frank (23). Thus Tellier and al (24) have recently proposed a vectorial analysis of the dissolution based on the Frank model where in two dimensions we can associate to a surface profile element of given orientation θ a dissolution slowness vector $\vec{L}(\theta)$. The propagation vector $\vec{P}$ which is directly connected with the movement of the profile element can be evaluated from the polar path $L(\theta)$. This paper continues along this line and presents a three dimensional model in which the magnitude of the dissolution slowness vector $\vec{L}$ is determined from a tensorial analysis of the dissolution.

2. THEORETICALS

2.1/ The kinematic model in two dimensions

In two dimensions a surface profile is supposed to be composed by a succession of linear surface profile elements $\Delta r(\alpha)$ (Figure 1) and a moving surface profile is a function of two independent variables the coordinate x_2 and the etching time t.

It is usual to deal with a velocity, $v_3 = (\partial x_3 / \partial t)_{x_2}$, measured in the direction of the normal, $\vec{n}_o$, to the reference surface. If the velocity v_3 depends only on the orientation α of the surface profile element, an assumption which for quartz crystals is sustained by a great number of experimental data (10-18), the moving profile element propagates into the bulk crystal along a linear trajectory called a "characteristic". Then after an etching of duration dt the point M is located at M' (Figure 2). At this point it is convenient to define a propagation vector P with cartesian components (dx_2, dx_3) which corresponds to $\vec{MM'}$. Then when we track the movement of the surface profile element we have just to evaluate $\vec{P}$. This may be done by considering a new vector, $\vec{L}$, called the dissolution slowness vector (24), whose direction coïncides with that of the inward normal n to the profile element, Δr, (Figure 2) and whose magnitude, L, is equal to the reciprocal of rate V_N (i.e. the normal velocity) (24). When the orientation α of the profile element varies the extremity of the vector, $\vec{L}$, moves along a polar path (Figure 3).

After some geometrical and mathematical manipulations (24) we can show that the vector $\frac{\partial \vec{L}}{\partial \beta}$ which is the derivative of the rotating vector $\vec{L}$ with respect to the polar angle β lies perpendicular to the propagation vector $\vec{P}$ (Figure 3). Since according to the polar representation of $\vec{L}$ the cartesian components of the slowness vector are expressed as

$$L_2 = f(\beta) \cos \beta \qquad (1)$$

$$L_3 = f(\beta) \sin \beta$$

with

$$\vec{L} \cdot \vec{L} = \left(f(\beta) \right)^2 \qquad (2)$$

we obtain after some calculations

$$dx_2 = \pm \frac{B(\beta)}{\left\{f(\beta)\right\}^2} dt \qquad (3)$$

$$dx_3 = \mp \frac{A(\beta)}{\left\{f(\beta)\right\}^2} dt \qquad (4)$$

where $A(\beta)$ and $B(\beta)$, the cartesian components of the tangent vector $\vec{\tau}$, are expressed in terms involving $f(\beta)$ and $\partial f(\beta)/\partial \beta$. Thus we see that we need to know the equation for the polar path i.e. the equation for L to derive the propagation vector $\vec{P}$.

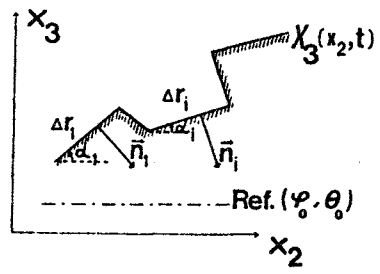


Figure 1 : Representation of a surface profile.

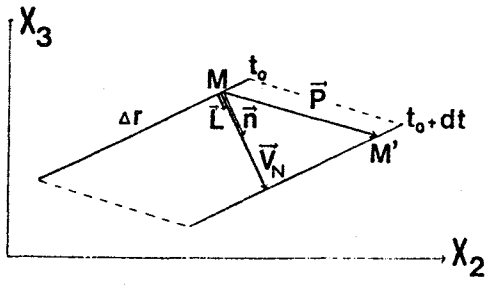


Figure 2 : Definition of the vectors $\vec{P}$ and $\vec{L}$.

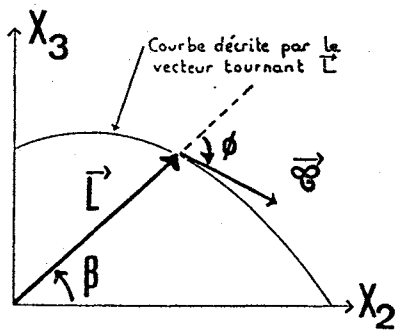


Figure 3 : The polar graph of $\vec{L}$.

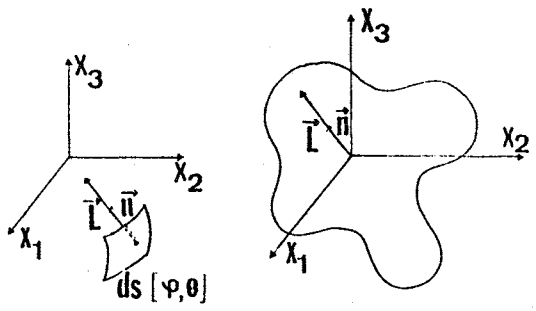


Figure 4 : The representative surface of $\vec{L}$.

2.2/ The representative surface of the dissolution slowness

In space (x_1, x_2, x_3) the dissolution slowness vector $\vec{L}$ is associated with a surface element ds whose orientation is defined either by the angles of cut (ϕ, θ) or by means of its inward normal $\vec{n}$ with cartesian components n_1, n_2, n_3 (Figure 4). As the orientation of the surface element varies both the magnitude and the direction of $\vec{L}$ are modified and it is easy to see that in three dimensions the dissolution slowness vector $\vec{L}$ moves along a representative surface called for convenience the slowness surface (25).

The problem is now to find the equation for this slowness surface, i.e. the equation expressing the magnitude, L, of the slowness vector in terms of the orientation (ϕ, θ) of a surface element. A rigorous solution seems at the first sight untractable, but some simplifications can be made. Firstly we can remark that the components n_i of the inward normal $\vec{n}$ are functions of the angles of cut (ϕ, θ) so we have finally to derive an equation in the form

$$L = L(n_1, n_2, n_3)$$

Secondly it is always possible to replace the exact function $L(n_1, n_2, n_3)$ by a polynomial representation in which three variables n_1, n_2 and n_3 are involved. In a compact form we can thus write

$$L = D_o + D_i n_i + D_{ij} n_i n_j + D_{ijk} n_i n_j n_k + D_{ijk\ell} n_i n_j n_k n_\ell + D_{ijk\ell n} n_i n_j n_k n_\ell n_n + \ldots \qquad (5)$$

where the coefficients D_o, D_i, D_{ij}, D_{ijk}, ... are components of tensors of rank 0, 1, 2, 3, ... which are designated as dissolution tensors.

It should be pointed out that from the definition of the dissolution tensors all the subscripts can be interchanged that is to say, for example

$$D_{ijk\ell} = D_{jik\ell} = D_{ij\ell k} = D_{i\ell jk} \qquad (6)$$

To determine the independent dissolution constants we consider the degree of symmetry of the quartz crystal. Thus we are concerned with two transformations from axes (x_1, x_2, x_3) to primed axes (x'_1, x'_2, x'_3) :

① The transformation involving the twofold axis x_1 (crystallographic X axis) in which the primed components of the inward normal are given by :

$$n'_1 = n_1 \; ; \; n'_2 = -n_2 \; ; \; n'_3 = -n_3 \qquad (7)$$

② The transformation following the threefold symmetry of the x_3 axis (crystallographic Z axis) which is specified by the transformation matrix

	1	2	3
1'	C	S	0
2'	- S	C	0
3'	0	0	1

(8)

with C = cos (120°) and S = sin (120°). Here we apply the usual transformation rule which for a 6th-order tensor is written as

$$D'_{ijk\ell mn} = a_{io}a_{jp}a_{kq}a_{\ell r}a_{ms}a_{nt}D_{opqrst} \qquad (9)$$

The calculations which are performed up to the 6th-order are rather laborious and give several relations between the independent constants which are listed in table 1.

Rank N	Relations between the dissolution constants and independent constants
0	D_o
1	All the constants are equal to zero
2	$D_{11} = D_{22}$; D_{33}
3	$D_{111} = -D_{122}$
4	$D_{1111} = D_{2222} = 3D_{1122}$; $D_{1133} = D_{2233}$; $D_{1123} = -D_{2223}$; D_{3333}
5	$D_{11111} = -5D_{11122} = -(5/3)D_{12222}$; $D_{12233} = -D_{11133}$
6	$D_{111111} = D_{222222} = 5D_{112222} = D_{111122}$
	$D_{111123} = 3D_{112223} = -(3/5)D_{222223}$
	$D_{111133} = D_{222233} = 3D_{112233}$; $D_{112333} = -D_{222333}$
	$D_{113333} = D_{223333}$; D_{333333}

Table 1 : Identification of the independent dissolution constants of a dissolution tensor of rank N.

Then we finally obtain $L(n_1, n_2, n_3)$ in the form (25).

$$L = D_o + D_{11}(n_1^2 + n_2^2) + D_{111}\,n_1(n_1^2 - 3n_2^2) + D_{1111}(n_1^2 + n_2^2)^2 + D_{3333}n_3^4 + 6D_{1133}n_3^2(n_1^2 + n_2^2) + 4D_{2223}n_2n_3(3n_1^2 - n_2^2) + D_{11111}n_1(n_1^2 + n_2^2)(n_1^2 - 3n_2^2) + 10\,D_{11133}n_1n_3^2(3n_2^2 - n_1^2) + D_{111111}(n_1^2 + n_2^2)^3 - 10\,D_{111123}\,n_2n_3(n_2^2 + n_1^2)(n_2^2 - 3n_1^2) + 15n_3^2(n_1^2 + n_2^2)^2\,D_{111133} + 20n_2n_3^3(3n_1^2 - n_2^2)\,D_{112333} + 15n_3^4(n_1^2 + n_2^2)D_{113333} + n_3^6D_{333333} \qquad (10)$$

Since the cartesian components of $\vec{n}$ are expressed as

$$n_1 = \sin\phi\cos\theta\ ,\ n_2 = -\cos\phi\cos\theta\ ,\ n_3 = -\sin\theta \qquad (11)$$

some manipulations of Equation (10) allow us to propose a generalized equation for the representative slowness surface :

$$L(\phi, \theta) = L_o + \sum_j A_j \cos^{2j}\theta + \sum_k B_k \cos^{2k+1}\theta \sin 3\phi + \sum_\ell C_\ell \cos^{2\ell+1}\theta \sin\theta \cos 3\phi \qquad (12)$$

which involves only the angles of cut ϕ and θ and the unknown coefficients L_o, A_j, B_k and C_ℓ.

3. DETERMINATION OF DISSOLUTION SHAPES

3.1/ Dissolution shapes of initially circular sections

Here some attempts are made to determine theoretically the dissolution shapes of an initially spherical crystal. Then to a point M located on the sphere we associate a planar surface element whose orientation (Φ , θ) coïncide with that of a plane tangent to the sphere at the point M so that (Figure 5) the slowness vector $\vec{L}$ (Φ , θ) related to this surface element lies parallel to MO. The successive dissolution shapes of the crystal sphere can be derived from the three dimensional analysis of the dissolution as soon as for a crystal of given symmetry we are able to determine the unknown coefficients appearing in the corresponding generalized equation of the slowness surface.

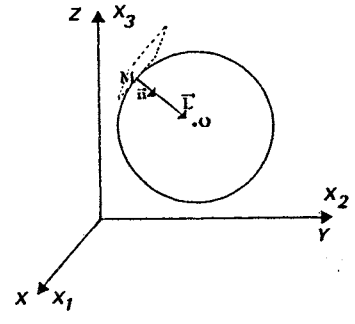


Figure 5 : The geometry for an initially circular section.

After prolonged etching spherical crystals are usually reduced to a limiting form which results from diverging and converging trajectories (21, 26, 27). Effectively when several dissolution trajectories corresponding to orientations lying in the range $[(\phi_1, \theta_1), (\phi_2, \theta_2)]$ intersect a crystal edge is produced (Figure 6). More large is the orientation range more sharp becomes the crystal edge. Conversely diverging trajectories lead to the formation of rounded surfaces.

To illustrate briefly this situation we start here with a cylindrical crystal specimen and we track the changes in shape of the initially circular cross section. Then the knowledge of a polar graph is only required for this application. Furthermore another paper is devoted to dissolution shapes of differently oriented circular sections of quartz crystal thus for an example we consider here the asymmetrical polar path of figure 7. The dissolution shapes are derived from a specific numerical and graphical simulation which allows us to deal with elements whose orientation varies from 0° to 360° and to undertake suitable tests to distinguish diverging and converging trajectories. Figure 8 shows the successive limiting forms of the cross section of an initially cylindrical crystal whereas the complementary dissolution shapes of figure 9 result from repeated etchings of an initially cylindrical hollow.

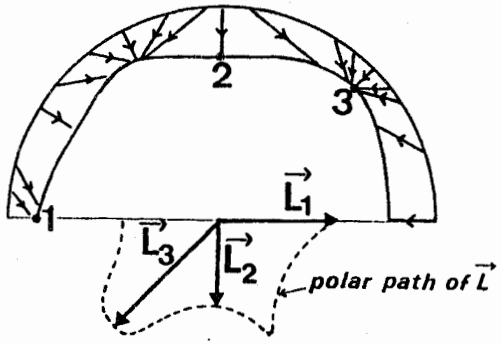


Figure 6 : Formation of a crystal edge.

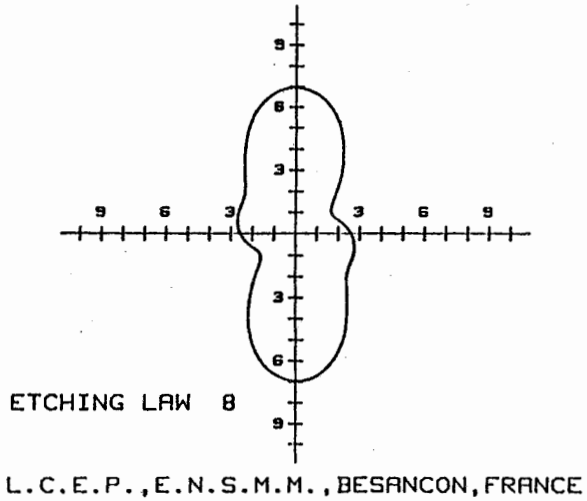


Figure 7 : An asymmetrical path for the dissolution slowness vector.

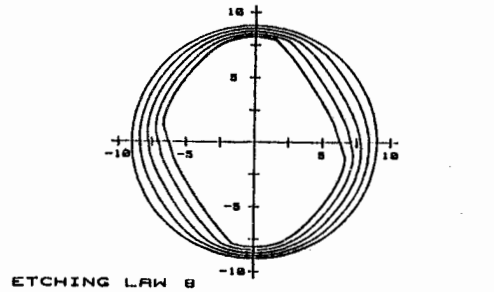


Figure 8 : Dissolution of a cylindrical crystal.

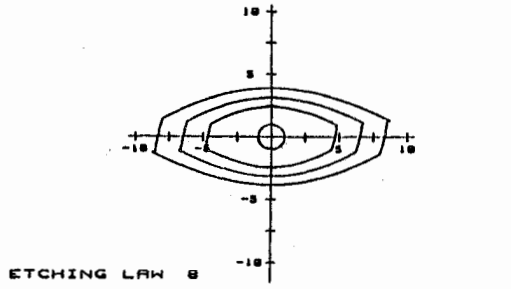


Figure 9 : Dissolution of a cylindrical hollow.

As the shape of a section element changes from convex to concave when we start respectively with a cylindrical crystal and a cylindrical hollow we observe that according to the reciprocity criterion previously proposed by Frank (28, 29) the elements at a relative maximum of dissolution slowness persist in the dissolution shapes of figure 8 while the converse is true for the dissolution shapes of figure 9.

3.2/ The Z' dissolution profiles of singly rotated quartz plates

A second application of the tensorial analysis consists of predicting the theoretical shapes of etched surface profiles. Here again we are concerned with a two-dimensional analysis and we decide for example to consider the case of the Z' profiles of singly rotated quartz plates. Since for differently oriented quartz plates the slowness vectors associated with elements of the Z' profiles belong all to the (x_2, x_3) plane (i.e. to the (Y, Z) plane) the treatment is reduced to the determination of the polar path, $L_{YZ}(\theta)$, which is expressed by

$$L_{YZ}(\theta) = L_o + \sum_j A_j \cos^{2j}\theta + \sum_\ell C_\ell \cos^{2\ell+1}\theta \sin\theta \quad (13)$$

Practically the unknown constants A_j and C_ℓ are numerically evaluated using experimental measurements on the dissolution rate of various singly rotated quartz plates (26). The evaluation is accomplished up to the 16th order by collecting data on at 30 different cuts. The result for θ in the range (- 60°, 60°) is illustrated in Figure 10 where for convenience we have plotted $L_{ZY}(\theta)$ vs the angle of cut, θ .

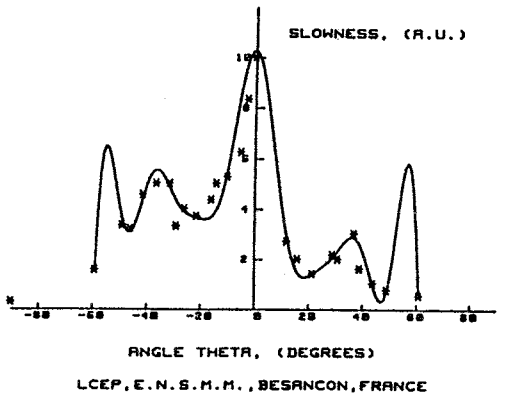


Figure 10 : $L_{ZY}(\theta)$ vs the angle of cut, θ .

With the help of a graphical simulation whose program has been presented elsewhere (19, 20, 24) we can investigate how the orientation of the reference surface (i.e. of the cut) influences the final shape of the etched Z' profile. For example, let us consider the AT-33, AT-39, AT-44 and AT-49 plates corresponding to a reference surface of orientation θ_o of 33°, 39°, 44° and 49° respectively. The theoretical evolution of the etched Z' profiles are for the various cuts illustrated in figures 11 to 14 starting from a triangular surface profile with slopes equal to tan (± 12°) whereas figures 15 and 16 give results when an initial randomized surface profile is etched. We can observe that the final shape of the dissolution profiles does not depend on the initial profile. Moreover these theoretical final shapes merit some comments.

① When the local dissolution slowness is faster than the slowness of the reference surface a quasi convex section is formed (figures 14 and 17). Conversely when the local dissolution slowness is slower than that related to the cut a concave background results as approximately observed for the AT-39 cut (figures 12, 16 and 18).

② When the dissolution slowness of the reference surface lies midway from a relative maximum and a relative minimum of the slowness (figures 11, 13, 15 and 19) the etching causes the development of a surface profile with an alternate convex-concave or concave-convex shape. A situation which occurs in the case of AT-33 and AT-44 quartz plates.

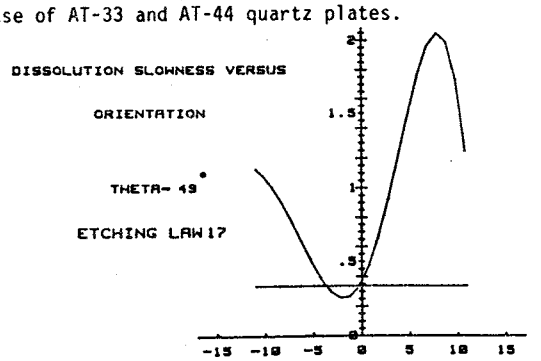


Figure 17 : The dissolution slowness vs orientation for θ_o = 49°.

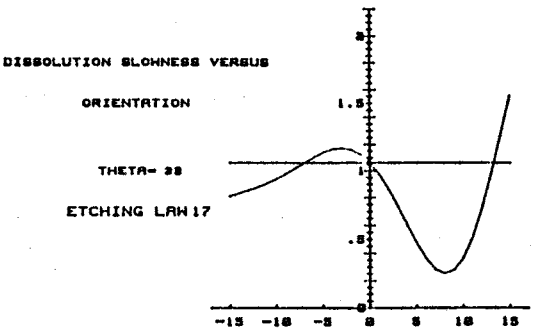


Figure 18 : The dissolution slowness vs orientation for a reference surface of orientation θ_o = 39°.

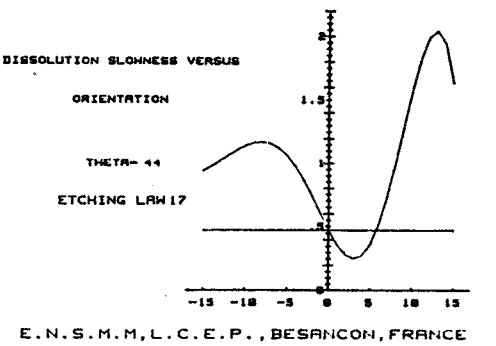


Figure 19 : The dissolution slowness vs orientation for a reference surface of orientation θ_o = 44°.

To give some useful informations on the consistency of the method we have studied the changes in shape of Z' profiles with the depth of etch, Δd_s, of the various AT-θ_o plates. As usual the plates were etched at constant temperatures in a concentrated $NH_4F.HF$ solution. After each isothermal etching profilometry traces were made along the Z' direction using a microprocessor based profilometer (30) (figure 20). To assess the validity of the model magnifications of the final Z' profiles are displayed in figure 21. In addition the final S.E.M. micrographs of the etched surfaces are also shown in figures 22 to 25.

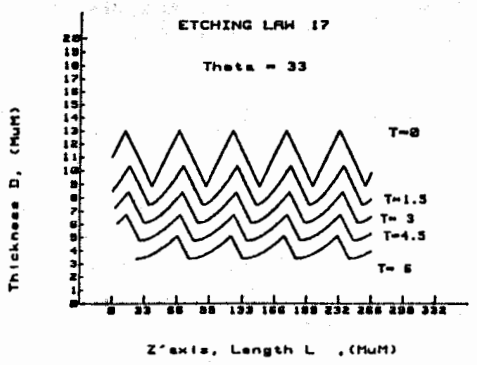


Figure 11 : Development of Z' surface profiles starting from a triangular profile : The case of an AT-33 plate (θ_0 = 33°).

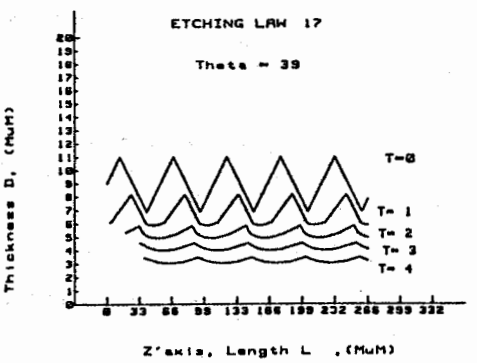


Figure 12 : Development of Z' surface profile for an AT-39 plate (θ_0 = 39°).

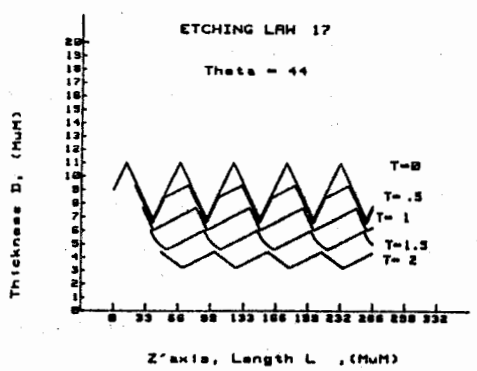


Figure 13 : Development of Z' surface profiles for an AT-44 plate (θ_0 = 44°).

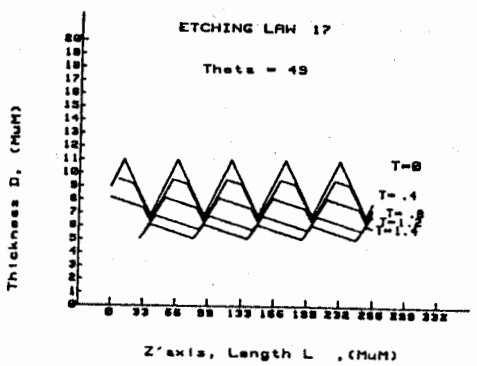


Figure 14 : Development of Z' surface profiles for an AT-49 plate (θ_0 = 49°).

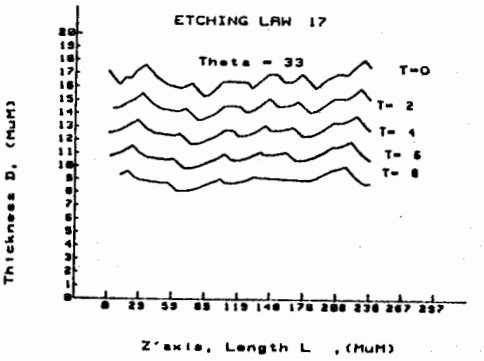


Figure 15 : Development of Z' surface profiles starting from a randomized profile : The case of an AT-33 plate.

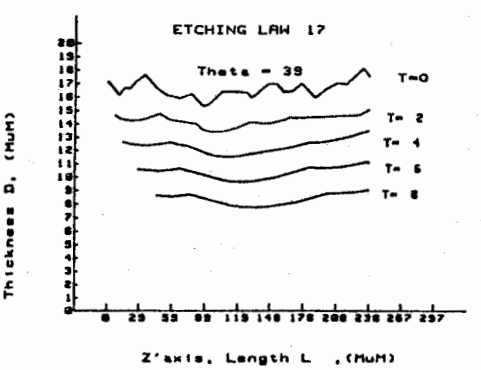


Figure 16 : Development of Z' surface profiles for an AT-39 plate.

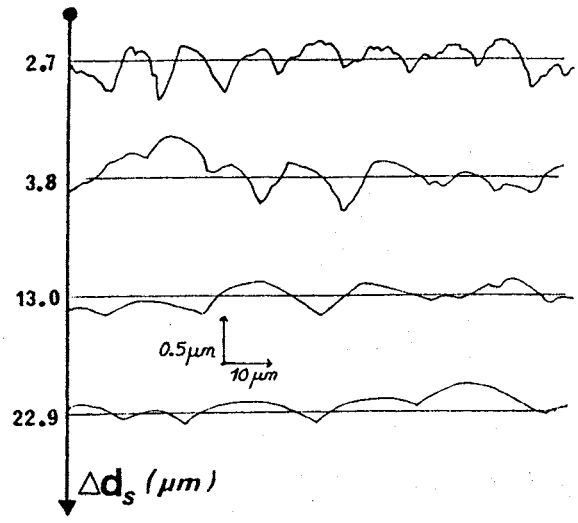


Figure 20 : **Changes in the Z' profilometry traces with the depth of etch of an AT-49 plate.**

We observe a very satisfactory agreement between the theoretical and the experimental shapes of the etched Z' profiles since

① quasi-concave and convex background structures develop on the AT-39 and AT-49 plates respectively

② final Z' profiles of alternate shape are formed on the AT-33 and AT-44 plates.

4. CONCLUSION

We have shown that at once the equation $L(\phi,\theta)$ of the slowness surface corresponding to a crystal of given symmetry is derived and that the unknown coefficients appearing in this equation are evaluated from data we can track the changes in shape with etchings of either a cross section of a spherical crystal or a surface profile. The equation $L(\phi,\theta)$ related to the quartz crystal is obtained by a tensorial method and the calculations are performed up to dissolution tensors of rank 6 in order to propose a generalized $L(\phi,\theta)$ equation. The unknown coefficients are then partially evaluated from data on singly rotated quartz plates. From a comparison of the theoretical dissolution profiles as given by a numerical simulation with the experimental Z' profilometry traces of various singly rotated quartz plates we can conclude that the experimental determination of the unknown coefficients is rather accurate. This confirms the great interest of this tensorial approach of the dissolution.

REFERENCES

(1) J. VIG, H. WASSHAUSEN, C. COOK, M. KATZ and E. HAFNER, **Proc. 27th Ann. Symp. on Frequency Control, Fort Monmouth, N.J. 1973,** Electronic Industries Association, Washington, DC, (1973), p 98.

(2) Y. SEKIGUCHI and H. FUNAKUBO, J. Mater. Sci., 15 (1980) 3066.

(3) H. FUKUYO and N. OURA, **Proc. 30th Annu. Symp. on Frequency Control, Fort Monmouth, N.J., 1976,** Electronic Industries Association, Washington, DC,(1976), p 254.

(4) J.R. VIG, J.W. LEBUS and R. FILLER, **Report ECOM-4548,** 1977 (US Army Electronics Command, Fort Monmouth, N.J.).

(5) J.R. VIG, C.F. COOK, K. SCHWIDTAL, J.W. LEBUS and E. HAFNER, **Proc. 28th Ann. Symp. on Frequency Control, Fort Monmouth, N.J., 1974** ; Electronic Industries Association, Washington, DC, (1974), p 96.

(6) J.R. VIG, R.J. BRANDMAYR and R.L. FILLER, **Proc. 33rd Annu. Symp. on Frequency Control, Fort Monmouth, N.J., 1979** ; Electronic Industries Association, Washington, DC, (1979), p 351.

(7) R.J. BRANDMAYR, J.R. VIG, **Proc. 39th Annu. Symp. on Frequency Control, Philadelphia, Pa, 1985** ; I.E.E.E., New York, N.Y., (1985), p 276.

(8) D. ANG, **32nd Annu. Symp. on Frequency Control, Fort Monmouth, N.J., 1978** ; Electronic Industries Association, Washington, DC, (1978), p 282.

(9) A.J. BERNOT, **Proc. 39th Annu. Symp. on Frequency Control, Philadelphia, Pa, 1985** ; I.E.E.E., New York, N.Y., (1985), p 271.

(10) C.R. TELLIER, Surf. Technol., 21 (1984),p83.

(11) C.R. TELLIER and C. BURON, Surf. Technol., 22 (1984),p287.

(12) C.R. TELLIER, **Proc. 38th Annu. Symp. on Frequency Control, Philadelphia, Pa, 1984** ; I.E.E.E., New York, N.Y., (1984) p 105.

(13) C.R. TELLIER, **Proc. XIth International Congress of Chronometry, Besançon, France, 1984** ; Société Française des Microtechniques et de Chronométrie, Besançon (1984) p 115.

(14) C.R. TELLIER, **Proc. 39th Annu. Symp. on Frequency Control, Philadelphia, Pa, 1985** ; I.E.E.E., New York, N.Y., (1985), p 276.

(15) C.R. TELLIER and F. JOUFFROY , J. Mater. Sci., 18 (1983) 3621.

(16) M. CASTAGLIOLA, C.R. TELLIER and J.L. VATERKOWSKI, J. Mater. Sci., 21 (1986) 3551.

(17) C.R. TELLIER, F. JOUFFROY and C. BURON, Mater. Chem. & Phys., 14 (1986) 25.

(18) C.R. TELLIER and J.L. VATERKOWSKI, Surf. Technol., 26 (1985) 275.

(19) C.R. TELLIER, N. VIALLE and J.L. VATERKOWSKI, **Proc. 40th Annu. Symp. on Frequency Control, Philadelphia, Pa, 1986** ; I.E.E.E., New York, N.Y., (1986) p 76.

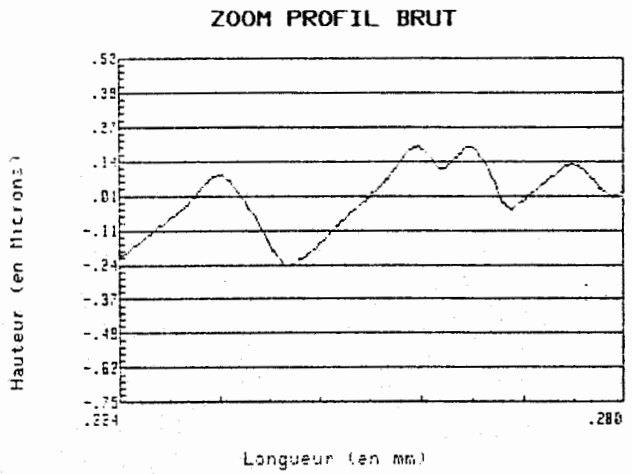


Figure 21a : The magnified Z' profilometry trace of an AT-33 plate.

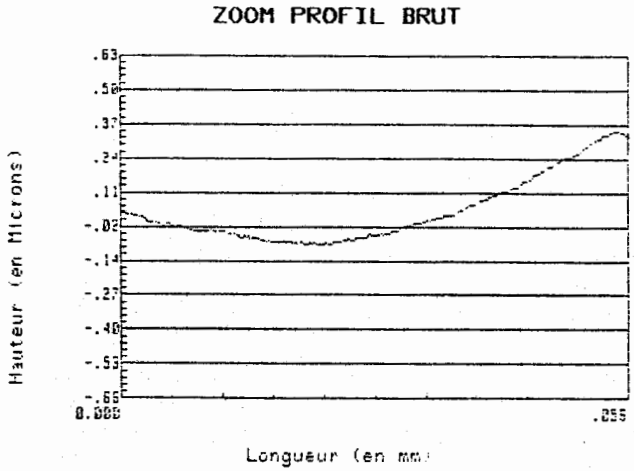


Figure 21b : The magnified Z' profilometry trace of an AT-39 plate.

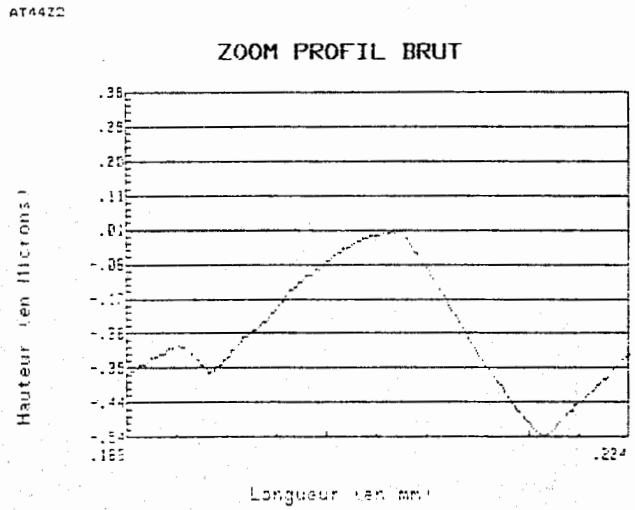


Figure 21c : The magnified Z' profilometry trace of an AT-44 plate.

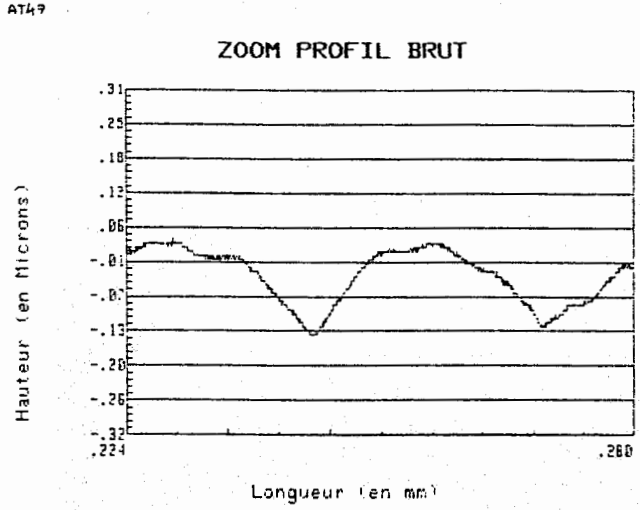


Figure 21d : The magnified Z' profilometry trace of an AT-49 plate.

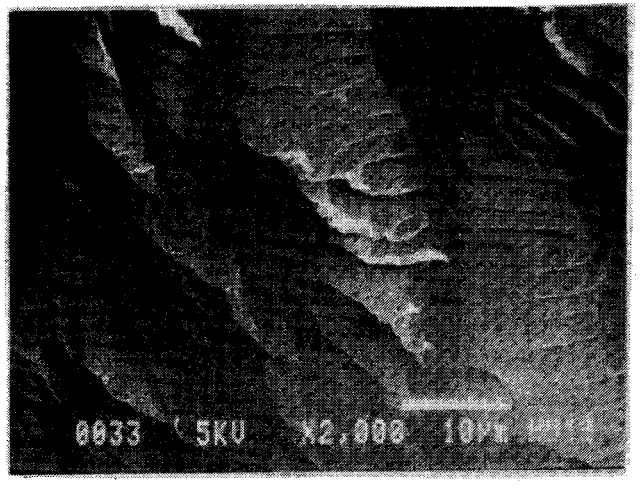


Figure 22 : SEM micrograph of a deeply etched AT-33 plate.

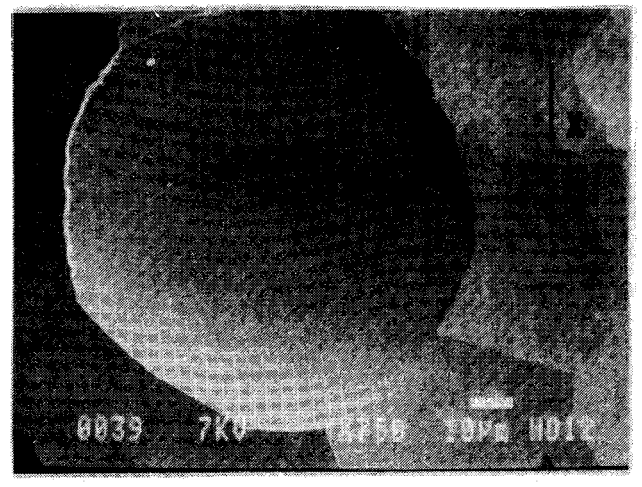


Figure 23 : SEM micrograph of a deeply etched AT-39 plate.

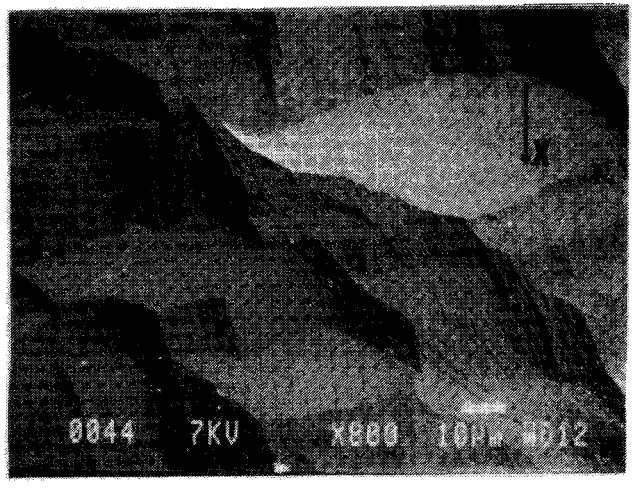


Figure 24 : SEM micrograph of a deeply etched AT-44 plate.

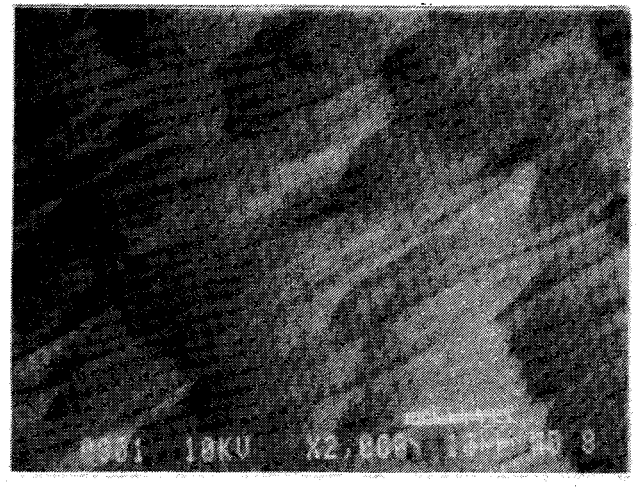


Figure 25 : SEM micrograph of a deeply etched AT-49 plate.

(20) C.R. TELLIER, N. VIALLE and J.L. VATERKOWSKI, **First European Time and Frequency Forum, Besançon, France, 1987** (Imprimerie du Conseil Général du Doubs, Besançon, 1987) p 159.

(21) R.B. HEIMANN, in J. GRAMAIER (Ed.), **Silicon Chemical Etching**, Springer, Berlin, (1982) pp 197-224.

(22) M.J. LIGHTHILL and G.B. WHITHAM, Proc. Roy. Soc., A229 (1955) 281.

(23) F.C. FRANK, in R.H. DOREMUS, B.W. ROBERTS and D. TURNBULL (eds.), **Growth and Perfection of Crystals**, John Wiley, New York (1958), pp 411-419.

(24) C.R. TELLIER, N. VIALLE and J.L. VATERKOWSKI, Surf. & Coatings Technol., 34 (1988) 417.

(25) C.R. TELLIER and J.L. VATERKOWSKI, J. Mater. Sci., (to be published).

(26) B.A. IRVING in P.J. HOLMES (ed.), **The Electrochemistry of Semiconductors** , Academic Press, London (1962) pp 256-289.

(27) K. SANGWAL, **Etching of Crystals** , North Holland, Amsterdam, (1987).

(28) F.C. FRANK and M.B. IVES, J. Appl. Phys., 31 (1960), 1996.

(29) B.W. BATTERMAN, J. Appl. Phys., 28 (1957) 1236.

(30) P. BLIND, D. JOZWICK and C.R. TELLIER, **2nd European Time and Frequency Forum, Neuchâtel, Switzerland, 1988** (Fondation Suisse pour la Recherche en Microtechniques, Neuchâtel, 1988) p 937.

* * * * * *

ACKNOWLEDGMENTS

The authors gratefully acknowledge financial support from the **Direction des Recherches Etudes et Techniques** which depends on the French Department of Defense.

* * * * * *

TRACE ELEMENT ANALYSIS IN QUARTZ

Bernard Viard, Jérôme Daumet, Mohamed Khayar, Henri Mérigoux

Laboratoire de Cristallographie et Chimie Minérale
U.F.R. des Sciences et des Techniques
25030 Besançon Cedex - France

Summary

Determination of the quality of a quartz used in a resonator may be obtained by I.R. spectrometry. In this paper we present a direct determination of the chemical impurities in a quartz crystal. By an ICP method we measure the concentration of Al, Fe, Ti, Co, Ca, Na, Li, K, B, Ba, Sc, Ge atoms with an accuracy of 0.1 ppm if the mass of the dissolved quartz in 10g.

In order to avoid unexpected contaminations we use a very severe method. All the chemicals and water we use are of very high purity. The quartz is not crushed but is split with a thermic shock. In order to reduce the amount of hydrofluoric acid the attack of the quartz is made using a reflux method. We add 10 ppm of Scandium in the solution as internal standard. This element is not known in quartz and there is no wavelength interference with the elements normally found in this crystal. During the attack, Boron and Silicon are lost because they give off volatile compounds with the fluorine atoms. Fortunately, we can obtain a very strong Boron complex with manitol, a sugar not destroyed in hydrofluoric acid. By this addition we are also able to know the Boron content of the quartz.

Analysis results of natural and cultured quartz are given. Inside a cultured quartz bar we can distinguish the X, Y, Z, S growth zones by the impurities they have trapped.

Introduction

Many works have been undertaken to understand the behaviour of quartz resonators several phenomenons are used for this knowledge and the impruvement of the resonator characteristics : sweeping and irradiation.

By IR spectroscopy we may observe their effects and determine the resonator quality using only a IR spectrum.

In this paper we present a new step in the description of Quartz crystals using a direct chemical analysis.

Choice of the analytical method

The quartz formula is SiO_2. This crystal is considered, for common utilisations, as one of the purest among the natural materials. Unfortunately, experiments show that resonator quality may be linked to chemical impurities. So we must be able to measure the foreign atoms in this crystal even if impurities are as small as trace elements.

Several methods are useful for determining trace contaminations in a material. Some of them permit a surface chemical analysis on the sample. Such methods are very sensitive to the local impurity concentrations.

The others more chemical permit a bulk analysis because the crystal is dissolved in a solution. These methods give an average value of the element concentration. On the reverse they add, during the sample preparation, the risk of an unknown contamination.

In spite of the difficulties linked to the solution preparation we have chosen the Argon plasma emission spectrometry called ICP. The plasma temperature which in about 6 to 10 000 Kelvin gives a complete ionisation. Then we have the possibility to measure a great number of elements at the same time.

Preparation and measurement

The quartz crystal is dissolved in a hydrofluorhydric solution. Silicon atoms give with fluorine a gazeaous componant leaving the solution during the attack. Other atoms, except Boron, remain in the solution so wa can measure them.

The quartz we want to study are quartz of a very high purity, so wa have tried to avoid all causes of contamination :

those associated with the grinding,
those associated witht the chemical compounds.

Preliminary results have shown that with the quartz we have foreign atoms coming from the bowl used for grinding the crystal before the chemical attack. Then instead of a grinding we produce a thermal shock which splits the crystal before the attack. The attack is made using a reflux method avoiding as much as possible all the contaminations.

Notice the boron determination needs a special preparation. This element gives as silicon a volatile compound with fluorine. Therefore boron is lost during the attack except if it makes a stong chemical complex with manitol. On the spectrometer the measure is made by comparison with standard solutions. With each element we use two or three emission bands.

To be able to compare results obtained with different samples we add the same amount of scandium atoms in the solution. We can do it because the presence of this element is not known in quartz.

Results

After controling the efficiency of the analytical method we started analysis on different natural or cultured quartz.

Among natural quartz we have chosen quartz of a very low quality and some others we consider as good quality quartz because of the quality of the blocks and their transparency.

With wultured quartz we also have two kind of samples : those with which we only have a precut block cut in a Z growth zone, and those we have cut in blocks as grown from the autoclave.

With the last one we have been able to analyse the main growth zones : those are the both Z, -X, +S and S growth zones. So we have been able to comparethe results of samples with the same make up.

As this study concerns only the quality of quartz crystals and not the comparison between the quartz manufacterers our results are anonymously presented.

In order to compare results, we present those with which the same analysis are made. The commun elements are Al, Ca, Na, K, Li, Ti, Fe, Co with an exception with boron. This element is not measured on every sample because its importance was not understood at the beginning of this study.

Results are expressed by the number of actual atoms of a given element versus 1 000 000 silicon atoms. We also add a number equal to the sum of all the foreign atoms. This number permits a quick comparison of the impurity containt of a quartz. This number is not an absolute criterium of the crystal purity because it is linked to the number of the measured elements.

We have 68 results : 15 on natural quartz and 53 on cultured quartz. Therefore, we have a great amount of values. Before a discussion we need an introduction of the values we may find in this work.

A first indication is given table 1 with the average values of the eight analysed elements calculated from the followings samples :

4 samples of low quality natural quartz
11 samples of a higher quality natural quartz
33 samples of cultured quartz coming from 7 blanks we have cut in blocks as grown from the autoclave. With these samples we distinguish the Z, -X, +X and S growth zones. The analytical results are presented in the appendix.

	Natural Quartz		Cultured Quartz			
	Common	Clear	Z zone	+X zone	-X zone	S zone
Al	260	33	15	46.2	63.02	150.5
Ca	30	3.1	8.2	8.8	11.8	7.12
Na	39	3.71	12	26.6	24.1	74.6
K	6	0.9	2.7	3.6	4.19	2.5
Li	199	16	4.8	23.3	22.2	91.08
Ti	0.9	1.3	0.1	0.19	0.17	0.3
Fe	13	1.6	1.7	2.3	2.49	1.08
Co	0.02	0.09	0.04	0.02	0.03	0.03
B	3	0.06	4.4	0.1	0	0
	550.9	59.75	49.04	110.51	128	327.29

Table 1 : *Average value of the eight elements in natural quartz and in the four growth zones obtained in Y cuts.*

The quartz we call low quality quartz are smoked or pink or amethyst quartz. Their impurity containt is very high compared to the others we have called pure quartz.

With the cultured quartz we find the expected results. The average content of foreign atoms is the lowest with the Z growth zones. Then we have the +X, -X and S zones. We can find the same classification if we compare the IR spectra of these growth zones.

Now consider the first two elements : aluminium and calcium. They have a very different behaviour. The Z growth zones have less aluminium then the others. With the calcium atoms we can find a more regular distribution between the growth zones.

Comparison of the Z growth zones of a same quartz

In table 2 we present a comparison of the foreign atom content of the both Z growth zones coming from the same quartz block. Seven differents quartz are presented.

	AZ1	AZ2	BZ1	BZ2	CZ1	CZ2	DZ1	DZ2
Al	20,45	18,23	3,78	5,33	28,33	52,01	11	13,8
Ca	1,79	15,86	3,59	7,93	6,63	7,58	7,63	16,2
Na	11,48	9,65	2,86	3,36	18,42	23,51	5,48	19,3
K	2,1	1,68	1,01	1,04	3,8	5,2	2,45	7,97
Li	0,86	0,21	0,86	0,86	12,36	0,99	1,72	1,72
Ti	0	0,12	0,12	0,12	0,28	0,21	0,12	0,37
Fe	0,85	2,14	0,96	1,39	2,82	2,42	1,71	1,5
Co	0	0,2	0,1	0	0,11	0,16	0,1	0
B	0	0	0	0	0	0	0	0
	37,53	56,09	12,48	20,03	72,75	100,1	30,21	60,8

	EZ1	EZ2	FZ1	FZ2	GZ1	GZ2
Al	14	37,58	12	15,89	2,22	2,22
Ca	12,9	28,44	2,94	3,44	12,7	5,68
Na	12,5	23,48	8,09	11,5	1,04	2,08
K	2,91	0,64	0,61	0,61	0,92	1,53
Li	0	8,64	5,18	6,39	0	21,61
Ti	0	0,12	0	0	0,12	0,25
Fe	2,9	2,68	3,13	2,36	1,18	2,14
Co	0	0	0,14	0,1	0	0
B	0	0	0	0	0	0
	45,2	101,6	32,09	40,29	18,2	35,51

Table 2 : *Chemical composition of both growth zones one seven Y cut.*

There is a very important difference between the two Z growth zones of a same cultured quartz. There is a ratio close to 2/3 between the total foreign atom content. If we consider this effect on each element there are no correlation between the observed variations. But a difference is visible between the two zones. This effect is not associated to a low quality quartz. Normaly a quartz with a containt of 3 or 2 aluminum atom versus 1 000 000 silicon atom is considered as a good quartz (samples B and G).

This result is a confirmation of our previous observations made with IR spectra. However such a difference between the Z growth zones is inexpected. Figure 1 we have both Z growth zones of the E sample spectra when the observation is made at the nitrogen temperature.

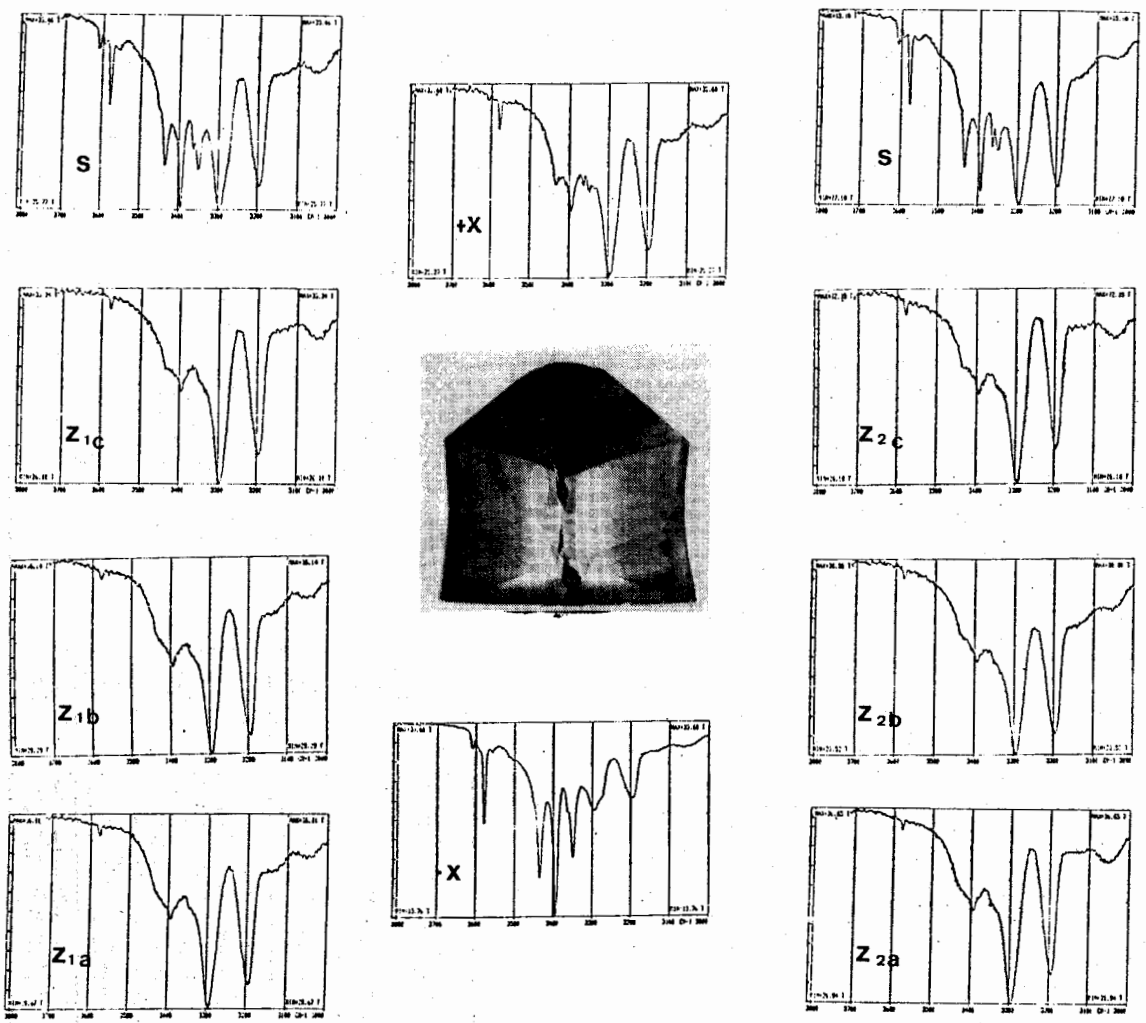


Figure 3 : *IR spectra of ten parts cut in a Y cut where the Z growth zones are blackened by gamma irradiation.*

		QUARTZ	H							
	H Z1A	H Z2A	H Z1B	H Z2B	H Z1C	H Z2C	H +X	H -X	H S1	H S2
Al	17,9	5,33	4,66	4,75	5,44	6	38,24	57,81	109,63	114,3
Ca	7,63	2,39	0,5	0,62	6,01	2,44	1,37	5,08	5,98	0,74
Na	37,68	3,75	2,92	3,86	4,56	9,91	14,19	22,94	36,35	41,23
K	12,36	0,21	0	0,69	0,66	2,96	0,64	0,32	1,05	1,79
Li	2,59	2,5	1,72	1,72	1,29	1,12	21,96	25,28	84,39	93,9
Ti	0,16	0	0,15	0,06	0	0	0	0,06	0,06	0,16
Fe	0,69	0,37	0,19	1,39	1,12	0,37	0	0,42	0	0,15
Ca	0	0	0	0	0	0	0,08	0,13	0,12	0
B	55,59	5,65	3,66	0	3,99	9,98	0,99	0	0	0
	134,6	20,2	13,8	13,09	23,07	32,78	77,67	113,04	237,58	252,27

Table 3 : *Chemical analysis of the ten parts presented figure 3.*

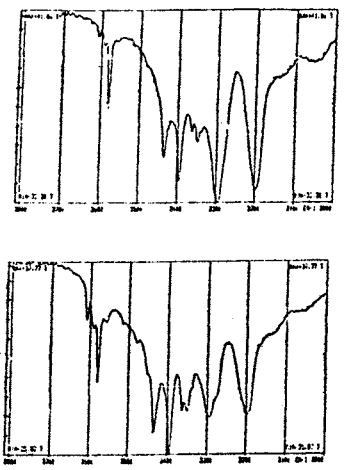
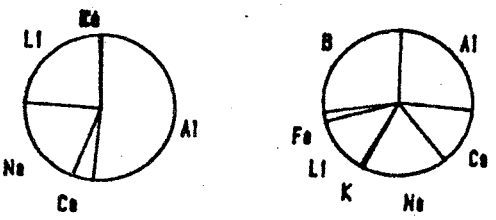


Figure 1 : *IR spectra of both Z growth zones, sample E.*

Electrical equilibrium

The electrical equilibrium inside the quartz may be evaluated. Remember that if we have substitution of a tetravalent silicon atom by a trivalent aluminum atom we must have in the neiborough of the aluminum atom a positive electrical charge.

Let us take two characteristic examples. In the first one we have almost as many aluminum atoms as alcaline atoms figure 2.a. In the second one we have nearly the same equilibrium but (figure 2.b) the silicon atom substitution is made by both aluminum and boron atoms (boron as aluminum is a trivalent atom). If we calculate precisely this electrical equilibrium or if we consider other examples we may observe that this equilibrium is not exact. We must not forget that the number of foreign atoms concerned is very small in comparison with those of silicon. In the presented examples there are respectively 77 and 20 atoms versus 1 000 000 silicon atoms. That only indicates that the actual equilibrium is obtained by several other ways as hydrogen atoms, not seen by ICP analysis, or electron holes.



Sample H - X Sample H Z2A

Figure 2 : *Chemical composition (in number of atoms) of two characteristic quartz.*

Study of a quartz showing coloration in the Z growth zone

Normally when a quartz is irradiated, the Z growth zones are not colored. With the sample called H we have observed a blackening in some parts of the Z zones. The sample has been cut in ten parts corresponding to the normal growth zones and to the abnormally colored zones in the Z zones, Figure 3.

Results of IR spectrometry at the liquid nitrogen temperature and ICP analysis are also presented table 3.

The chemical analysis also presents some unexpected behaviour. We have a very important amount of boron atoms. The strange result is that we can believe that some parts of the quartz are good if we only take into account the observed aluminum content.

Observation of precut Z growth zones

Again the presented samples come from different manufacturers. The results table 4 indicate that we may find some very good quartz with a very low aluminum content. We observe again boron and calcium atoms their presence is also unusual. But we are unable to observe the previous result saying both Z zones are different.

		precut	QUARTZ		Z zone	
	I	J	K	L	M	N
Al	3,78	4,82	8,67	16,9	3,11	8
Ca	31,88	10,97	32,33	23,5	2,69	10,53
Na	12,31	3,39	4,69	68,11	0,52	15,58
K	3,37	0	1,07	22,24	0,76	1,87
Li	0,86	0	0,86	6,05	1,29	0
Ti	0,15	0	0,37	0,25	0,12	0
Fe	1,07	1,13	4,19	2,14	1,28	4,13
Co	0,5	0	0	0	0,4	0
B	0	0	0	0	0	0
	53,92	20,31	52,18	139,19	10,17	40,11

	O	P	Q	R
Al	3,2	2,46	5,69	5,78
Ca	2,9	3,87	2,76	4,02
Na	7,04	1,25	0,57	0,86
K	2,45	0,69	0,72	0,49
Li	0	0,6	0	0
Ti	0	0,01	0,06	0
Fe	0,39	0,4	0,27	0,23
Co	0	0,03	0	0,24
B	0	0	0	0
	15,98	9,31	10,07	11,62

Table 4 : *Chemical analysis of some precut blocks in the Z growth zone.*

Observation of natural quartz

The chemical analysis on the sample we have considered as good quality quartz are presented table 5. The aluminum content of these quartz is high if we compare to the previous results concerning cultured quartz. The calcium content in reverse seems to be small.

	Natural	Quartz			
	QN 2	QN 3	QN 4	QN 5	QN 6
Al	17,56	20,28	72,27	49,58	33,8
Ca	2,18	0,92	18,09	2,09	5,08
Na	4,12	3,65	14,43	3,36	6
K	0,12	0,5	4,57	0,7	2,45
Li	15,39	23,34	33,37	15,3	16,42
Ti	0	0,33	1,11	4,42	1,37
Fe	0	0	14,47	0	1,61
Co	0	0	0	0	0
B	0	0	0	0	0
	39,37	49,02	159,11	75,45	66,73

QN 7	QN 8	QN 9	QN 10	QN 11	QN 12
37,13	27,57	32,02	25,12	27,35	27,35
0	0	1,79	1,12	1,27	1,04
0	1,04	3,39	0,44	4,36	0
0	0,55	1,22	0	0,53	0
16,42	12,97	16,42	10,54	18,5	0
1,37	1	1,37	0,6	0,56	2,5
0,48	0,64	0,75	0	0	0,32
0,3	0,25	0,25	0	?	0,25
0	0	0	0,33	0,33	0
55,7	44,02	57,21	38,15	53,92	31,46

Table 5 : *Chemical analysis of natural quartz of "good quality".*

Conclusion

All these chemical analysises we have presented indicate that quartz is a very strange material.
Two main results may be considered :

First in a cultured quartz the both Z zones are not equivalent. It is possible to distinguish them. This is important if we need the best quartz we may have.

Second in a quartz we may have many unexpected atoms as calcium and boron.

Calcium is principally found in cultured quartz. Perhaps it is possible to avoid such a contamination during the preparation of the autoclave. If we have calcium atoms in quartz, the calcium may be associated to aluminum atoms giving locally an association close to the plagiaoclase $CaAl_2Si_2O_8$.

The presence of some boron atoms in quartz seems to be quite natural. This atom is as aluminum a trivalent atom it may enter in the quartz structure replacing a silicon atom.
We do not know the consequences of boron atoms on the behaviour of a quartz resonator, but is we have boron atoms we have less aluminum atoms and the quartz seems less contaminated.
All theses results are preliminary results we must analyse more quartz samples and correlate their properties as resonator with their impurity content.

Acknowledgement

We would like to express our thanks to the D.R.E.T (Direction des Recherches Etudes et Techniques) of the Ministry of Defense who sponsored this research.

Appendix

Chemical analysis of seven samples cut in Y cuts.

		QUARTZ A		
	A Z1	A Z2	A +X	A -X
Al	20,45	18,23	37,13	117,41
Ca	1,79	15,86	3,14	6,28
Na	11,48	9,65	15,92	61,87
K	2,1	1,68	3,37	12,12
Li	6,86	8,21	24,21	79,98
Ti	0	0,12	0	0
Fe	0,85	2,14	2,36	1,5
Co	0	0,2	0	0
B	0	0	0	0
	37,53	56,09	86,13	278,36

		QUARTZ B			
	B Z1	B Z2	B +X	B -X	B S
Al	3,78	5,33	52,7	47,81	78,72
Ca	3,59	7,93	9,88	21,1	5,83
Na	2,06	3,36	24,14	24,81	31,21
K	1,01	1,84	5,67	6,32	2,3
Li	0,86	0,86	25,07	20,23	53,57
Ti	0,12	0,12	0,37	0,37	0,25
Fe	0,96	1,39	1,5	1,82	1,61
Co	0,1	0	0	0	0
B	0	0	0	0	0
	12,48	20,83	119,33	122,46	173,49

		QUARTZ C		
	C Z1	C Z2	C +X	C S
Al	28,33	52,91	86,95	231,93
Ca	6,63	7,58	14,85	16,94
Na	18,42	23,51	56,45	125,4
K	3,8	5,2	4,23	5,53
Li	12,36	8,99	50,41	139,56
Ti	0,28	0,21	0,65	1,32
Fe	2,82	2,42	4,69	3,39
Co	0,11	0,16	0,07	0,04
B	0	0	0	0
	72,75	100,88	217,4	515,11

		QUARTZ E			
	E Z1	E Z2	E S	E +X	E -X
Al	14	37,58	172,78	57,15	76,05
Ca	12,87	28,44	28,2	15,86	16,76
Na	12,52	23,48	23,74	33,92	0,26
K	2,91	0,64	2,76	2,76	0
Li	0	0,64	10,37	21,61	0
Ti	0	0,12	0	0	0,25
Fe	2,9	2,68	2,57	2,47	7,73
Ca	0	0	0	0	0
B	0	0	0	0	0
	45,2	101,58	232,42	133,77	101,05

		QUARTZ D			
	D Z1	D Z2	D +X	D -X	D S
Al	11	13,78	36,69	64,94	127,19
Ca	7,63	16,16	11,37	11,67	6,88
Na	5,48	19,31	27,14	58,37	57,41
K	2,45	7,97	8,74	7,97	3,52
Li	1,72	1,72	13,83	21,61	59,23
Ti	0,12	0,37	0,25	0,25	0,37
Fe	1,71	1,5	2,9	2,57	1,5
Ca	0,1	0	0	0,1	0
B	0	0	0	0	0
	30,21	60,81	100,92	158,58	256,1

		QUARTZ G			
	G Z1	G Z2	G -X	G +X	G S
Al	2,22	2,22	26,9	19,79	76,94
Ca	12,72	5,68	11,97	9,37	15,71
Na	1,04	2,08	9,13	7,82	60,28
K	0,92	1,53	2,6	2,14	23,24
Li	0	21,61	7,78	7,78	21,61
Ti	0,12	0,25	0,25	0,18	0,75
Fe	1,18	2,14	2,95	2,57	4,4
Ca	0	0	0,03	0	0,3
B	0	0	0	0	0
	18,2	35,51	55,61	49,65	203,23

		QUARTZ F			
	F Z1	F Z2	F +X	F -X	F S
Al	12	15,89	42,2	57,15	169,67
Ca	2,94	3,44	5,01	10,19	5,08
Na	8,09	11,5	33,14	PAS	113
K	0,61	0,61	1,22	PAS	0,67
Li	5,18	6,39	21,87	PAS	87,33
Ti	0	0	0,12	0,07	0
Fe	3,13	2,36	2,36	0,45	0,36
Ca	0,14	0,1	0	0	0
B	0	0	0	0	0
	32,09	40,29	105,92	67,86	376,11

ION BEAM MILLING : CHARACTERIZATION AND PREVENTION OF SURFACE DAMAGE.

N. VIALLE - J-P AUBRY - D. BLONDE

C.E.P.E.
44, avenue de la Glacière - 95100 ARGENTEUIL - FRANCE

ABSTRACT

Ion Beam Milling is a technology commonly used to manufacture very thin quartz plates, so that the final frequency obtained exceeds several hundred MHz in fundamental mode with AT-cut.

The traditional chemical surface damage occuring during chemical removal is no longer observed with this process. Nevertheless, certain typical imperfections such as hillocks and pits remaining are incompatible with the properties required for UHF and VHF resonators. So, it is important to understand their origin to overcome this major drawback.

In this paper, a few experiments are reported to show the influence of the initial surface finish on the final topography of the quartz plate. Futhermore, Scanning Electronic Microscopy (SEM) is used to point out the connection between the defects introduced by mechanical lapping and polishing and the surface damage resulting from ion beam etching.

A preparation of the quartz plates is proposed to achieve a smooth surface despite the high rate of material removed. Moreover, we will show that this technology does not generate imperfections. On the contrary, it may polish the surface in some particular cases. In conclusion, Transmission Electronic Microscopy (TEM) is applied to prove that these imperfections are not connected with a crystallographic problem. A graphic model is proposed to explain the appearance and the evolution of these typical ion beam pits.

INTRODUCTION

La fabrication de résonateurs à quartz implique la connaissance et la maîtrise des technologies spécifiques de son usinage. Les méthodes mécaniques habituelles, rodage et polissage, suffisent à la réalisation de structures simples, pour des lames "plan-plan" ou "plan-convexe" par exemple. Mais des configurations plus complexes sont conçues à partir d'autres procédés comme l'usinage ultra-sonore, l'usinage ionique et l'usinage chimique.

La dissolution chimique dans un bain de bifluorure d'ammonium est tributaire de son caractère anisotrope qui détériore, plus ou moins rapidement selon l'angle de coupe du matériau, l'état de surface des lames **(1-3)**. Or l'obtention, par procédé mécanique d'abord puis la conservation du poli optique s'il intervient d'autres usinages ultérieurement, est une garantie d'un bon fonctionnement et d'un bon coefficient de qualité Q du résonateur.

Les méthodes classiques de rodage et de polissage ne permettent, quant à elles, que l'obtention de fréquences beaucoup plus basses : 50 MHz en coupe AT:35, soit une épaisseur finale de 35 microns pour les lames.

L'usinage ionique a permis de réaliser des résonateurs matérialisés par des lames extrêmement fines, pouvant fournir des fréquences de l'ordre de plusieurs centaines de Mégahertz en coupe AT **(4-9)**, voire du Gigahertz **(10)**. C'est ce qui a permis son développement dans le milieu industriel, notamment pour réaliser l'amincissement central des lames.

II - OBSERVATIONS DES DEFAUTS

La structure "MESA inversé" (cf. Figure 1) avec des coupes AT:35 est au centre de notre étude.

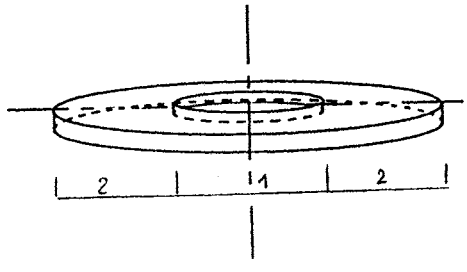


Figure 1 : Structure "MESA inversé"

Un iris masque le pourtour de la lame et l'amincissement extrême se fait uniquement au centre, au niveau de la zone vibrante où seront déposées les électrodes.

Nous pouvons donc observer à la fois la surface, usinée ioniquement (zone 1) et la surface d'origine (zone 2) pour comparaison.

Nous avons concentré notre recherche sur les motifs apparus dans la zone 1 . Ils sont caractéristiques mais peuvent être très différents : ce peut être des bosses (cf. Figure 2), des creux et des cônes (cf. Figure 3).

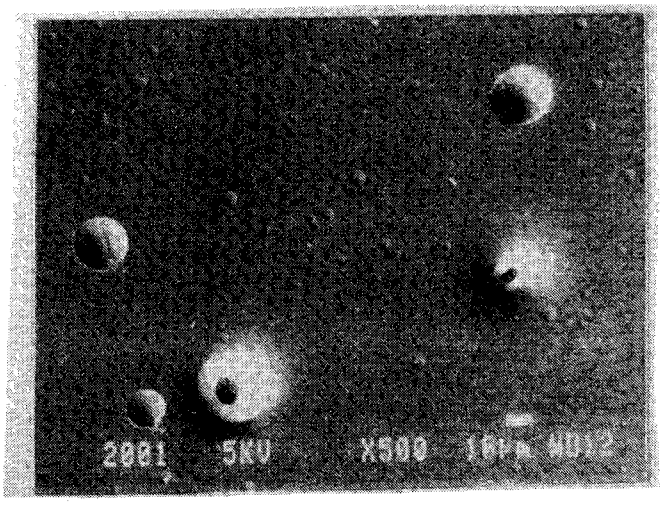


Figure 2 : Bosses révélées par usinage ionique.

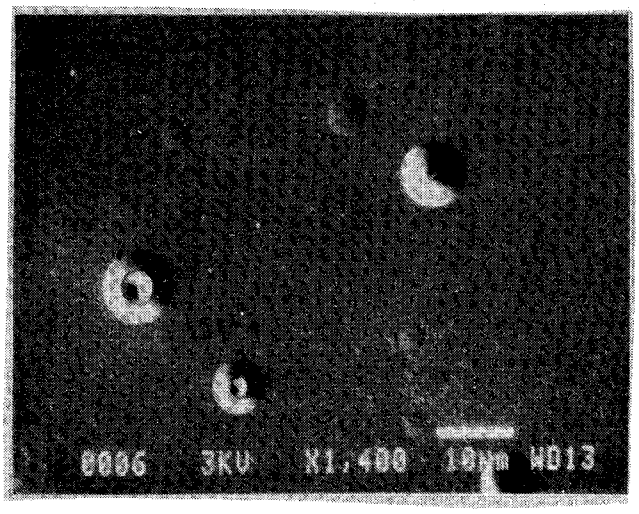


Figure 3 : Cuvettes et cônes révélés par usinage ionique.

Toutes ces formations possèdent un axe de symétrie à cause de la rotation du plateau supportant les lames au cours de l'usinage.

L'usinage ionique arrache les atomes de la cible exposée au bombardement et donc, il est étroitement rattaché au phénomène de pulvérisation de la matière. C'est ainsi que la plupart des auteurs qui ont étudié le problème des structures apparues après un usinage ionique, notamment la formation des cônes sur la quartz **(11-14)** ou d'une façon plus générale sur les matériaux à fort taux de pulvérisation, attribue ce problème à un phénomène de masquage local de la surface exposée.

En fonction de la nature de la cible et de la matière évaporée, nous pouvons classifier ces états de surface en deux catégories : ceux qui possèdent des structures en crêtes, et ceux qui révèlent des structures côniques.

Ceci donne lieu à une application pratique, en anglais sous le nom "ion beam texturing" **(12)** où l'on constate en effet la génération de cônes en forte densité. Cependant, nos essais laissaient supposer qu'un autre phénomène pouvait être à l'origine de ces défauts.

III - LES ESSAIS

Un grand nombre de lames qui étaient soit directement extraites d'une production industrielle, soit préparées spécialement à partir de celle-ci, ont fait l'objet de deux préparations d'état de surface. Néanmoins, pour tenter de déterminer exactement ce qui donnait naissance aux imperfections incriminées, nous avons fait jouer les paramètres suivants :

- Qualité du matériau : standard, "premium Q", swept.
- Origine du matériau : de différents fournisseurs.
- Préparation de la surface avant l'usinage ionique : polissage mécanique traditionnel, polissage mécano-chimique.

Les phases de préparation étant d'une extrême importance, il est nécessaire de les décrire avec précision. Dans le souci de comparer les résultats, nous avons travaillé sur des lames initialement à 21 MHz (79,5 microns d'épaisseur), et qui sont portées jusqu'à 30 MHz (56 microns) par rodage puis polissage. L'usinage ionique les amène enfin aux alentours de 115 MHz (14,5 microns).

3.1/ Le polissage mécanique

Après un prérodage suivant directement le sciage des lames de quartz, s'ensuit un rodage et un polissage mécanique avec des abrasifs adaptés, dont la taille des grains diminue régulièrement afin d'aboutir à un état de surface qualifié de "poli optique". Mais, une observation soigneuse des lames en microscopie optique d'abord puis électronique à balayage, a permis de remarquer que tous les défauts qui apparaissent avec l'usinage ionique étaient en étroite relation avec le nombre de perturbations résiduelles (rayures, creux, éclats, ...) après les étapes mécaniques de préparation. Notons soigneusement qu'à ce niveau, aucun des défauts que nous avons pu observer, n'étaient protubérants.

3.2/ Le polissage mécano-chimique

Nous avons donc développé une méthode de préparation extrêmement soignée pour éliminer au mieux toutes ces imperfections, non seulement celles visibles, mais aussi celles éventuellement cachées **(15)**. Entendons par là que la couche amorphe qui résulte du polissage mécanique ne doit en aucun cas dissimuler des défauts sous un aspect "miroir" uniquement superficiel. Celui-ci résulte de ce que nous pourrions assimiler par analogie à un écrouissage

ou encore à un glaçage. Nous proposons en conséquence une préparation mécano-chimique des surfaces, incluant des trempages dans le bifluorure d'ammonium, pour éliminer toute amorphisation au fur et à mesure du polissage et éviter son effet dissimulateur.

3.3/ L'usinage ionique des lames

Les lames sont collées sur le plateau tournant, incliné à 35° par rapport au faisceau incident, avec une graisse au Silicone. La rotation du support élimine les effets d'hétérogénéité d'usinage et amène ainsi une qualité d'expérience parfaitement constante pour les différentes lames. Celles-ci sont placées suivant une disposition "camembert" (Cf. Figure 4). L'impact ionique perd sensiblement de son intensité quant on s'éloigne du centre et nous obtenons ainsi différentes profondeurs usinées qui interviennent dans la taille des motifs apparus. Par ailleurs, on pourra estimer que tous les spécimens placés sur un cercle de même rayon auront subi une même attaque et en conséquence, pour ces échantillons, si des différences de texture apparaissent, elles seront dues essentiellement à la nature des lames. Certaines d'entre elles ont été retirées alors qu'elles avaient atteint la fréquence de 70 MHz, les autres à 110 MHz. Les observations ont été faites en microscopie électronique à balayage.

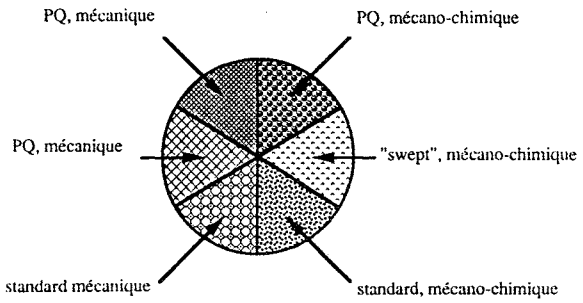


Figure 4 : Essai en disposition "camembert"

Dès l'inspection en microscopie optique, nous avons pu constater que les lames ayant subi la préparation mécano-chimique avaient un taux très bas, voire nul, de défauts, contrairement aux autres lames du plateau.

IV - RESULTATS ET INTERPRETATION

4.1/ Résultats des essais

L'observation en microscopie électronique à balayage a fourni des clichés très significatifs (cf. figures 5, 6, 7, 8).

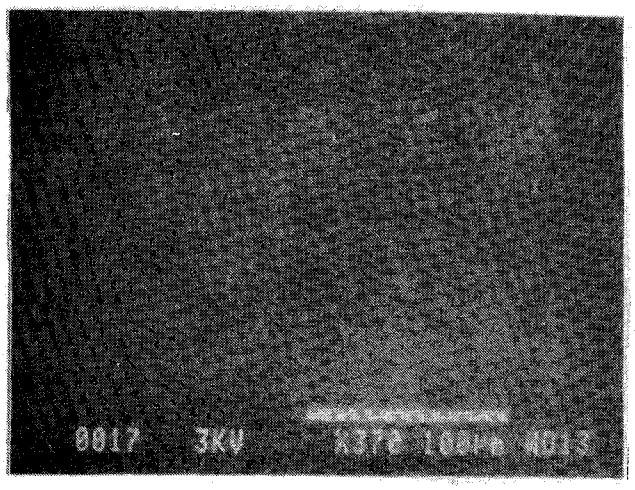


Figure 5 : Surface initiale "polimécanique".

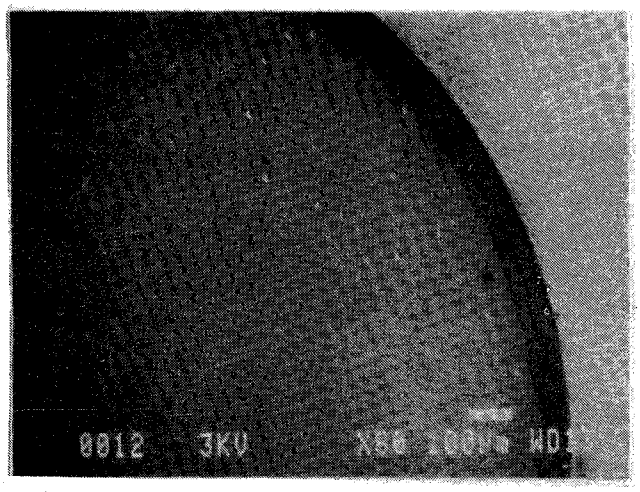


Figure 6 : Surface "poli mécanique" après usinage ionique.

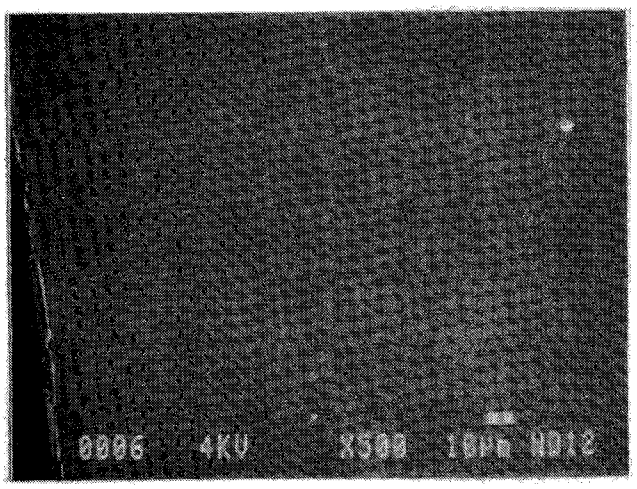


Figure 7 : Surface initiale "poli mécano-chimique".

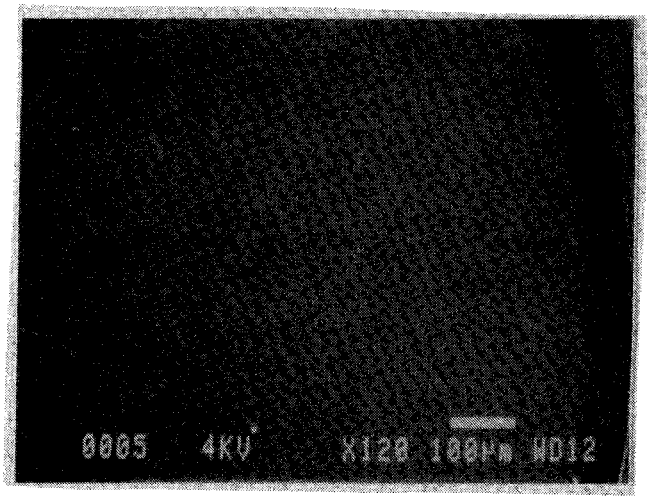


Figure 8 : Surface "poli mécano-chimique" après usinage ionique.

Nous remarquons nettement que l'apparition des cônes est en lien étroit avec la présence de défauts résiduels sur la surface, essentiellement des creux, et nous pouvons dire dès à présent que :

L'état de surface final obtenue après un usinage ionique dépend principalement de l'état de surface initial de la lame de quartz.

L'explication par masquage local est exclue du fait de la disposition "camembert". Suivant les secteurs, nous pouvions observer soit un fort taux de défauts (cônes ou autres), soit quasiment aucun, respectivement pour une préparation mécanique ou mécano-chimique de la surface initiale. Toutefois, le défaut initial lui-même doit être considéré, comme le montre la figure 9.

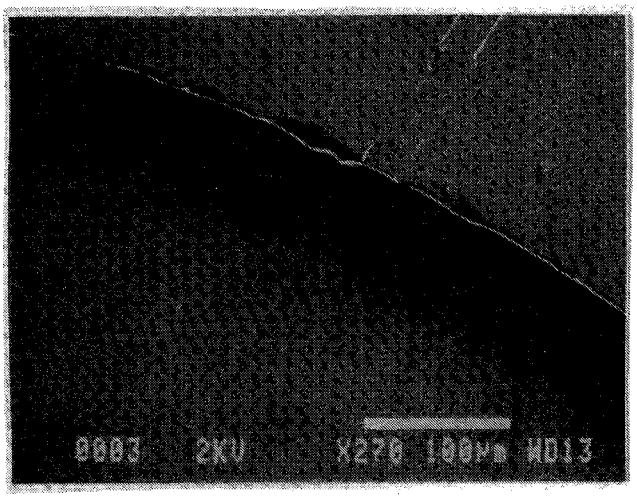


Figure 9 : Défauts de rayures.

Nous constatons en effet sur ce cliché que toutes les rayures présentes initialement n'ont pas été translatées par l'usinage ionique. La profondeur de la perturbation et l'inclinaison de ses flancs par rapport au plan de référence interviennent. Nous allons donc développer une hypothèse sur l'évolution de ces défauts que nous avons appelée **"l'effet d'ombrage"**. En effet, une analyse par spectroscopie rayons X n'a décelé aucun élément chimique résiduel, susceptible d'avoir protégé localement la surface. D'autre part, une étude en microscopie électronique à transmission n'a révélé aucune perturbation cristallographique interne sur les sites des défauts.

4.2/ Effet d'ombrage

C'est celui qui est a priori à l'origine des défauts que nous avons observés lors de nos essais.

Après un polissage imparfait, il subsiste généralement à certains endroits, des petits creux que nous représenterons en coupe par deux flancs (figure 10). Sous une incidence non verticale par rapport à la surface de quartz, les bords de ces trous peuvent protéger le fond du bombardement ionique. Cela se produira si la pente de ces flancs, mesurée par rapport à la surface de référence, est supérieure à l'angle d'incidence du faisceau, soit d'après le dessin ci-dessous, si $\gamma > \alpha$. C'est ce que nous avons appelé le phénomène d'ombrage.

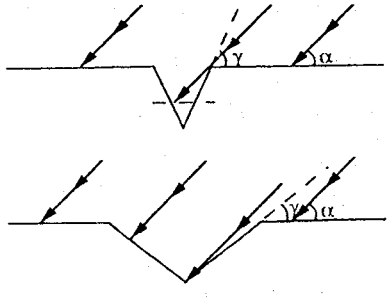


Figure 10 : Effet d'ombrage.

Dans notre étude, nous avons pris en compte l'effet de l'angle de coupe du matériau au niveau de l'apparition des défauts de l'usinage ionique, en exploitant une courbe **expérimentale**. Cette courbe établie pour une coupe AT:35 inclut :

- l'effet microscopique du degré d'orientation du faisceau d'ions,
- la variation de la densité du faisceau lié à l'orientation de l'échantillon.

Le développement des cônes peut donc s'envisager suivant le processus décrit ci-après (figure 11).

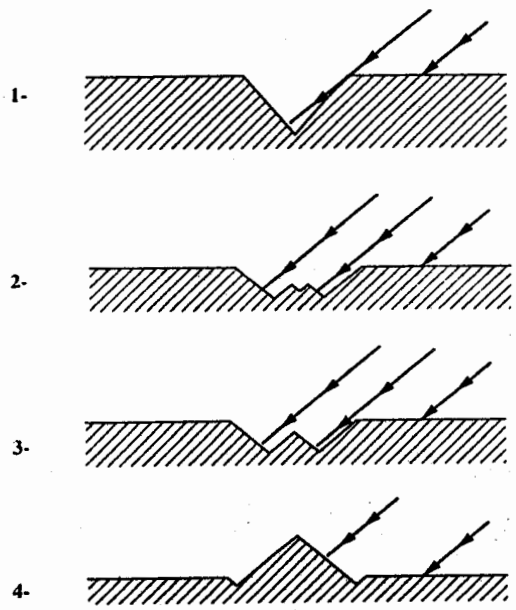


Figure 11 : Principe de croissance des cônes par l'effet d'ombrage.

Nous en déduisons deux conséquences. La première est que les cônes peuvent avoir différentes hauteurs correspondant à leur temps de croissance au centre de la cuvette. La seconde conséquence est que le flanc des cônes possède des pentes liées à l'angle d'incidence du faisceau d'ions. Les travaux de certains auteurs **(16)** confirment cette dépendance "angle du cône - angle d'incidence) à travers leurs observations.

En fait, si à l'origine l'angle d'incidence du faisceau est de α, l'angle d'incidence maximum qui frappera le flanc du cône lors d'une révolution du plateau tournant sera de 2α (figure 12).

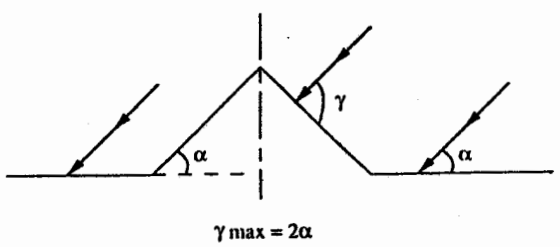


Figure 12 : Dépendance "angle du cône - angle d'incidence".

Pour un cône donné, toutes ces génératrices seront frappées par le faisceau sous une incidence variant de 0° à 2α (figure 13).

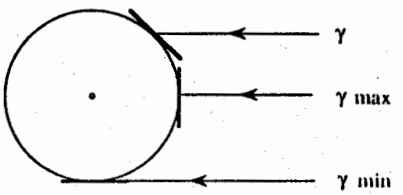


Figure 13 : Evolution de l'angle d'incidence sur le cône.

Nous pouvons donc penser que l'attaque du flanc des cônes sera maximale, et donc entrainera leur diminution, voire leur disparition totale, si pour une rotation du plateau de support des lames, ces flancs subissent un usinate maximal, donc si :

$$I = \frac{1}{2\alpha} \int_{\gamma=0}^{\gamma=2\alpha} V_I(\gamma).d\gamma$$

est maximale.

V_I est la vitesse d'usinage du quartz en fonction de l'angle d'incidence du faisceau ionique. A partir de la courbe expérimentale de cette vitesse (cf. figure 14), nous avons tracé la courbe $I = f(\alpha)$, et nous constatons que **pour un maximum d'usinage à l'angle d'incidence** $\alpha_{max} = 50°$, I **est maximale pour** $\alpha_{Imax} = 33°$ (cf. figure 15).

Les observations expérimentales confirment ce résultat **(4, 8, 17)** : on emploie couramment un angle = 35° environ, pour justement minimiser au maximum l'apparition des défauts ioniques.

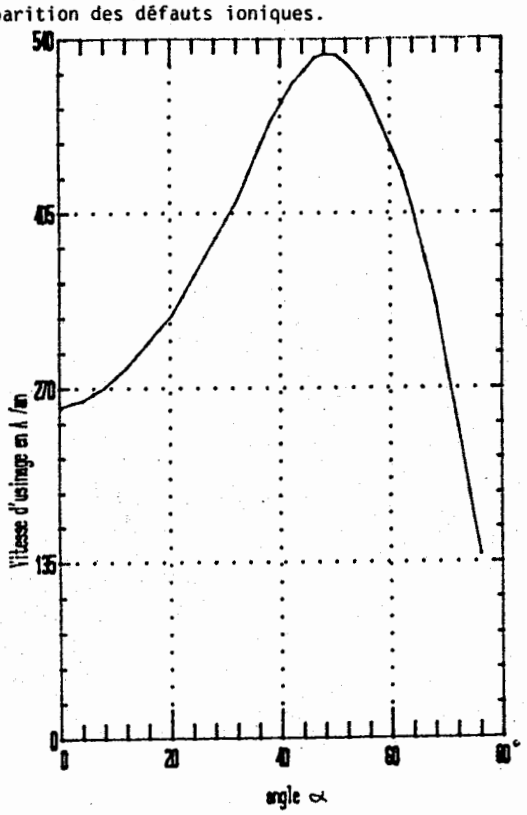


Figure 14 : Courbe V_I (α)

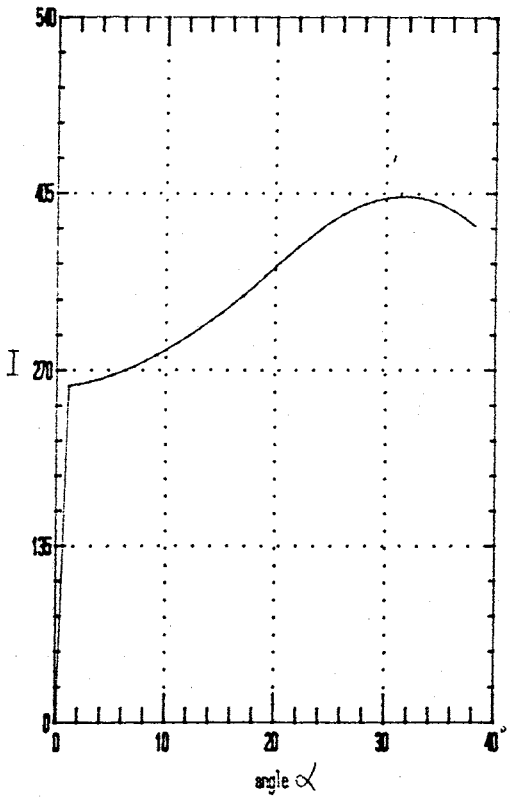


Figure 15 : Courbe I = f(∝).

V - MISE EN EVIDENCE DU "POLISSAGE" IONIQUE

Maintenant qu'il est possible d'obtenir des lames usinées par voie ionique possèdant un bon état de surface, nous avons tenté, sur un certain nombre d'échantillons, une dissolution chimique.

Ainsi, la dissolution dans le bifluorure d'ammonium s'effectue simultanément sur la surface polie mécano-chimiquement (surface initiale) et sur la surface "ionique" d'une même lame. Nous pouvons comparer la densité des motifs chimiques révélés (figures 16 et 17) sur chacune.

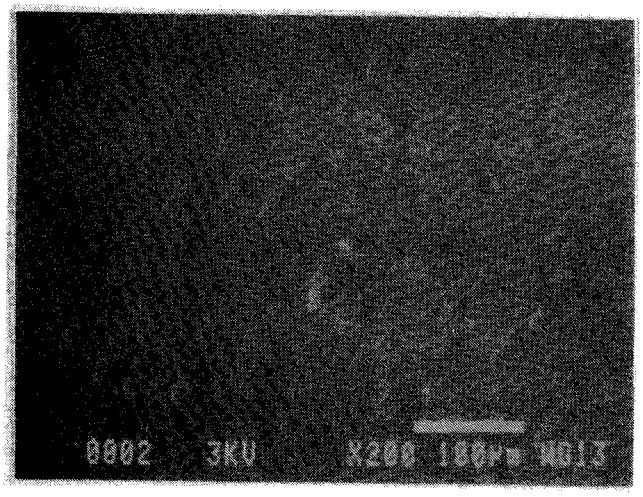


Figure 16 : Surface initiale après dissolution chimique.

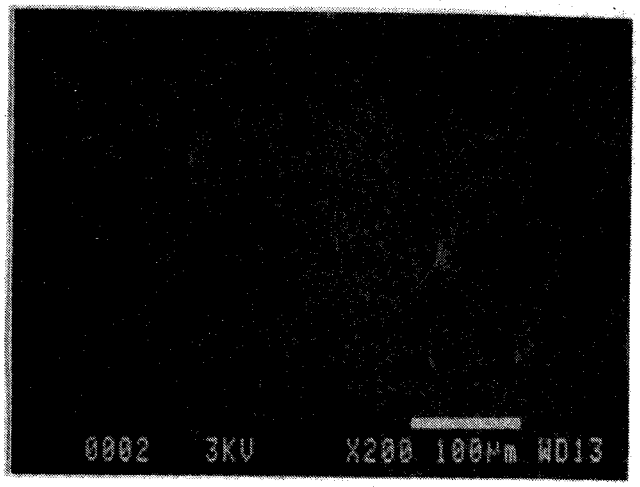


Figure 17 : Surface "ionique" après dissolution chimique.

Nous constatons que la densité des motifs est nettement supérieure sur la surface initiale "mécano-chimique". Comme ces motifs ont tendance à apparaître sur les perturbations de la surface au cours de sa dissolution **(2)**, nous pouvons donc parler d'un réel **polissage ionique**. En fait, il ne sera efficace qu'à partir d'une surface initiale déjà soignée, et en sera d'autant plus perceptible.

VI - CONCLUSIONS

Ces interprétations sur l'apparition de défauts ioniques sont toutes deux basées sur des constatations expérimentales. Elles ont fourni une aide à la compréhension de leur origine, ombrage ou masquage, et permettent donc d'envisager des solutions concrètes. Pour l'effet dit d'ombrage, la méthode du polissage mécano-chimique a fourni des résultats satisfaisants. Pour éviter le masquage, des solutions techniques comme l'usinage ionique réactif ont été développées **(18)**. Le contrôle des états de surface en usinage ionique est capital. Tout d'abord, il permet la fabrication de résonateurs VHF et UHF possèdant un bon coefficient de qualité, mais il ouvre également d'autres horizons au niveau de l'usinage chimique.

REFERENCES

(1) TELLIER - 1984
Surface Technology, 21, p 83-89.

(2) TELLIER, VIALLE, VATERKOWSKI - 1986
40th Annual Frequency Control Sumposium, p 76-85.

(3) TELLIER, VIALLE, VATERKOWSKI - 1987
1er Forum Européen Temps-Fréquence, Besançon, mars 1987, p 159-167.

(4) VALENTIN - 1980
Colloque MICRONORA, octobre 1980, LCEP/ENSMM, La Bouloie, Route de Gray, 25030 Besançon Cedex.

(5) WANG, WATSON - 1984
38th AFCS, p 101-104.

(6) VALENTIN, MICHEL, BESSON - 1979
CR Académie des Sciences, Paris, t. 289, 22 octobre 1979, Série B, p 155-158.

(7) CASTELLANO, HOKANSON - 1975
29th AFCS - p 128-134.

(8) BERTE - 1977
Electronic Letters, Vol. 13, n° 9, p 248.

(9) BERTE, HARTEMANN - 1977
Thomson SCF, Labo Centre de Recherche Corbeville, 91401 Orsay, Communication.

(10) TANSKY - 1978
Ultrasonics Symposium Proceedings, Cat. 78, Ch. I., p 433.

(11) HUDSON - 1976
Nasa Technical Memorandum.

(12) HUDSON - 1977
Journal of Vacuum Sci. Technology, Vol. 14, n° 1, janvier/février 1977, p 286-289.

(13) WITCOMB - 1974
Journal of Materials Sciences 9, p 551-563.

(14) STEWART, THOMSON - 1969
Journal of Materials Sciences 4, p 56-60.

(15) VIG, LEBUS, FILLER - 1977
31th AFCS, p 131-143

(16) WITCOMB - 1974
Journal of Materials Sciences 9, p 551-563.

(17) WEISSMANTEL - 1976
Le Vide, n° 183, mai/juin/juillet 1976, p 107-116.

(18) WANG, WATSON - 1984
38th AFCS, p 101-104.

SAW VOLTAGE CONTROLLED OSCILLATOR USING HYBRID TECHNOLOGY

L. PENAVAIRE and H. GAUTIER

THOMSON-SINTRA-ASM/DTAS
399, route des crêtes - Parc de Sophia-Antipolis
06561 Valbonne Cedex - France

ABSTRACT

Surface Acoustic Wave (SAW) technology allows building of medium to high Q oscillators which operate in the VHF and UHF frequency range and which are very compact. The new family of VCOs described here employs a SAW delay line as the stabilizing element and makes use of hybrid techniques to limit the overall size of the component and exhibit high reliability.
Custom design of the acoustic delay line controls the center frequency, chosen between 150 and 450 MHz, and the tuneability range, below 0.6%. Very good linearity (10%) together with low phase noise floor (-140 dBc/Hz) are obtained. The source is housed in a flat 16-pin package and MTBF over 300,000 hours is predicted at 60°C.
This line of product is qualified for airborne and ground environment. Major applications are presented.

INTRODUCTION

Surface Acoustic Wave (SAW) technology allows the building of two classes of high quality oscillators : very low-noise sources which make use of resonators, and medium stability fixed or variable oscillators which employ a delay line as the stabilizing element (Fig. 1).
We wish here to deal with the latter class and to present a new family of voltage controlled oscillators (VCOs) where hybrid technology is used to reduce the overall size, achieve very high reliability and allow for a full qualification of the product.

1. - DESIGN

The new product family bears the name SLC5000. The major design goals are summarized in Table 1.

Center frequency	150 to 450 MHz
Freq. control range	0.1 to 0.7 %
Freq. control linearity	< 10%
Output power level	> 10 dBm
Phase noise floor	< -130 dBc/Hz
Temperature range	-55 to +85°C
Bias and control voltage inputs	
Flat Package with thickness < 5 mm	

TABLE 1 - Design goals of SLC5000

The general architecture of the oscillator is that schematically represented in Fig. 1 : the oscillating loop consists of a wideband amplifier in series with a SAW delay line and an electronic phase shifter. The delay line operates as the stabilizing element and the equivalent Q-factor of the loop is given by

$$Q = \pi F \tau$$

where τ is the acoustic delay.

A delay line oscillator is in principle multi-frequency, the various frequencies being $1/\tau$ apart. To achieve single frequency operation, one makes use of the bandpass filter characteristic of the delay line : the 6-dB bandwidth will be chosen approx. equal to $1/\tau$ (see Fig. 2). Finally, the phase shifter will allow the frequency of oscillation to be varied provided the loop gain margin remains positive. The maximum frequency deviation is

$$\Delta F = \Delta\Phi_{max} / 2\pi\tau$$

where $\Delta\Phi_{max}$ is the maximum phase shift.

To obtain a good to medium stability source, the minimum Q-factor is set at 200. In addition the maximum practical phase shift is around 140 deg. The resulting maximum relative frequency deviation of the VCO is therefore limited to approximately 0.6%.
The minimum frequency deviation relates to frequency stability : temperature, initial setting and ageing effects amount to approximately 400 ppm and the practical lower limit of the frequency deviation is around 0.1%.

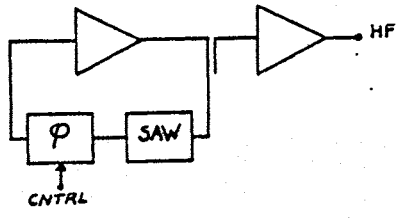


Fig.1 - Block diagram of a SAW delay line oscillator

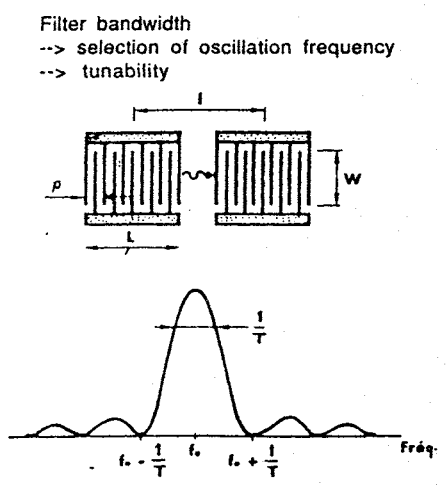


Fig.2 - Role of the delay line in a SAW oscillator.

2. - MANUFACTURING

The oscillator is made of three sub-assemblies : the delay line and its tuning circuits, the loop and output electronics, the phase shifter (see Fig. 3).

The delay line is built on quartz, according to the basic principles presented above. The crystal cut chosen corresponds to a very high turn-over temperature close to 90°C. The high linear temperature coefficient of the electronic circuitry will make the final overall turnover point go down to near 0°C. Input and output tuning inductances built on alumina are added.

The loop electronic circuits and the output buffer amplifier are designed and built using hybrid technology. Thick film resistors and thin film inductances are integrated on the alumina substrate ; chip capacitors, transistors and one silicon IC are added after.

The phase shifter makes use of a coupled microstrip hybrid junction loaded by individually packaged varicaps. The external "bias" voltage is directly applied to the varicaps, however the "control" voltage is first amplified using an operational amplifier before being fed to the other side of the varicaps. In addition to greater flexibility in the choice of the voltage control range, this allows the frequency vs voltage response of the VCO to be adjusted. This operation is done through laser trimming of resistors prior to sealing.

The complete circuit is housed in either a 16-pin single-in-line flat-pack case as shown in Fig. 3 or in a 32-pin dual-in-line plug-in type case with the identical overall size (42.9 x 19.6 x 4.4 mm). Hermetic sealing under dry N_2 is required to guarantee stable characteristics.

3. - PERFORMANCES

Various oscillator models have been built. They operate between 195 MHz and 305 MHz and the frequency control range varies from 0.24% to 0.64%.

Fig.3 - Photograph of the assembled VCO.

The response of a representative delay line is given in Fig. 4 : frequency is 303 MHz, equivalent Q is 175 and loss is below 30 dB. This line is use in oscillator model SLC5103 where the maximum frequency deviation is 750 kHz. The frequency control characteristic is shown in Fig. 5 together with the 10% linearity template. The same characteristic is given in Fig. 6 for a wider band oscillator covering the range 195.55 ± 0.63 MHz.

Careful design of the electronics and acoustics ensures that temperature variations have almost no effect on control linearity (see Fig. 7) but center frequency is shifted in accordance to the quartz parabolic coefficient of 3.10^{-8} $(°C)^{-2}$. The turnover temperature is generally below 0°C due to the large linear temperature coefficient of some of the electronic components.

The phase noise power density of a free running oscillator is plotted in Fig. 8 as a function of frequency offset from carrier. A classical $1/F^2$ curve is measured from a few kHz onward. The noise floor is reached at F/2Q which corrsponds to a few hundred kHz and the floor level is -140 dBc/Hz.

A full qualification test program has been conducted following MIL-STD-883C. This included solderability, electrostatic discharge sensitivity, steady state life, temperature cycling, acceleration, vibration, shocks, moisture and salt tests. Conformity to overall level B is guaranteed.

Mean Time Between Failure (MTBF) has been computed following MIL-HDBK-217D under AIF (Airborne Inhabited Fighter) envionment conditions. MTBF over 300,000 hours was found at 60°C and 80,000 hours at 100°C.

Performances are summarized in Table 2.

Center frequency	180 to 400 MHz
Freq. setting accuracy	± 20 ppm
Temperature stability	$3.10^{-8}(\Theta - \Theta_0)^2$ $\Theta_0 = 0°C$
Output power level	8 to 12 dBm
Frequency control :	
- frequency range	0.1 to 0.7 %
- voltage range	0 - 5 V
- control linearity	<10%
- control bandwidth	DC-200kHz
Harmonic ratio	- 20 dBc
Phase noise floor	< -140 dBc/Hz
Power supply	+12V/80mA -12V/10mA
Temperature range	-55 to +85°C
MTBF (at 60°C)	300,000 hours
Package	SIL-16

TABLE 2 - PERFORMANCES OF SLC5000 SAW VCO SERIES

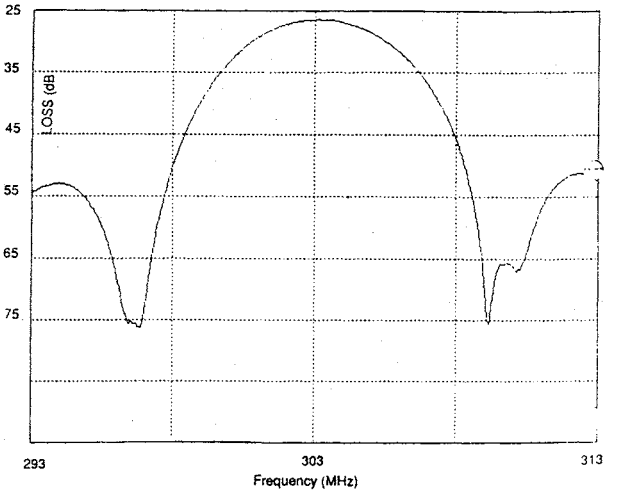


Fig.4 - Measured transfer function of a 180 ns SAW delay line centered at 303 MHz.

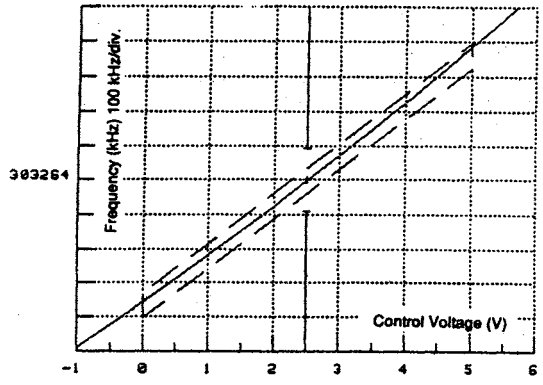


Fig.5 - Frequency control characteristic at 30°C of a narrovband VCO (F = 303.3 MHz, ΔF = 730 kHz).

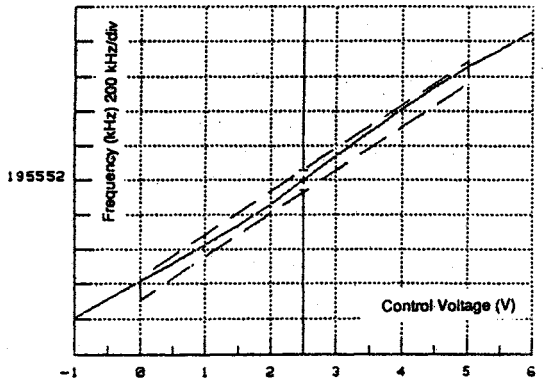


Fig.6 - Frequency control characteristic of a wideband VCO (F = 195.5 MHz, ΔF = 1250 kHz) at 30°C.

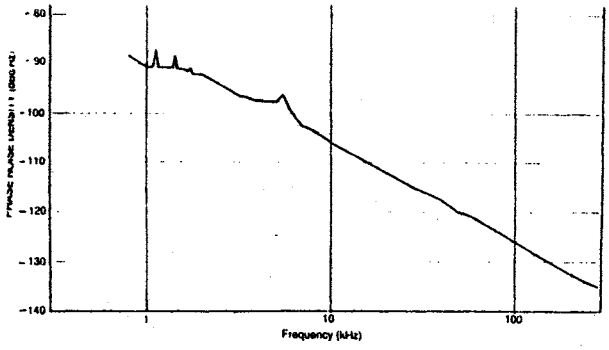


Fig.8 - Phase noise characteristics of VCO centered at 210 MHz with nominal ΔF = 700 kHz.

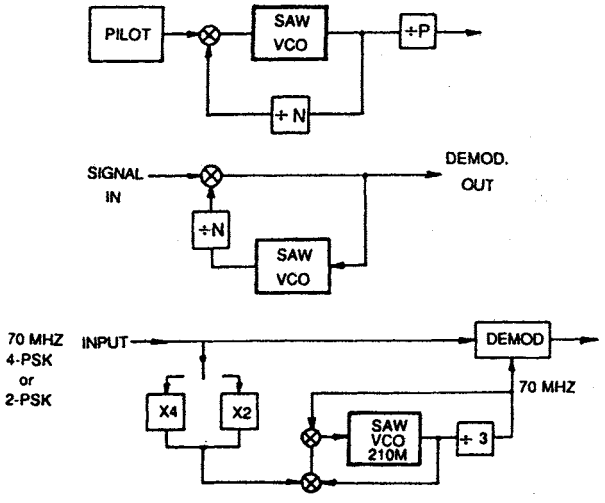


Fig.9 - Major applications of the compact SAW VCO

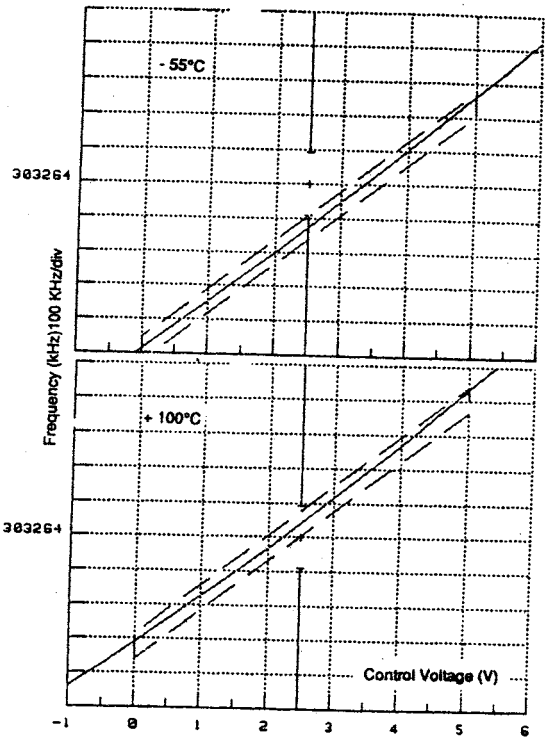


Fig.7 - Frequency control characteristics of a narrowband VCO (F = 303.3 MHz, ΔF = 730 kHz) at -55°C and +85°C.

4. - APPLICATIONS

These good electrical performances and the small size of the SAW VCO series SLC5000 make it well suited for ground, mobile and airborne applications under medium to severe environment.

Examples of actual implementations are schematized in Fig. 9. The oscillator is used in a phase-locked loop to benefit from its VCO capability.

(a) Locking the VCO on a low frequency pilot allows compact and high quality clocks to be built. The VCO bandwidth is chosen to exceed the sum of all instability causes.
(b) Frequency (or phase) demodulation of narrowband signals is simply performed through phase-locking of a VCO and reading of the control voltage signal. This application often requires large control bandwidth and values of 1-2 MHz are not uncommon.
(c) The last example relates to 70 MHz carrier recovery in a digital communication system employing 4-phase and 2-phase PSK modulation. A VCO operating around 210 MHz was developed and it now equips many modems in the French telecom network.

CONCLUSIONS

A family of Voltage Controlled Oscillators has been built using hybrid and SAW technologies. They operate in the VHF range and they offer low phase noise, good control linearity, together with small size and ruggedness. Present state-of-the-art in SAW design techniques and high frequency hybrid circuitry will allow the frequency range to be extended to 2.5 GHz. For higher frequency of operation, it is likely that the existing YIG and dielectric resonator technologies will continue to meet most requirements.

ACKNOWLEDGEMENTS

The authors wish to thank their colleague J. Elzière for his helpful collaboration in the definition of the custom hybrid circuits.

FILTRE FI FAIBLES PERTES A ONDE DE SURFACE

JP. MICHEL

C.E.P.E. 44 avenue de la Glacière 95100 ARGENTEUIL

La réalisation de filtres entre 30 et 120 MHz avec des bandes passantes voisines de 300 KHz (bandes relatives aux alentours de 0,5 %) pose des problèmes aux concepteurs car ils ont le choix entre deux technologies : 1/ les filtres à onde de volume 2/ les filtres à onde de surface.

Les deux technologies sous leurs formes traditionnelles sont distinctes, mais la progression de l'une et de l'autre les amènent à se recouper dans le domaine considéré. Elles présentent chacune leurs avantages et leurs inconvénients :

1/ Les filtres à onde de volume : ils sont traditionnellement limités à des fréquences inférieures à 35 MHz par des bandes passantes < 0,5 %.
Ils sont constitués autour de résonateurs à quartz usinés mécaniquement couplés entre eux à l'aide de capacités et de selfs.
Les limites sont repoussées grâce à des technologies spécifiques : a) l'emploi de matériaux à coefficient de couplage piézoélectrique supérieur au quartz permet d'élargir les bandes passantes. Ainsi, la berlinite autorise des bandes passantes voisines de 1 % et le tantalate de lithium s'adresse à des bandes de quelques pourcents. b) la fréquence centrale des résonateurs peut être élevée au delà de 35 MHz en prolongeant l'amincissement mécanique des lames de résonateurs par un usinage ionique ou un usinage chimique. On parvient ainsi à commercialiser des filtres à fréquence supérieure à 200 MHz.

2/ Les filtres à onde de surface : ils sont adaptés au domaine de fréquence et de bande passante, mais présentent traditionnellement une perte d'insertion élevée. Les techniques proposées dans la littérature pour surmonter cet inconvénient sont désormais nombreuses.

La nécessité de répondre à des demandes de filtres accrues dans le domaine fréquence x bande passante considéré et en particulier les demandes de filtres FI pour le radiotéléphone numérique cellulaire paneuropéen nous ont conduit à développer une technologie faible perte pour nos filtres à onde de surface afin de permettre une comparaison équitable de performances avec les filtres à onde de volume.

Les caractéristiques objectives que nous nous sommes fixées sont les suivantes

- fréquence centrale 30 à 120 MHz
- bande passante $\sim$ 300 KHz
- variation du temps de groupe < 300 ns
- raideur équivalente : 4 pôles
- réjection : 60 dB
- perte d'insertion < 6 dB sur 50 ohms
- encombrement : minimal
- quantité : > 100.000 pièces par an

I/ Choix d'une technologie faible perte

Les technologies faibles pertes pour les filtres à onde de surface peuvent être classées en 3 catégories 1/ les transducteurs à réflecteurs répartis 2/ les filtres à transducteurs d'entrée et de sortie imbriqués 3/ les transducteurs triphasés.

Ces trois principes ont été discutés selon les critères suivants 1/ performances de pertes faibles démontrées 2/ coûts de fabrication réduits compatibles avec des séries de 100.000 filtres 3/ développement réduit pour permettre une réponse rapide au marché.

a) Les transducteurs à réflecteurs imbriqués

Ils se présentent sous deux formes : 1/ le transducteur est composé de blocs de doigts séparés par des espaces occupés par des réflecteurs partiels [1]. 2/ le transducteur est continu et composé de doigts dédoublés. Un des deux doigts jumeaux est surchargé par une métallisation lourde donnant à ce doigt un pouvoir réflecteur supérieur [2].

Dans les deux cas, l'effet unidirectionnel est obtenu par un décalage de $\lambda/8$ entre le barycentre des émissions et le barycentre des réflexions. Cette disposition donne en théorie la perte faible d'un résonateur associée à la largeur de bande du transducteur. L'avantage majeur de cette technologie est l'absence de réseau de déphasage pour alimenter le transducteur. Mais la technologie semble délicate : les deux niveaux de masquages doivent être réalisés

avec une grande précision de superposition et la performance du produit dépend de l'épaisseur de la surchage sur le réflecteur, difficile à maîtriser en pratique.

b) Les filtres à transducteurs d'entrée et de sortie imbriqués

Cette technique a été proposée par plusieurs auteurs ; la présentation la plus élaborée a été faite par Hitachi [3]. Cette technique semble très prometteuse tant par les performances annoncées que par la facilité de fabrication : pas de réseaux de déphasage et un seul niveau de masquage.

Cependant, les études de base ne nous ont pas semblé encore assez avancées pour permettre une mise en fabrication compatible avec le délai que nous prévoyons.

c) Les transducteurs triphasés

Ce principe est le plus traditionnel des principes faibles pertes. Il se présente sous deux formes : 1/ les doigts des transducteurs sont alternativement réliés à trois phases espacées de 120° ou de 90°. Dans ce cas, les pistes doivent se chevaucher ce qui impose une technologie à plusieurs niveaux de masquage [4]. 2/ les transducteurs sont formés d'une succession de blocs alimentés alternativement entre une phase et la masse puis entre l'autre phase et la masse [5]. Pour éviter les croisements de pistes, la piste de masse présente alors un tracé en forme de grecque. Il n'y a plus qu'un niveau de masquage, mais la périodicité des blocs apporte des réponses indésirables hors bande.

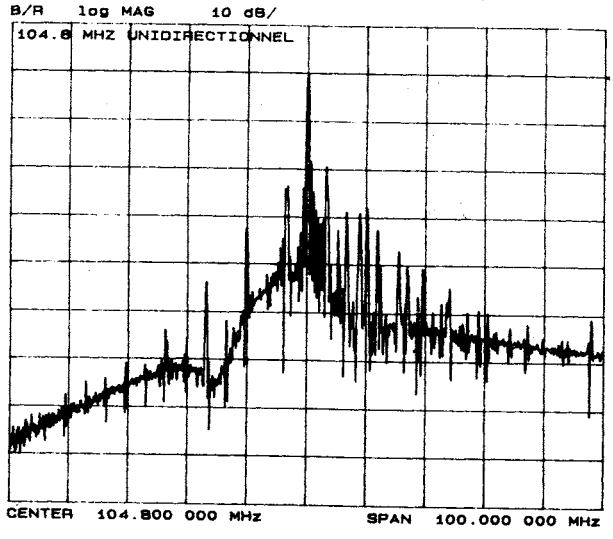
Les deux dispositions présentent en outre l'inconvénient de nécessiter des réseaux de déphasage. Cependant les performances en perte d'insertion sont bonnes, la synthèse est réalisable et les principes sont bien maîtrisés autorisant une mise en oeuvre industrielle rapide.

C'est donc cette dernière technologie que nous avons retenue.

Dans l'optique d'une minimisation des coûts de fabrication, la technologie à plusieurs niveaux de masquage nous est apparue comme une lourdeur. Nous avons donc opté pour la technologie des transducteurs à blocs.

II/ Optimisation du transducteur faibles pertes

Un premier filtre a été réalisé à 104,8 MHz. Chaque transducteur comportait 16 blocs de 32 doigts. L'alimentation en masse était réalisée par une grecque . Aucun transducteur n'est pondéré. Les réponses obtenues sont sur la courbe 1.Il est apparu -ce qui était prévisible- deux défauts majeurs vis à vis de l'objectif : 1/ la périodicité des blocs apporte de sérieuses réponses parasites hors bande. 2/ la grecque de masse, conducteur long et étroit était trop résistive ; il a fallu faire des rappels de masse par bondings en de nombreux points pour obtenir une bonne performance en perte d'insertion.



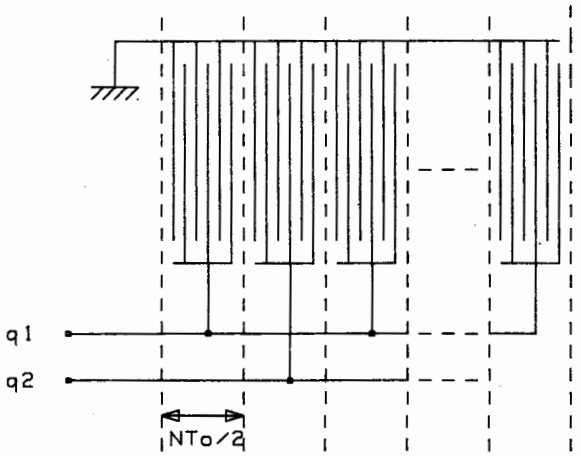
COURBE 1 : Filtre unidirectionnel à 104,8 MHz

Il est donc apparu nécessaire d'établir un certain nombre de règles de conception visant à minimiser ces inconvénients.

a) Analyse simplifiée du transducteur

Le transducteur à blocs peut être représenté dans le domaine temps par la combinaison de 3 fonctions : 1/ la fonction trame t(t) 2/ la fonction bloc b(t) 3/ la fonction enveloppe e(t)

La description du transducteur est la suivante : il comporte 2α blocs constitués chacun de N sources espacées de To/2 (fig.1).



bloc1

q1 et q2 sont les 2 alimentations électriques à phases décalées.

FIGURE 1 : Description du transducteur

La fonction de transfert du transducteur h(t) s'écrit alors :

$$h(t) = (t(t) * b(t)).e(t)$$

* est le symbole du produit de convolution.

Les trois fonctions sont décrites dans le tableau 1.

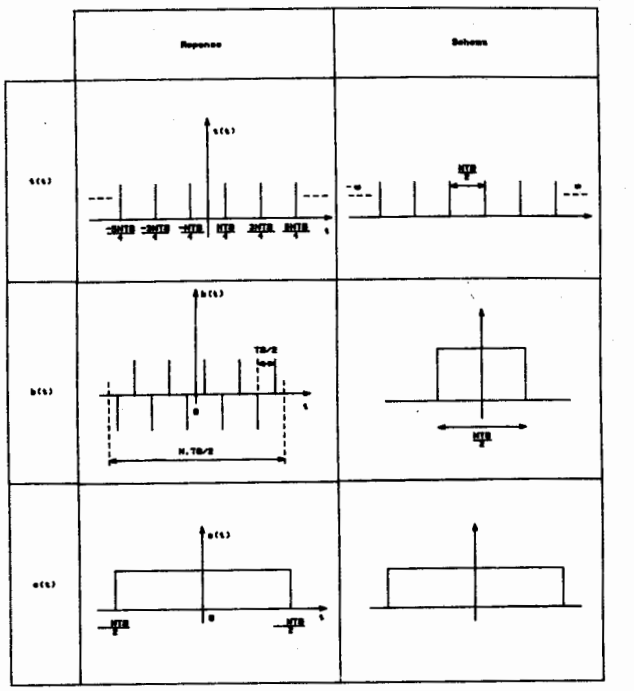


TABLEAU 1

Remarque :

- δ est le symbole de la fonction de Dirac
- e(t) dépend de la réponse en fréquence désirée. Dans le cas présent, nous supposons la réponse obtenue sans pondération

Ces fonctions se tranforment dans le domaine fréquence selon le tableau 2

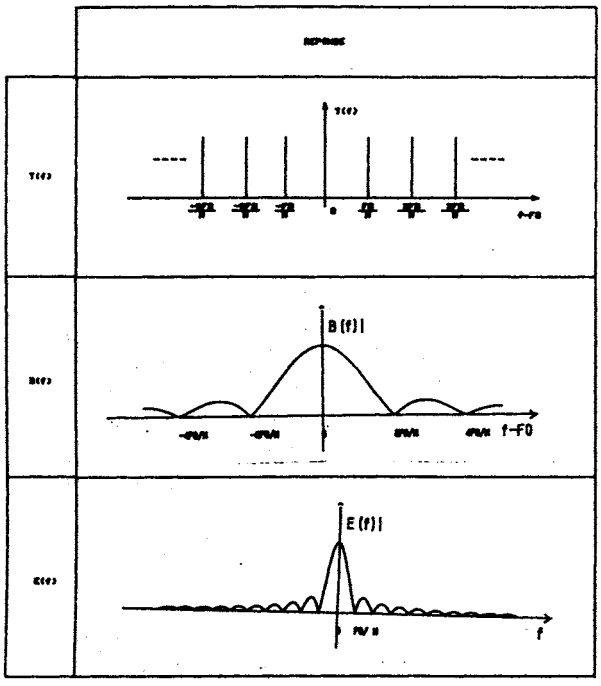


TABLEAU 2

La réponse globale en fréquence s'écrit

$$H(f) = (T(f).B(f)) * E(f)$$

La fonction T(f).B(f) dite fonction de structure, donne les différentes réponses de la structure. On convient d'appeler réponse parasite d'indice k les réponses pour lesquelles k # 0 et réponse principale celle pour k = 0.

On constate que la structure élimine les réponses parasites d'ordre pair (coïncidence d'un zéro de B(f) et d'un pic de T(f)) et ne concerne que les réponses parasites d'ordre impair d'amplitude décroissant en 1/k.

Par le biais de la convolution, la réponse d'enveloppe E(f) se manifeste sur chaque réponse de la structure selon la figure 2

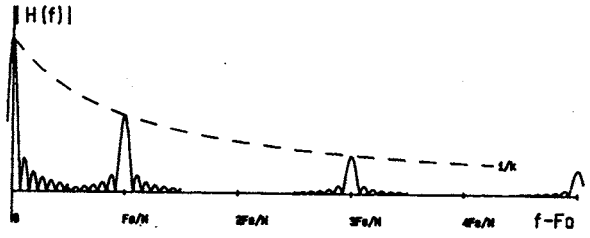


FIGURE 2

Cette allure de réponse est bien vérifiée sur la courbe réelle (courbe 1).

Les réponses supplémentaires obtenues pour f > Fo sont des réponses d'onde de volume facilement éliminables.

b) Optimisation de la dimension des blocs

La présence des réponses parasites étant due à la périodicité de la structure, une première hypothèse consiste à détruire cette périodicité. Des simulations ont été effectuées en donnant aux blocs des longueurs non plus constantes mais aléatoires : l'amélioration est effective mais insuffisante.

Dans la suite, les blocs sont à nouveau supposés de longueur constante. On examine l'influence de la longueur du bloc. Plus les blocs sont petits, plus la première réponse parasite est éloignée de la réponse principale. A priori, cette situation est favorable pour le filtre. Une analyse plus fine a montré le contraire, par suite de la prise en compte du nécessaire espacement entre les blocs.

Soit la structure s(t) définie par le schéma de la figure 3.

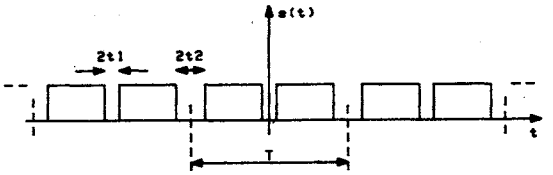


FIGURE 3

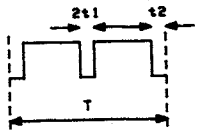
Cette structure comporte une succession infinie de blocs identiques espacés alternativement de $2t_1$ et $2t_2$.

$$s(t) = t'(t) * b'(t)$$

avec :

$$t'(t) = \sum_{k=-\infty}^{+\infty} \delta(t-kT)$$

$b'(t)$ =



La réponse de la structure s'écrit :

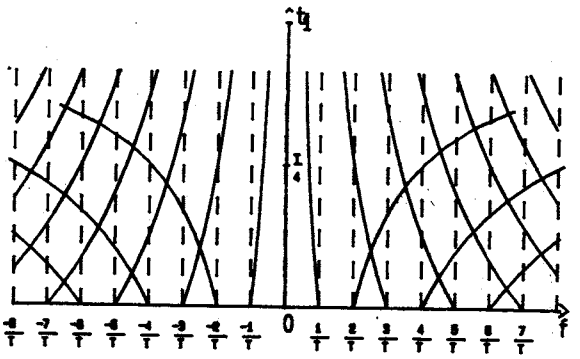
$$S(f) = T'(f)\, B'(f)$$

$$T'(f) = \sum_{k=-\infty}^{+\infty} \delta(f - 2\pi k/T)$$

$$B'(f) = \sin \pi f (t_1 + t_2 - T/2) \cos \pi f (T/2 + t_1 - t_2)$$

On recherche t_1 et t_2 pour lesquels un nombre maximum de pics de T'(f) se trouve annulé par des zéros de B'(f).

La figure 4 présente le lieu des zéros de B'(f) pour t_1 variant de 0 à sa valeur maximale possible T/2 (en posant $t_2 = 0$)



En trait plein: lieu des zeros de $E_1(f)$
En pointilles: lieu des pics de $E'_2(f)$

FIGURE 4

On constate que plus t_1 augmente, moins les annulations de pics de trame sont fréquentes. L'optimum est pour $t_1 = 0$.

Le tracé des courbes comparables pour une valeur non nulle de t_2 ne modifie ni l'allure générale des courbes, ni la conclusion.

L'optimum est donc obtenu pour t_1/T et t_2/T voisins de zéro.

En pratique, il faut donc que l'espacement entre les blocs soit nul, ou que les blocs soient très grands devant l'espacement.

c) Dessin du transducteur

Pour obtenir l'effet d'unidirectionnalité, il est nécessaire que le barycentre des blocs alimentés par la phase 1 et le barycentre des blocs alimentés par la phase 2 soient déphasés dans le temps d'un angle opposé à l'angle de déphasage des alimentations (en pratique 90°).

Le nécessaire décalage entre les blocs en résultant empêche par suite de considérations géométriques l'annulation de

t_1 et t_2. Il est toutefois possible de les minimiser : 1/ en supprimant la grecque de masse, dont les trajets verticaux entre les blocs sont causes d'accroissement de t_1 et t_2 2/ en adoptant une disposition des doigts qui minimise t_1 et t_2. Cette disposition est décrite en figure 5 :

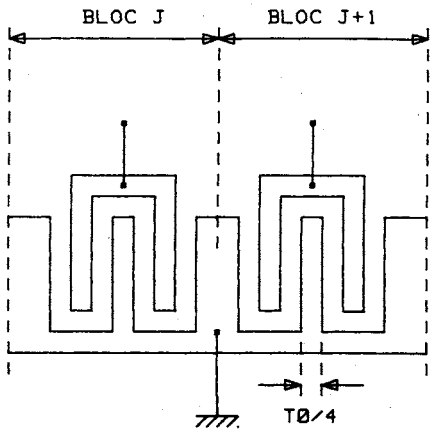


FIGURE 5

Elles est telle que $2t_2 = 3/4To$ et $2t_1 = 1/4\ To$

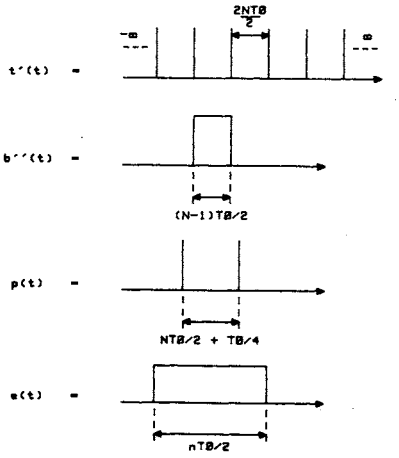
On a vérifié que $1/2\ (2t_2 - 2t_1) = To/4$ apporte bien le déphasage souhaité de 90°.

d) Analyse détaillée du transducteur

La règle de dessin définie pour le transducteur amène une nouvelle analyse du transducteur donnée par :

$$h(t) = (t'(t) * b''(t) * p(t)).e(t)$$

avec les schémas suivants :



où la fonction p(t) est la fonction de déphasage des blocs.

La réponse fréquentielle est alors :

$$H(f) = (T'(f).B''(f).P(f)) * E\ (f)$$

Elle est décrite par le haut du tableau 3

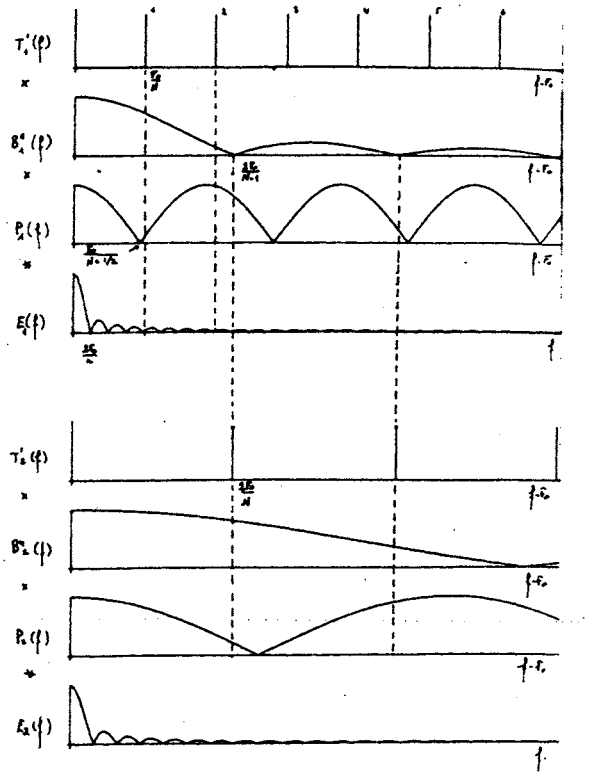


TABLEAU 3

On constate : 1/ que les réponses d'indices pairs de T'(f) ne sont plus exactement annulées par B"(f). Cependant, si N est grand (ie les blocs sont longs), les pics resteront dans des zones de fort affaiblissement de B"(f). Toutefois, on peut prévoir que le pic d'ordre 2, le moins atténué de tous, restera encore apparent. 2/ les zéros apportés par P(f) sont très rapprochés des pics impairs de T'(f).On peut prévoir qu'ils apporteront une atténuation pas parfaite mais tout de même très bonne des pics impairs de T'(f). Le pic 1 en particulier, le moins atténué de tous les impairs, sera certainement plus atténué que le pic 2

e) Analyse globale du filtre

Le filtre est composé de 2 transducteurs. Chaque transducteur a une réponse $H_1(f)$ et $H_2(f)$. Le transducteur 1 est non pondéré. Le transducteur 2 est pondéré. La réponse totale est :

$$H(f) = H_1(f) \cdot H_2(f)$$

On choisit la périodicité des blocs de $H_2(f)$ afin d'obtenir la situation décrite par l'ensemble du tableau 3.

On remarque que pour le transducteur 2, les blocs sont plus petits et donc le rapport espacement sur bloc est plus grand. L'annulation des pics pairs de $T_2'(f)$ par les zéros de $B_2"(f)$ est alors très médiocre. Pour remédier à ce défaut, les pics de $T_2'(f)$ ont été mis en correspondance avec les zéros de $H_1"(f)$. Ainsi tous les pics de $T_2'(f)$ sont parfaitement annulés.

Donc, sur l'ensemble du filtre le seul pic restant médiocrement annulé est le pic 2 de $T_1'(f)$ situé à $2Fo/N$

f) Optimisation de l'enveloppe

La première application de l'enveloppe est de fournir la réponse correspondant au gabarit de filtrage souhaité. La définition de l'enveloppe est alors obtenue par des méthodes de synthèses classiques.

Dans le cas du transducteur unidirectionnel à blocs, il est intéressant d'imposer une condition supplémentaire sur la bande atténuée visant à accentuer encore l'atténuation des pics parasites, notamment les pics 1 et 2 de $T_1'(f)$

Cette optimisation n'a abouti que pour le transducteur non pondéré :

$$H_1'(f) = [T_1'(f)\, B_1"(f)\, P_1(f)] * E_1(f)$$

$$H_1'(f) = \int T_1'(f')\, B_1"(f')\, P_1(f') * E_1(f' - f)\, df'$$

$$H_1'(f) = \int \sum_k (f' - Fok/N)\, B_1"(f')\, P_1(f')\, \frac{\sin \pi n/2(f'-f)/Fo}{\pi n/2(f'-f)/Fo}\, df'$$

$$H_1'(f) = \sum_k B_1"(Fok/N)\, P_1(Fok/N)\, \frac{\sin\ n/2\ (k/N - f/Fo)}{n/2\ (k/N - f/Fo)}$$

Afin d'annuler les pics de $T_1'(f)$, on impose $H_1'(f) =$ pour

$f = mFo/N \qquad m = 1,2,\ldots$

$$\sum_k B_1"(Fok/N)\, P_1(Fok/N)\, \frac{\sin \pi n/2\ (k - m)/N}{\pi n/2\ (k - m)/N} = 0$$

Les solutions $n = 2 \alpha N$ (α = entier) satisfait cette équation pour toutes les valeurs de m.

III/ Application

L'étude décrite a défini un dessin du transducteur et induit les règles suivantes :

$$N_1 = 4\beta + 3$$
$$N_2 = 2\beta + 1$$
$$n_1 = 2\alpha_1 N_1$$
$$n_2 = 2\alpha_2 N_2$$
$$\beta = \max$$

En outre :

est borné supérieurement par le fait que le nombre de blocs par transducteur doit être suffisant pour assuré l'unidirectionnalité. Ce nombre a été fixé arbitrairement à 4.

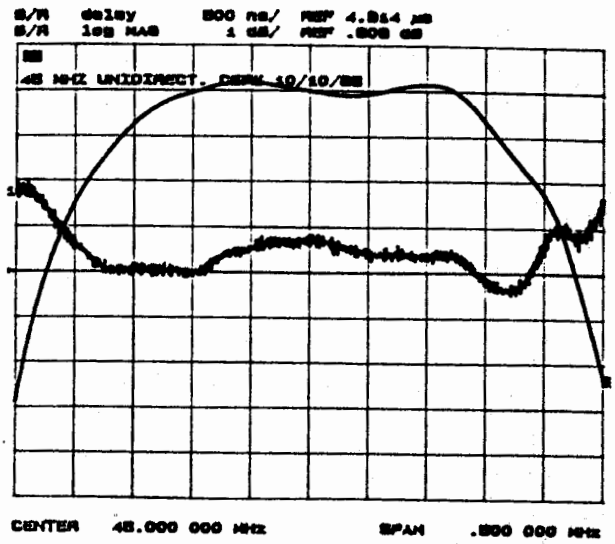
n1 et n2 sont définis à quelques unités près par la synthèse du filtre.

La définition finale de β a été obtenue grâce à une optimisation par simulation.

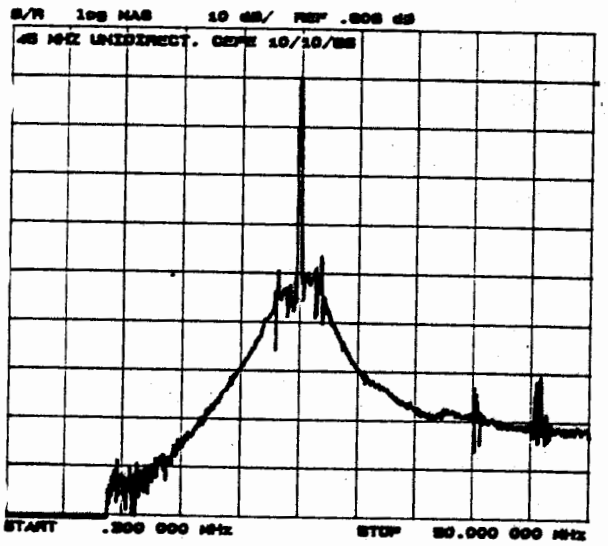
Les valeurs les plus satisfaisantes ont été :

$$\beta = 6 \quad \alpha_1 = 4 \quad \alpha_2 = 27$$

La réponse du filtre réalisé est donné par les courbes 2 et 3.



COURBE 2



COURBE 3

On vérifie bien une bonne réjection (de l'ordre de 40 dB) de toutes les raies parasites. Celle qui reste malgré tout la plus marquée est la raie 2 de $T_1'(f)$ située à $2Fo/N_1$ à 3,5 MHz.

C'est ce que prévoyait le raisonnement.

IV/ Conclusion

L'étude menée a eu pour but de rechercher les performances maximales pour des filtres à onde de surface faibles pertes applicables au filtrage FI dans le radiotéléphone numérique pan- européen.

Le choix du design a été orienté vers de technologies rapides à mettre en oeuvre et minimales en nombre d'opérations de fabrication afin d'être compatible en délai et en coût avec l'objectif visé.

En fonction de ces critères, la disposition à transducteur triphasé et un seul niveau de masquage a été retenue.

Les travaux menés ont montré qu'une optimisation de cette technologie permet de réduire les réponses parasites associées à cette disposition et d'obtenir des niveaux de réjection supérieurs à 40 dB.

L'optimisation de la fonction d'enveloppe du transducteur non pondéré et l'étude de blocs fractionnés permettraient d'améliorer encore ce résultat dans l'avenir.

BIBLIOGRAPHIE

[1] Low loss SAW device employing single stage fabrication
LEWIS
1983 Ultrasonics Symposium Proc. p 104-108

[2] Low loss SAW filter using internal reflection types of new single phase unidirectionnel transducer
YAMANOUCHI and Al
1984 Ultrasonics Symposium Proc. p 68-71

[3] High performance SAW filters with several new technologies for cellular radio
HIKITA and Al
1984 Ultrasonics Symposium Proc. p 82-92

[4] Low loss SAW band pass filter
DEFRANOULD
1981 Ultrasonics Symposium Proc. p 17-20

[5] Low insertion loss acoustic surface wawe filter using group type unidirectionnel interdigital transducers
YAMANOUCHI
1975 Ultrasonics Symposium Proc. p 317-321

SAW FILTERS FOR DIGITAL CELLULAR RADIO APPLICATIONS *

J.M.HODE, P.DUFILIE, J.DESBOIS, H.GAUTIER

THOMSON-SINTRA ASM / D.T.A.S
399, Route des Crêtes - Parc de Sophia-Antipolis
06561 VALBONNE CEDEX FRANCE

ABSTRACT

Two Surface Acoustic Wave (SAW) filters have been developed for both the front end and the IF sections of the future European digital cellular mobile telephone system.

The first one operates at 71 MHz with 200 kHz bandwidth; it exhibits insertion loss as low as 2.6 dB, 45 dB close-in rejection and 70 dB far-out rejection. This filter provides both channel filtering and blocking signal rejection.

The second one operates in the transmit or receive band (about 900 MHz). Very high performances were achieved : -1 dB bandwidth greater than 25 MHz with less than 0.5 dB and 20 ns peak to peak in-band ripple, close-in and far-out rejection greater than 50 dB, shape factor better than 3:1 and insertion loss as low as 7 dB.

Theoretical predictions for both filters and other results are also presented, showing the capability of the SAW technology to achieve high performance low loss RF and IF filters.

INTRODUCTION

Great efforts are being made all over Europe to develop a Pan-European mobile communication system. Many countries are involved in this ambitious project, such as the Scandinavian countries, Germany, U.K, Italy and France.

The architecture of this future cellular mobile radio system is rather different from the present 400 MHz and 900 MHz all-analog systems. Digital processing provides better S/N ratio (about 13 dB) and the TDMA technique allows for time separation of transmitted and received signals with the following consequences :

* suppression of the duplexer which was one of the largest circuits in the system

* the RF filter specifications can be relaxed

However such a complex mobile radio system requires various high performance technologies regarding both active and passive elements.

CIRCUIT ARCHITECTURE

A possible architecture of the RF and IF sections of both mobile and base station systems could be represented by the block diagram presented in fig. 1. The transmit and receive bands for the mobile are respectively 890 - 915 MHz and 935 - 960 MHz. Four RF filters can be identified, two in each section (transmission and reception). Insertion loss must be less than 2.5 dB for the two receive filters, as low as 1.5 dB for the transmit front end filter, and less than 8 dB for the transmit preamplifier filter.

For technological reasons the single IF scheme is the most intensively investigated, with a probable IF close to 71 MHz. The IF section processes a 270 kBit TDMA channel (containing 8 time-multiplexed voice channels) with a 90 dB sidelobe GMSK modulation. The corresponding bandwidth is 200 kHz allowing 1000 voice channels to be allocated.

The aim of the IF filter is to protect the IF section against both blocking signals and adjacent TDMA channels. Since channels $n\pm1$ and $n\pm2$ are not used while channel n is used, the ultimate rejection requirement is specified from IF $\pm$ 800 kHz (channel $n\pm3$) to IF $\pm$ 25 MHz (limit of the receiving band). Low insertion loss is also required for both power consumption, circuit size and cost considerations.

Surface Acoustic Wave (SAW) filters are known for miniaturization, mass production capability and high reproducibility. These points make the SAW technology a very attractive solution for both RF and IF filtering in mobile telephone systems.

This paper presents a 71 MHz and a 900 MHz SAW filter which were developed for digital radio applications.

LOW LOSS TECHNIQUES FOR SAW TECHNOLOGY

Generally conventional SAW filters show fairly large insertion loss (IL), typically 20 dB, which restrict their applications. Front end or first IF stage applications require lower losses, below 10 dB and even 5 dB.

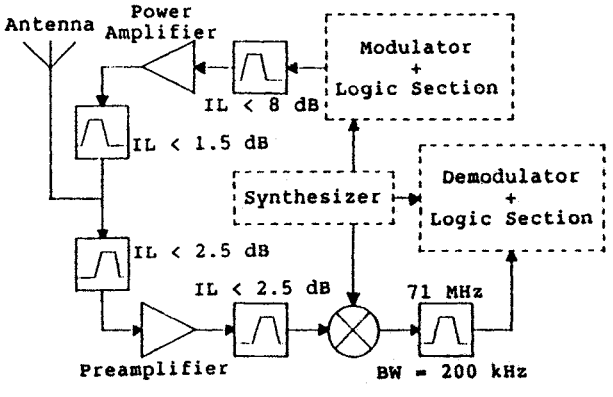


FIGURE 1 : Possible block diagram for Pan-European mobile communication system.
Receive band : 935 - 960 MHz
Transmit band : 890 - 915 MHz

* This work was partially sponsored by D.A.I.I (France)

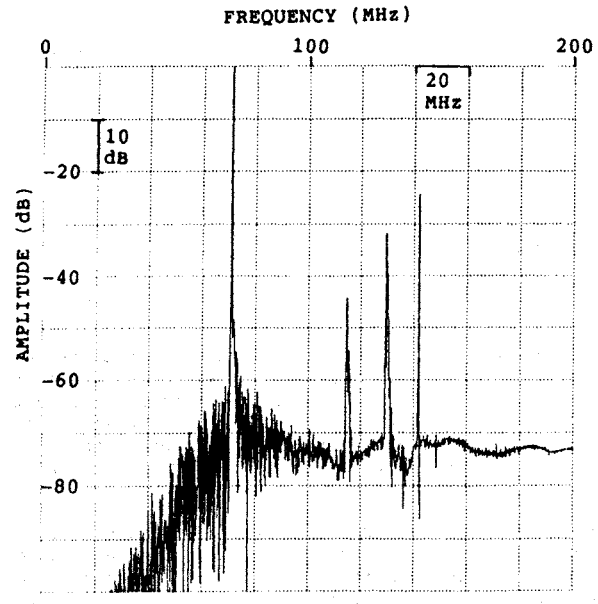


FIGURE 2 : Measured frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 100 MHz. Plot scales are 20 MHz/div horizontal and 10 dB/div vertical.

Numerous low loss filter structures have been developed during the past 15 years, such as 3-phase unidirectional transducers (3PUDT) [1,2], single phase unidirectional transducers (SPUDT), including both internal reflection transducers [3] and folded multistrip couplers transducers [4]), repetition structures [5,6] or track coupling via image impedance transducers [6-8].

For lower frequency applications the 3-phase technique is the most accurate for controlling in-band shaping and ripple. However air gap cross-over technology and minimum linewidth restrict the use of this technique to below 300 MHz.

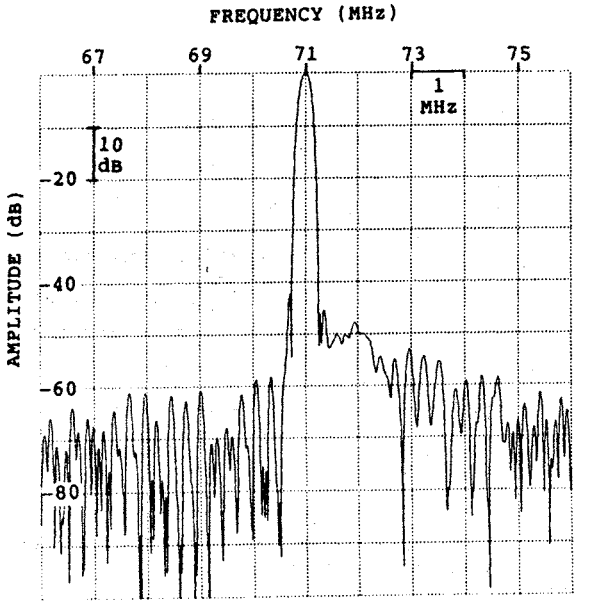


FIGURE 3 : Measured frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 1 MHz/div horizontal and 10 dB/div vertical.

High frequency applications require simple tuning, single phase and single layer technologies. These requirements are achievable at 1 GHz with a standard 1 µm lithography process by using a filter structure [8] derived from the repetition structure proposed by Hikita et al [6,7].

71 MHz IF FILTER

The center frequency of the IF filter was chosen to be 71 MHz. This choice represents a good compromise between the system requirements and the possibilities of the SAW technology.

Because of the narrow bandwidth requirement, quartz was chosen as the substrate material because of its high thermal stability ($3.5\ 10^{-2}$ ppm/°C²) and its low coupling efficiency ($k^2 = 0.116$ %).

The filter was designed and built using the 3-phase technology for easy in-band shaping and ripple control and for possibility of spurious signals compensation. In order to minimize insertion loss and diffraction distortion, withdrawal weighting was used.

The minimum cavity size required for this filter is 4.5 mm by 4 cm with a height of only 2 mm.

Figs. 2 to 4 give the measured frequency response of the filter. The insertion loss is only 2.7 dB, the close-in rejection is better than 45 dB and the far-out sidelobe level exceeds 70 dB from 0 to 200 MHz. Spurious bands occur at 115, 130 and 142 MHz with levels greater than 40, 30 and 20 dB from the main signal corresponding to SSBW (Surface Skimming Bulk Wave) radiation at 1.6 and 1.8 times f_0 and the 2nd harmonic of the SAW.
Figs. 3 to 4 can be compared with the analysis given in figs. 5 to 6. We can notice that, the agreement between experimental measurements and theoretical calculations is very good for both in-band and stop-band responses.

Group delay variation in the passband of the filter is parabolic in shape and is the result of the tuning. Such a shape is predicted by our simulations and can be approximated by a very simple relation, which is valid for many 3-phase filters :

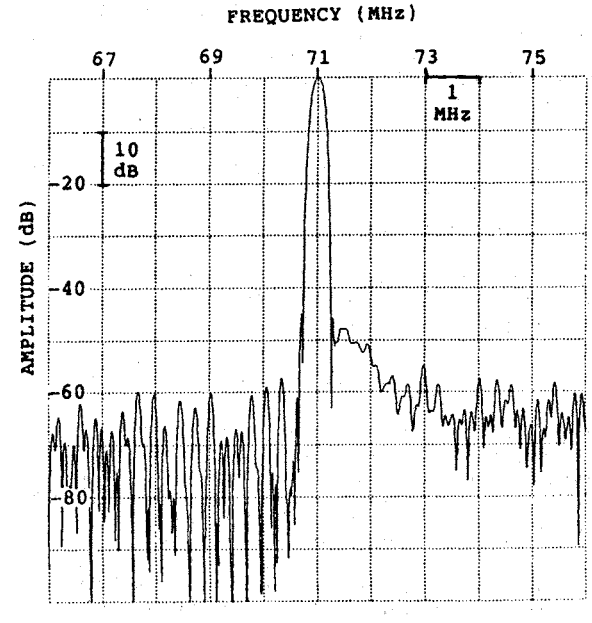


FIGURE 5 : Theoretical frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 1 MHz/div horizontal and 10 dB/div vertical.

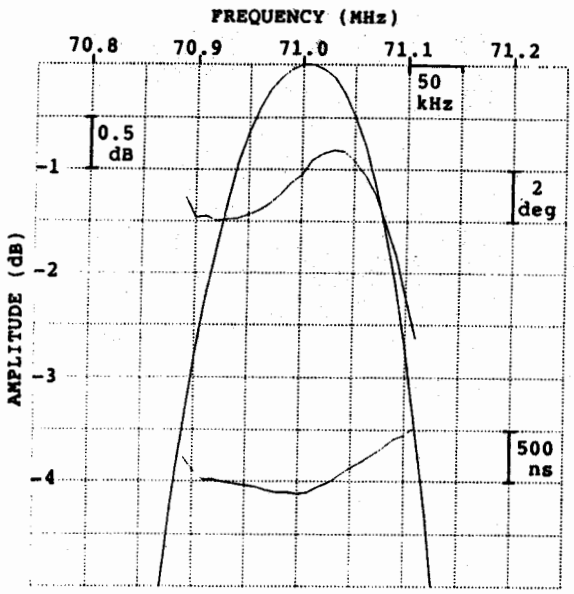


FIGURE 4 : Measured frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 50 kHz/div horizontal and 0.5 dB/div, 2 deg/div and 500 ns/div vertical.

$$\Delta\tau * B_{-3} \approx \frac{1}{10}\left(\frac{f - f_0}{0.5\, B_{-3}}\right)^2$$

This variation can be compensated for and reduced to ± 100 ns in the 3 dB band as shown in the simulation of fig. 7.

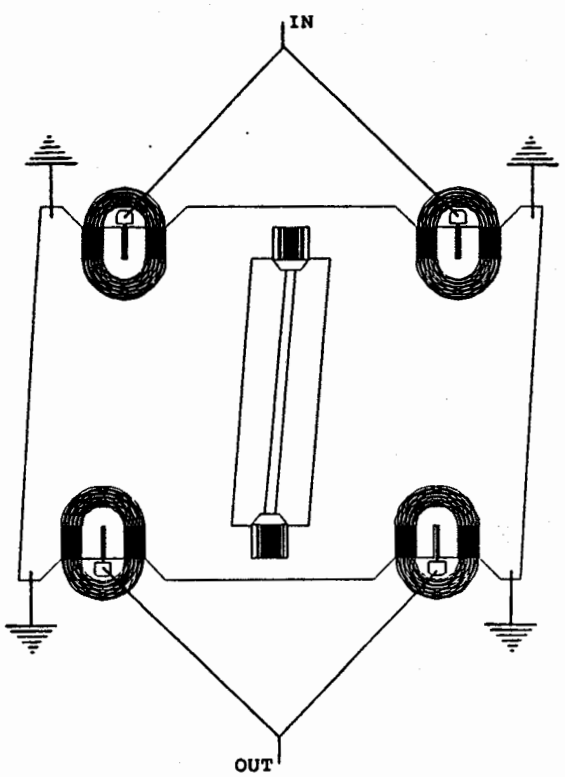


FIGURE 8 : 900 MHz RF filter configuration.

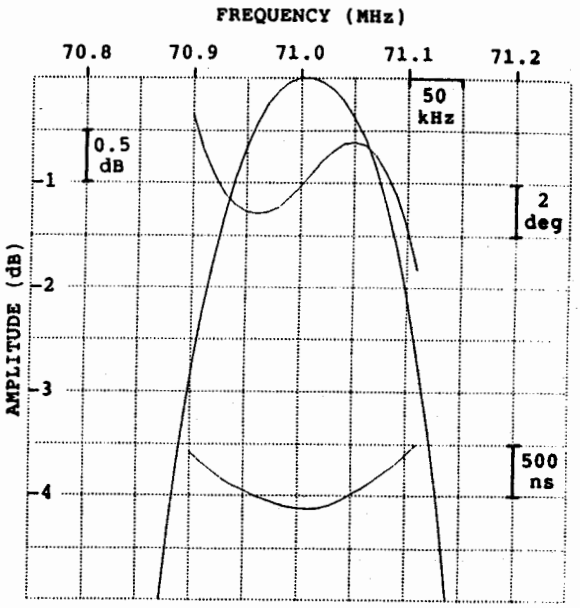


FIGURE 6 : Theoretical frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 50 kHz/div horizontal and 0.5 dB/div, 2 deg/div and 500 ns/div vertical.

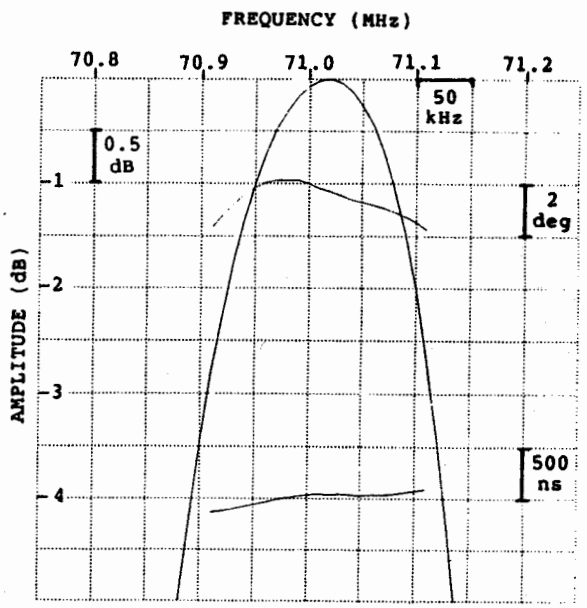


FIGURE 7 : Theoretical frequency response of the 71 MHz IF filter after phase error compensation. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 50 kHz/div horizontal and 0.5 dB/div, 2 deg/div and 500 ns/div vertical.

902 MHz RF FILTER

We have already reported [8] an improved structure for low loss, sharp cutoff, high rejection filters which can operate up to 1.5 GHz with standard optical processing techniques. This structure, derived from the repetition structure proposed by Hikita et al [6,7], combines several low loss techniques to achieve a single phase, single layer and simple tuning low loss filter.

The basic configuration for this structure is shown in fig. 8. It consists of a pair of unidirectional

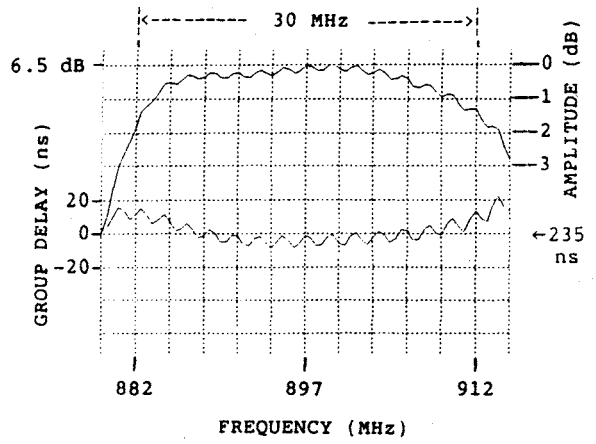


FIGURE 9 : Theoretical frequency response of the 900 MHz RF filter. Amplitude and group delay reference are 6.5 dB and 235 ns. Plot center at 897 MHz. Plot scales are 3 MHz/div horizontal and 1 dB/div and 20 ns/div vertical.

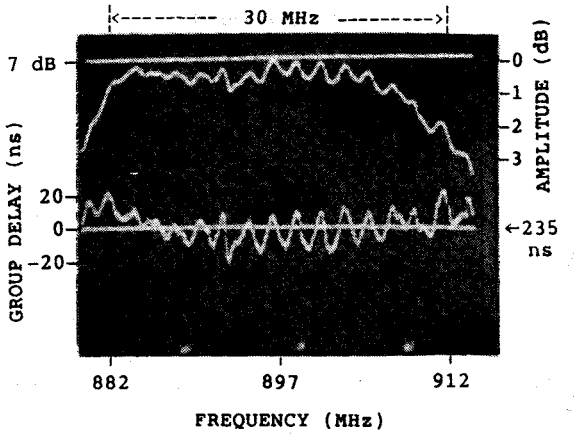


FIGURE 11 : Measured frequency response of the 900 MHz RF filter. Amplitude and group delay reference are 7 dB and 235 ns. Plot center at 897 MHz. Plot scales are 3 MHz/div horizontal and 1 dB/div and 20 ns/div vertical.

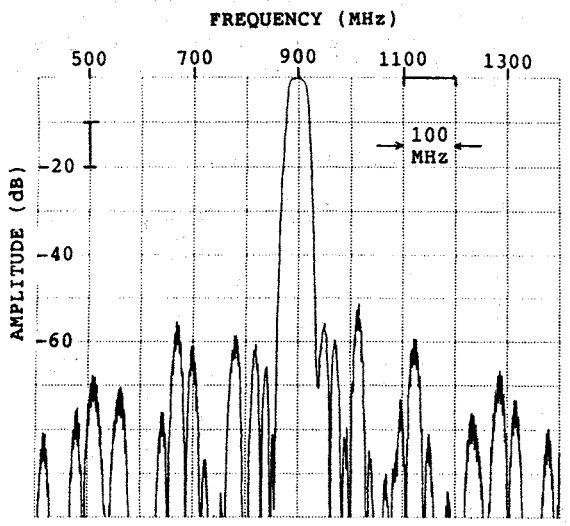


FIGURE 10 : Theoretical frequency response of the 900 MHz RF filter. Amplitude reference is 6.5 dB. Plot center at 900 MHz. Plot scales are 10 dB/div vertical and 100 MHz/div horizontal.

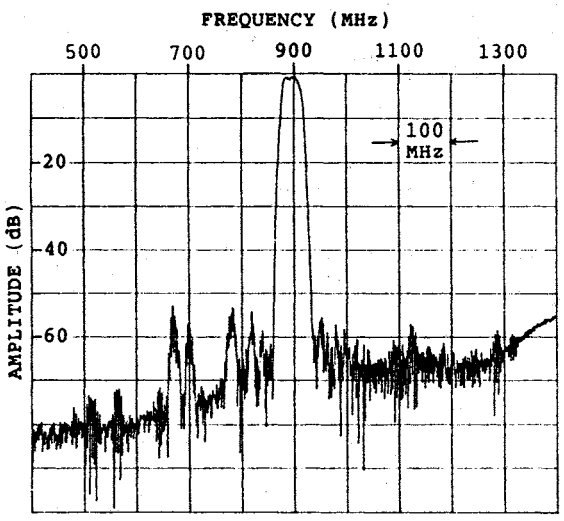


FIGURE 12 : Measured frequency response of the 900 MHz RF filter. Amplitude reference is 7 dB. Plot center at 900 MHz. Plot scales are 100 MHz/div horizontal and 10 dB/div vertical.

transducers (UDT) for both electrical input and output. These transducers are comprised of a bidirectional transducer inside a 3 dB multistrip coupler (MSC) which is folded in the form of a 0 (OMSC-UDT structure). Input and output are acoustically coupled via a pair of electrically coupled bidirectional transducers (TCs). Lossless coupling and passband shaping are obtained by a proper image impedance connection of the TCs as explained in reference [6]. Standard withdrawal weighting on the TCs is sufficient to obtain 50 dB rejection with the input/output transducer responses included as shown on figs. 9 and 10.

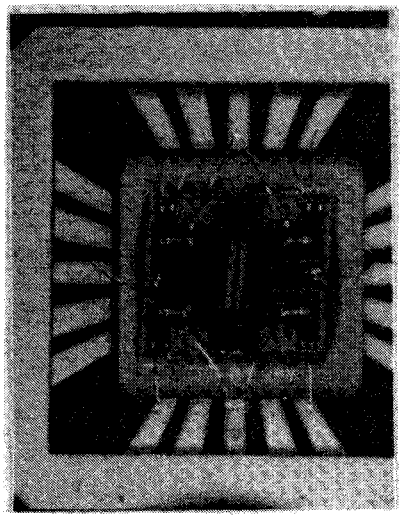
The filter was manufactured on 36° YX-$LiTaO_3$ because of its low temperature coefficient (35 ppm/°C) and high coupling coefficient (> 5.6 %). Figs. 11 and 12 show the measurement of its frequency response which appears to be in good agreement with the theoretical predictions shown in figs. 9 and 10.

High performances are obtained since the achieved insertion loss is 7 dB with a 0.5 dB and 20 ns fast peak to peak amplitude and group delay in-band ripple. Both close-in and far-out rejection are greater than 50 dB over a 1 GHz bandwidth for a shape factor better than 3:1.

Chip size is only 2.2 x 2.2 mm^2 which allows mounting in a TO-5 can. Flat pack packaging affords improved electromagnetic feedthrough rejection and chip carrier mounting has also been investigated (fig. 13) for surface mount applications.

OTHERS LOW LOSS RESULTS

The "image impedance" technique described above and explained in reference [6], allows an achievable fractional bandwidths close to 0.67 times k^2. Using an improved transducer design technique we can vary both effective coupling and stopband width in order to extend the domain of achievable relative bandwidths on a given material.



4 mm

FIGURE 13 : Photograph of the 900 MHz RF filter mounted in a 9 x 9 mm² chip carrier.

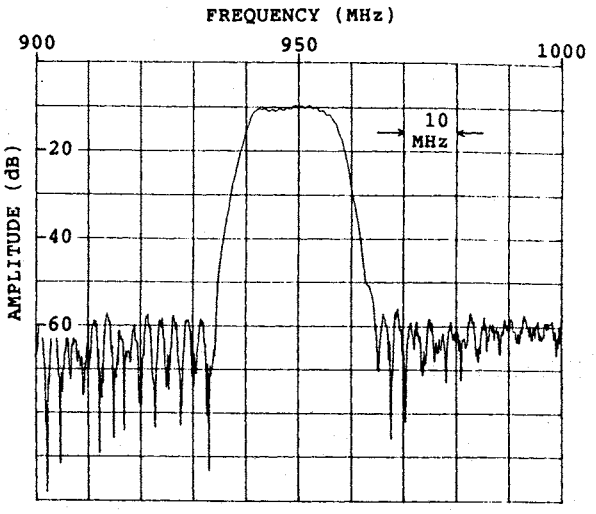


FIGURE 14 : Measured frequency response of the 950 MHz narrow band filter. Amplitude reference is 0 dB. Plot center at 950 MHz. Plot scales are 10 dB/div vertical and 10 MHz/div horizontal.

With such a technique we were able to design a narrow band filter having a relative 3 dB bandwidth of 1.7 %, operating at 950 MHz. Fig.14 shows the measured frequency response of this filter. The in-band and stop-band characteristics are very similar to those of the 900 MHz filter, except for the 10 dB isertion loss caused by a strong diffraction and propagation attenuation.

An improved weighting technique described in ref. [6] was used to design a 2 MHz bandwidth filter operating at 72 MHz with only 3 dB insertion loss (fig. 15 and 16). Proper design of the input/output transducers combined to the weighting technique results in both 70 dB far-out rejection (50 dB close-in with a shape factor of 3:1) and simple tuning (one inductor on each side of the filter). In addition the chip size is only 8.5 x 8.5 mm².

CONCLUSION

Two high performance low loss SAW filters operating at 71 MHz (IF) and 900 MHz (RF) have been developed for the future Pan-European digital cellular radio system.

Theoretical predictions for both filters are in perfect agreement with the experimental measurements.

The main specifications of the mobile radio system are satisfied. The results presented in this paper show the capability of the SAW technology to achieve high performance low loss RF and IF filters.

Future efforts are toward even lower losses and higher frequencies (up to 1.5 GHz).

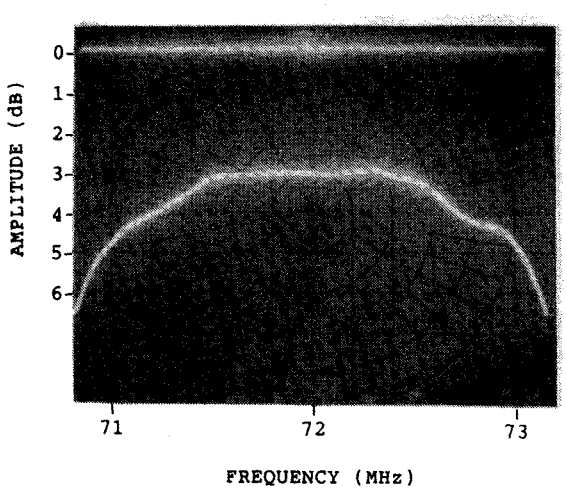


FIGURE 15 : Measured frequency response of the 72 MHz filter. Amplitude reference is 0 dB. Plot center at 72 MHz. Plot scales are 200 kHz/div horizontal and 1 dB/div vertical.

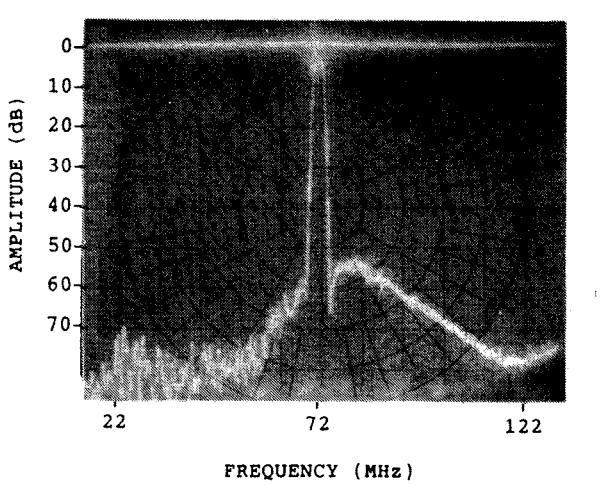


FIGURE 16 : Measured frequency response of the 72 MHz filter. Amplitude reference is 0 dB. Plot center at 72 MHz. Plot scales are 10 MHz/div horizontal and 10 dB/div vertical.

REFERENCES

[1] R.C ROSENFELD et al - "Low loss Unidirectional Acoustic Surface Wave Filters" Proceedings of 28th Annual Symposium on Frequency Control, New Jersey 1974

[2] Ph. DEFRANOULD, J. DESBOIS - "Low Loss SAW Bandpass Filters" 1981 IEEE Ultrasonics Symp. Proc. p. 17

[3] M. LEWIS - "Low Loss SAW Devices Employing Single Stage Fabrication" 1983 IEEE Ultrasonics Symp. Proc. p. 104

[4] F.G. MARSHALL et al - "New Unidirectional Transducer and Broadband Reflector of Acoustic Surface Waves" Elect. Lett., Vol. 7, p. 638

[5] M.LEWIS - "SAW Filter Employing Interdigitated Interdigital Transducers, IIDT" 1982 IEEE Ultrasonics Symp. Proc. p. 12

[6] M. HIKITA et al - " High Performance SAW Filters with Several New Technologies for Cellular Radio" 1984 IEEE Ultrasonics Symp. Proc. p. 82

[7] M. HIKITA et al - "New Low Loss Broadband SAW Filter using Unidirectional IDT's with U-Shaped MSCs" Elect. Lett., Vol. 20, p. 453

[8] P. DUFILIE, J.M. HODE, J. DESBOIS "A Low Loss High Performance Filter Structure" 1988 IEEE Ultrasonics Symp. Proc. (to be published)

A STRATEGY FOR A LOW COST IF FILTER FOR DIGITAL RADIOTELEPHONE

JP AUBRY, JP MICHEL, JL LE CORRE, D BLONDE & P de la FOURNIERE

C.E.P.E. 44 avenue de la Glacière 95100 ARGENTEUIL

SUMMARY

The next generation of digital radiotelephone which is expected to cover the entire EUROPE asks for many up to date technologies.

An intermediate IF filter will be required with the following basic specification :

- high center frequency (between 40 and 75 MHz)
- wide band (from 160 to 300 KHz)
- phase linearity < 5°
- low insertion loss (<6 dB)
- high out of band rejection
- low shape factor
- low ripple
- small size

The competition between SAW and BAW devices is still opened, even if SAW will have some inherent problems with rejection, size and cost.

BAW will be able to compete, but they have to overcome some limitation :

- high frequency fundamental
- piezoelectric substrate not in quartz
- small size
- low cost

The proposed strategy is based on four major points

- four poles on a single blank
- berlinite or lithium tantalate piezoelectric substrate
- automatique global filter adjustement
- low profile package

Two preliminary 4 poles designs have been fabricated and tested either with discrete or coupled resonators on a single wafer. Results meeting the specification are presented.

This is probably the first time that 4 poles single wafer are shown on $LiTaO_3$ or $AlPO_4$, and this is a very promising technology.

CLASSICAL TECHNOLOGY

Even if cost and size considerations are neglected, the IF filter for cellular digital radiotelephone is not an easy filter for most of crystal filters manufacturers.

The reasons are requirements for 1/ a wide relative bandwith 2/ a high center frequency 3/ a high out of band rejection 4/ a very good phase linearity.

1/ A wide relative bandwidth : Depending on the various system specifications, the required relative bandwidth is 0.25 to 0.7 %. This range of bandwidth is theoretically feasible with quartz resonators. But, practically, it will be a complicated structure involving such techniques as Jaumann design with two or more resonators per arm. This design needs complex transformers, is very hard to tune and has a poor rejection for a given number of resonators. In addition, the iterative impedance of the filter remains at a high level and matching the output of the IF mixer requires high quality factor and consequently large and costly inductors.

For wide bandwidth filters (above 0.3%) lithium tantalate is a well fitted piezoelectric crystal. Its high piezoelectric coupling coefficient makes possible a decrease of the iterative impedance compared to quartz (approximative ratio : 60).

So the required shape of the filter can be met with a very simple design like scales with serial only resonators. So it is a very promising material for the present filter. Nevertheless, for the lower bound of the bandwidth range there will be some limitations due to a slightly high insertion loss (~ 7 dB) caused by the limited Q factor of that material.

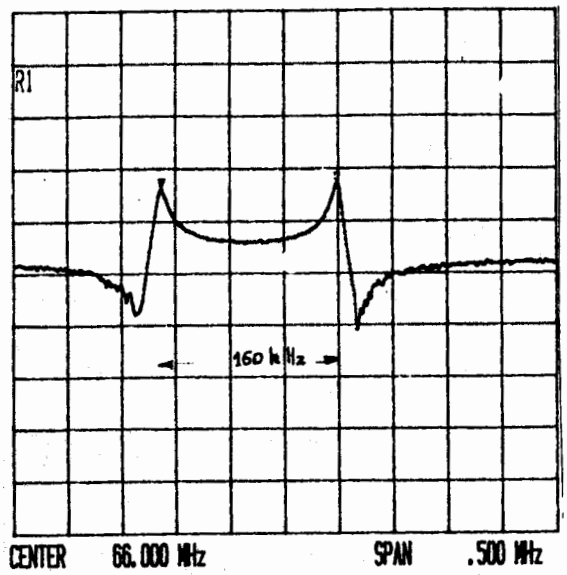


FIG. 1 : a 71 MHz - 100 KHz wide berlinite monolithic cell response

In order to cover the lower bound of the bandwith range we decided to investigate berlinite [1,2]. Berlinite is an aluminum phosphate with the same crystallographic structure as quartz (ie 32). But the piezoelectric coupling coefficient is approximately twice as compared to quartz and consequently achievable iterative impedances for filters are only half that of quartz. In addition monolithic cells up to 75 MHz proved to be feasible in the part (Fig.1). This makes possible a monolithic cells filter which is a very simple design (fig. 2).

So our choice has been lithium tantalate for the upper part of the bandwith range and berlinite for the lower part.

2/ A high center frequency : Either 40 MHz or 70 MHz are too high in frequency to make crystals by conventional mechanical etching. The current state of the art at CEPE would let them made by ionic etching. But we thought it is not convenient for mass production. We will see that there are other solutions.

3/ A high out of band rejection : The IF filter shall reject at a - 60 dB level. The shape factor required corresponds approximately to a four poles filter. But crystal resonators generally exhibit important spurious responses which compromise the high level of rejection. The classical mean to prevent this inconvenience is to increase the number of resonators and to achieve a higher order filter

4/ A good phase linearity : To get a good phase linearity is not properly speaking a great difficulty for crystal filters manufacturers. It acts more on the sensibility of the filter and it needs a more accurate trimming.

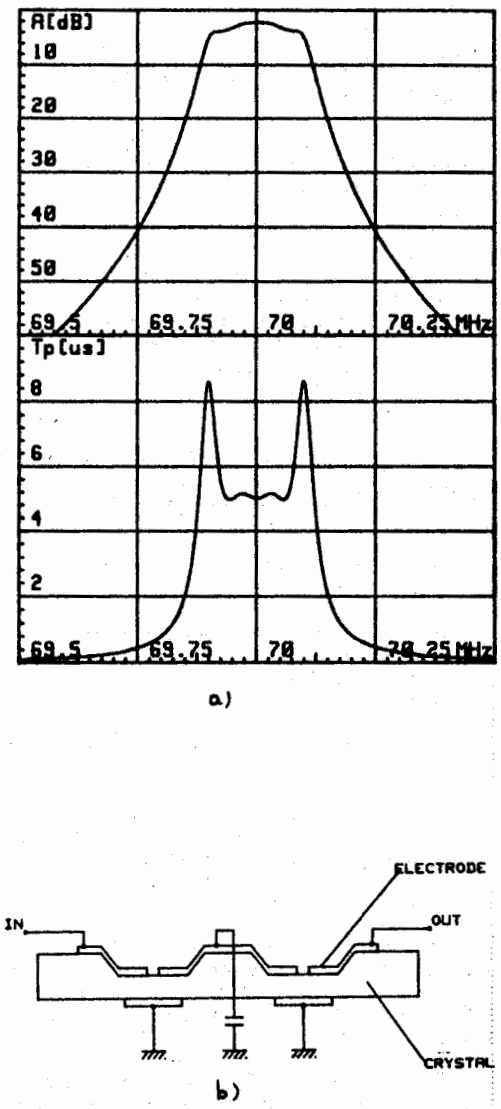


FIG. 2 : 4 poles berlinite filter at 70 MHz
a/ simulated response
b/ design

DEDICATED TECHNOLOGY FOR IF FILTER FOR DIGITAL RADIOTELEPHONE

The IF filter for radiotelephone will have to overcome the above difficulties. In addition it will need to be low cost and low volume. In order to achieve this, we made the following design choices :

1/ four poles design
2/ all the resonators on a single blank
3/ a maximally low iterative impedance
4/ a low profile package

1/ Four poles : As above stated, 6 poles would be better for spurii rejection. But the manufacturing costs for a filter are always more or less increasing with an increasing number of poles. Especially the curve tuning cost versus number of poles gets a strong upward bend around five poles. So it appeared necessary to limit at four (may be five) the number of poles. Fitting the required shape with only four poles however requires in certain cases a slightly bridged structure in order to produce attenuation maximum for a convenient closed rejection (fig.3).

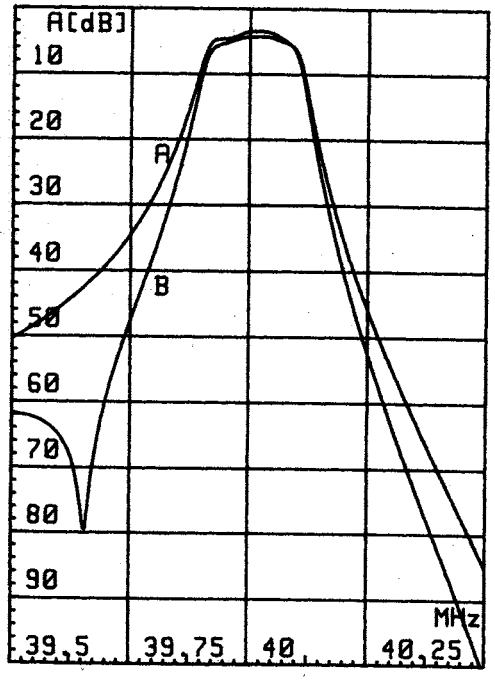


FIG.3 : Normal versus brigded filter design

When making filters this way with no special care as regards to spurious suppression, we got - what was expected - strong enough spurious responses whose rejection were only 30 dB.(fig. 4)

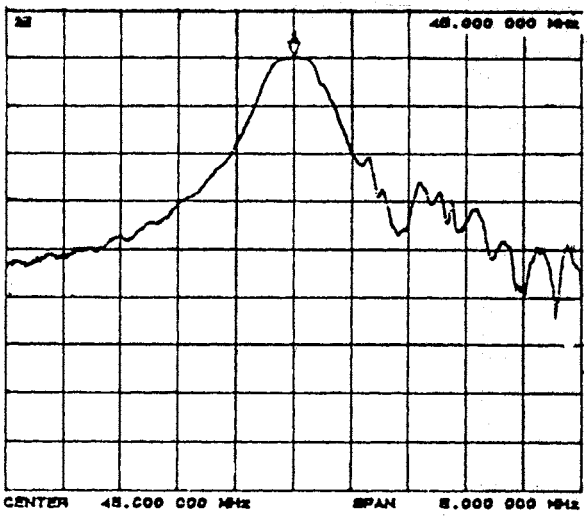


FIG. 4 : LiTa03 filter without spurii suppressor

To overcome this inconvenience we have studied resonators with special electrodes shape. On the resonator, the different mechanical modes of vibration are responsible for the main reponse and for the spurious ones too. One of them exhibits a strong coupling coefficient : it is the main and the useful one. Another specificity is the displacement figure caused by each different mode in the blank. These figures exhibit nodes and maxima in well defined places of the blank. A common method to avoid spurious is to place acoustical absorbant on the maxima of spurious modes provided they are not at the

same place as a main mode maximum. Results are good enough for low piezoelectric coupling materials such as quartz or berlinite. But even for these materials, and for materials with higher coupling coefficients, it can be improved by replacing the acoustical absorbant by an appropriate metallic load. The absorbing effect of spurious is got from mass loading added to a piezoelectric effect, especially strong for high coupling coefficient materials. Good results have been obtained with rejection up to 60 dB on spurii(see fig.10). Ease of implementation is also very good, because the metallic loads are deposited in the same operation as main electrodes.

2/ All the resonators on a single blank : Crystal costs are the main contribution of the cost of a crystal filter. So it is a great deal to reduce the cost of a resonator.

Multipoles resonators have already been shown [3].

The flow chart for a resonator's manufacture is given on figure 5. We made the following analysis : there are only two operations needing an individual action :

Materials

Sawing

Etching

Mounting

Tuning

Enclose

Final check

FIG. 5 : Flow chart for a crystal resonator

1- **Tuning** : as regard to achievable manufacturing process, one cannot hope that the resonator may reach the required precision without any individual tuning.

2- **Control** : each resonator is individually measured.

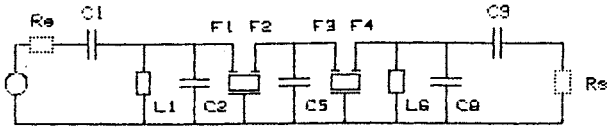
All other operations may find a solution in a collective or semi-collective production.

So the best should be done to individually process only for above both operations and to avoid individual manufacturing for all of the others. One solution for that last point was to study a device with all of the resonators on a same plate.

This disposal actually allows a collective processing on most of the manufacturing operations. In addition, it is cost saving on materials : obviously on mechanics, but also on crystal : the reduced needed quantity of crystal material is a very interesting point for lithium tantalate resonators, because this material is very dependant of the tantalum oxyde cost.

For lithium tantalate filters the design we have studied is a design with four resonators on a same plate. The diameter of the plate is approximatly 5 mm. The resonators are placed at a given distance of each other in order to avoid any unwanted coupling effects on the main response as well as on spurii. Results are very promizing.

For berlinite, we are studying a design based on a double monolithic cell. The frequency of the filter is > 60 MHz, the thickness of the plate is < 20 um. The two cells are placed in a depression made by ion etching in order to achieve the required thickness (see figure 6).



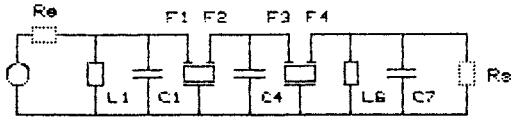


FIG. 6 : Monolithic two cells filter

3/ A maximally low iterative impedance : A high impedance filter may be a disadvantage for the manufacture of the filter but also for the user : the user generally needs an impedance value of 50 to 200 ohms. If the filter exhibits a much higher impedance, matching components between filter and user need to be high Q. The practical limitation for low cost low size inductors are Q = 30. Thus, we choose to use materials for crystal leading to the lowest impedance. The other end of the compromise is given by the filter insertion loss which shall not exceed 6 dB.

4/ A low profile package : The radiotelephone will be a portable device and even a portative one. Size will be a critical parameter. The volume for a classical filter meeting the required shape is approximately 35 x 12 x 10 mm. Our objective for this work was to reduce the volume in order to reach small imensions. This aim is fully consistant with our previously described choices. Firstly, placing the four resonators on the same crystal blank is a great advantage as regard to size reduction. The resonators contribution to the whole size is only one fourth of the usual size. Secondly, the maximally low iterative impedance for the filter leads to a small size for the coils, due to the limited Q value required. Thirdly, the design - either in the scale, either in the monolithic form - avoids the need for such size consuming components as transformer.

But even so, the filter requires a special internal design to allow an easy tuning. This is yet under study. Nevertheless, benches models have yet been produced.

For the benches models we choose to use a TO8 package. Nevertheless, the technology we developped may be easily arranged in a SMD package.

RESULTS

Some response curves for encapsulated filters are shown on figure 7 to 10. They are very closed to our target specifications. The pass-band is conform for amplitude as well as for group delay.

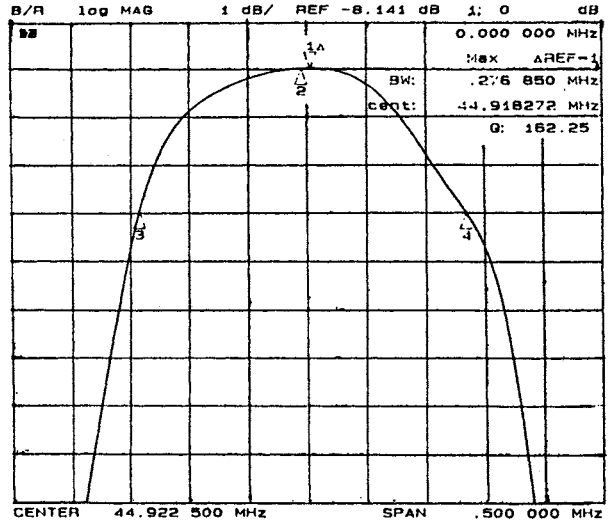


FIG. 7 : Pass band response for a lithium tantalate filter/ 4 poles per plates

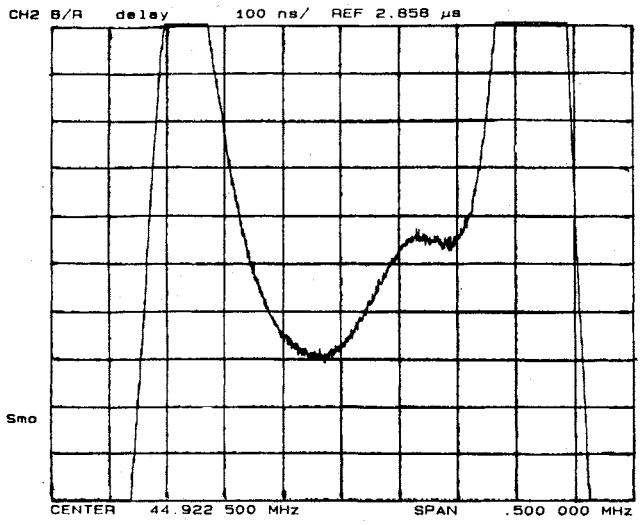


FIG.8 : Phase response for a lithium tantalate filter 4 poles per plate

The - 60dB shape factor is good. Compared to the specification generally required by customers, the response is still slightly too broad at - 20 dB and - 30 dB. This will be improved by a better filter synthesis.

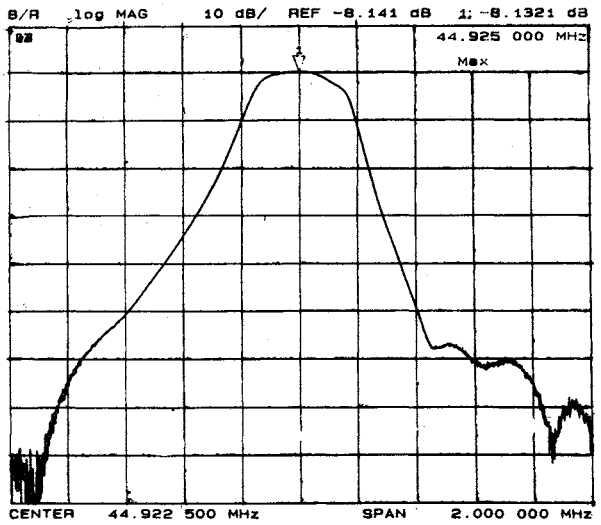


FIG 9 : Closed to band rejection for a lithium tantalate filter - 4 poles per plate

Rejection is - 60 dB. This is a good result because one may fear that the reduced distance between resonators or between resonators and components due to the small volume inside the package may have lead to a poor ultimate rejection.

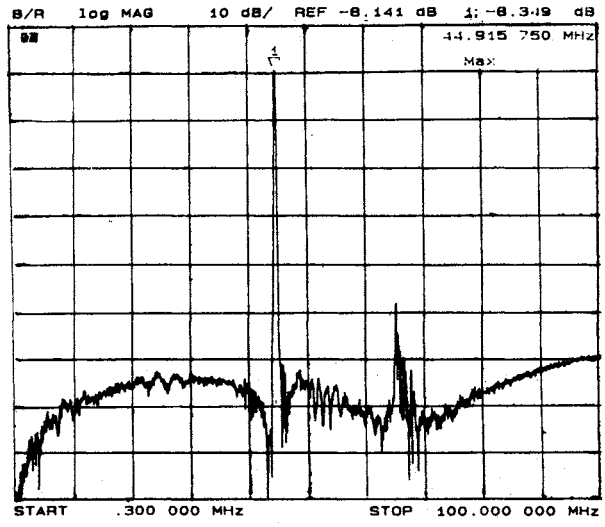


FIG.10 : Out of band rejection for a lithium tantalate filter / 4 poles plate

Spurious responses have been drastically reduced in amplitude as well as in number. Only one response remains at a level of - 45 dB at a distance of 20 MHz frome the center frequency. This response corresponds to a longitudinal propagation mode which is well known on lithium tantalate. We continue studies to reduce this response by modifying the physical arrangement of the resonators on the crystal blank. Compares the resulting filter to a classical equivalent filter.

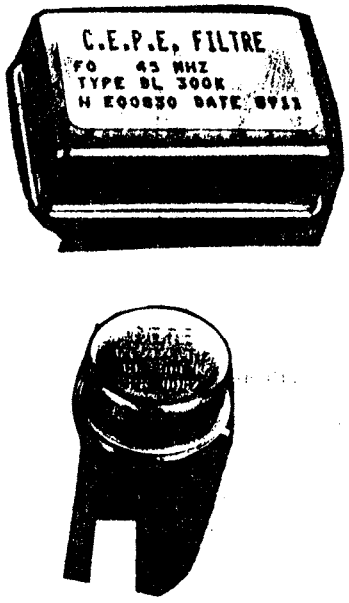


FIG. 11 : Compared dimension classical filter versus new filter

INDUSTRIALISATION

We have shown that numerous choices have been made in order to make the filter easily produceable at a low cost without jeopardizing the small overall size. We are also doing work on the manufacturing process in order to lead to reduced manufacturing costs.

The main lines for that work are :

1- materials saving, especially crystal saving by improved row material processing

2- etching automatic control

3- internal assembly design and mounting technology

4- automatic metallization of the resonators and tuning of the filter

Package is not yet concerned by this work because we don't know if customers will require a SMD package or a plug-in package.

COMPETITIVE TECHNOLOGIES

We have been led to the conclusion that crystal filter is a well fitted technology for the digital radiotelephone IF filter. But we needed before to confirm this opinion by looking at potential competitors, for example SAW filter. In order to make up our mind we have made a low loss SAW filter with such a design we have condidered as cost saving. This realization has been described in a previous paper during this conference [4].

The results are quite interesting : insertion loss is low enough, amplitude and group delay in the pass band are very good. Nevertheless it seems that rejection will hardly reach lower values than 50 dB which is, according to to-day customer specifications, not yet enough. Another main inconvenient is size, which is about 30 mm long.

Manufacturing cost is quite the same as for crystal filters mainly due to the need of a large substrate and of a 3 phases signal generation to achieve a low loss effect in the SAW filter.

CONCLUSION

After having compared the two main technologies seen to meet the specification for an IF filter for digital radiotelephone, we made our mind that crystal filter was a well fitted technology. For that purpose we have made choices to improve cost and size for the filter and to reduce manufacturing costs.

The results are benches models which are very closed to the required electrical specification and target size.

Work is continuing to develop the options above described in order to improve manufacturing process.

BIBLIOGRAPHY

[1] - A new Approach to Berlinite Crystal Growth
E. PHILIPPOT and Al
2nd EFTF 1988 - P. 843

[2] - Evaluation and Application of high purity Berlinite Crystals
J. DETAINT and Al
2nd EFTF 1988 - P. 873

[3] - An LCC Monolithic crystal filter
R.C SMYTHE
2nd EFTF 1988 - P. 637

[4] - Filtre RF faibles pertes à ondes élastiques de surface
JP. MICHEL
These Proceedings

OBSERVATION OF PIEZOELECTRIC ENVELOPE SOLITONS GENERATED IN THE BULK ACOUSTIC WAVE RADIATION FROM INTERDIGITAL TRANSDUCERS DEPOSITED ON QUARTZ PLATES

M. Planat

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

Solitons can be roughly thought as stable propagating wave packets resulting from a balance between nonlinear distortion and dispersive spreading. They have been observed recently in numerous fields like hydrodynamics, electrical transmission lines, plasma physics and optical fibers, but until now attempt to excite solitons in solid state acoustics have failed. In a previous work we showed that the bulk acoustic wave radiation from interdigital transducers deposited on a quartz crystal is angularly focused and that the direction of emission is scanned by the applied excitation frequency [1]. Using pulse excitation, it is found that wave packets propagate dispersively under the transducer lattice. On the other hand it is shown that the wave propagation is also highly nonlinear. Pulse instability, pulse narrowing, pulse splitting and dispersion compensation are demonstrated for 1 to 15 V amplitude and 25 ns to 200 ns duration pulses from a transducer with a 17.2 µm spatial period on a Y cut - Z propagation quartz substrate. Main features of experiments are interpreted by establishing the nonlinear Schrödinger equation which governs the pulse envelope function for the mechanical displacement. A reasonable agreement with experiments is found.

Introduction

Most physical phenomena in solid state electronics can be explained within the frame of a linear wave propagation theory. Nonlinear effects are generally introduced to explain the sensitivity of devices to temperature, pressure, acceleration, humidity ... Nonlinear effects are also introduced to explain propagation of harmonics of the main signal, amplitude-frequency effect in resonators and intermodulation products in the bandwidth of filters. In all cases nonlinearities are small and the problem can be satisfactory solved by a perturbation approach [1].

In this paper we will study a novel effect which cannot be explained with the perturbation approach because involved nonlinearities are strong : this is the process of nonlinear dispersion and of solitons. Solitons can be roughly thought as stable propagating wavepackets resulting from a balance between the broadening due to dispersion and the distortion due to nonlinearities. They have been already introduced in numerous fields like hydrodynamics [2], electrical transmission lines [3,4], plasma physics [5] and optical wave guides [6,7] to explain the observed space-time structures. Our goal is to show that nonlinear dispersion should be the key concept to explain space-time in stabilities in acousto-electronics as well. In this paper observation of modulational instability and soliton propagation of shear horizontal waves in quartz is reported for the first time.

First remarks on the experimental setup

A schematic of the experimental setup is shown on Fig. 1. We used a Y-cut quartz substrate (10 × 9 mm2 and thickness 1 mm). The bottom was unpolished and two interdigital transducers were deposited on the top normally to Z-axis. Each transducer had 100 periods, a 4.5 mm width, a spatial periodicity $\lambda_0/2$ = 17.2 µm and 800 Å thick aluminum.

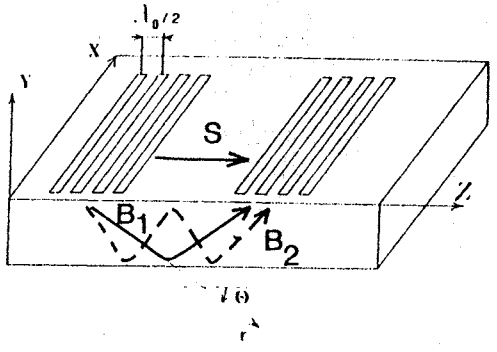


Fig. 1 : Schematic of the experimental setup

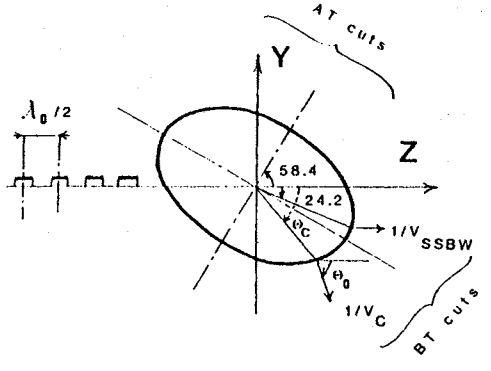


Fig. 2 : Slowness curve for singly rotated Y-cuts of quartz

In the linear regime such a cut supports only shear horizontal motion. Its slowness curve is shown on Fig. 2.

Since the excitation transducer is very long, propagation takes place with the wavenumber k along the direction of constructive interference θ_c given by the relation

$$k_Z = 2\pi/\lambda_o = k\cos\theta_c \qquad (1)$$

In other words the propagation angle θ_c can be scanned by the applied frequency f according to the relation

$$V_c = f\lambda_o\cos\theta_c \qquad (2)$$

The angle θ_0 for the propagation of energy is normal to the slowness curve (Fig. 2) and consequently there will be a cut-off in the device response when $\theta_0 = 0$. This angle corresponds to the propagation of the so-called surface skimming bulk wave (SSBW). In our device this happens for $\theta_c = 24.2°$ and the corresponding cut-off frequency is f_0 = 126 MHz.

The radiation efficiency for the shear wave in a Y-Z cut quartz half space was computed at the ray approximation [8] (Fig. 3). At the cut-off frequency the main lobe is broad and parallel to the surface ; with increasing applied frequency the lobe rotates and gets thinner.

The lobe broadening due to the finite length ΔL of the transducer can be estimated from the uncertainty relation

$$\Delta k \, \Delta L = 2\pi \qquad (3)$$

In our device, this leads to an uncertainty $\Delta f = V_c/\Delta L$ in frequency of the order 1 MHz. Broadening and secondary lobes due to the finite length of transducer are then to be expected for times larger than $\Delta t = 1/\Delta f = 1$ µs.

To avoid such modulations we will perform pulse experiments in a range of time scales between 10 ns and 100 ns. This means that these input pulses will propagate under a transducer which can be assumed to be infinite. To explain experiments both dispersion due to the strip lattice (transducer) and nonlinearity have to be taken into account.

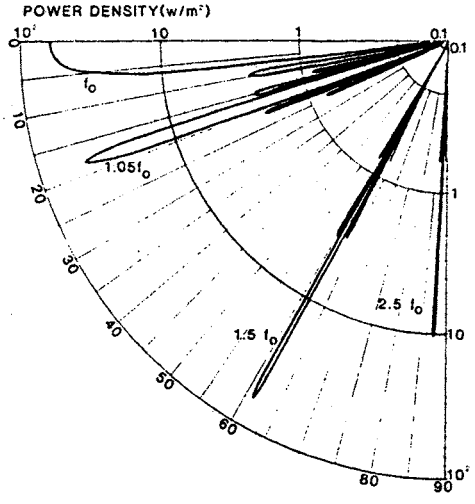


Fig. 3 : Power distribution for a 50-electrode-pair IDT on a Y-cut Z-propagation quartz crystal
f_0 = 126 187 MHz, input voltage ϕ_0 = 1 V

Derivation of dispersion relation and of the linear propagation equation for the pulse

We start with the equation for the slowness curve :

$$\rho\omega^2 = C_{55}\, k_Z^2 + C_{44}\, k_Y^2 + 2\, C_{45}\, k_Y\, k_Z \qquad (4)$$

with $C_{44} = 4 \times 10^{10}$ N/m2 ; $C_{55} = 5.8 \times 10^{10}$ N/m2 and $C_{45} = -1.8 \times 10^{10}$ N/m^2 for the Y-Z cut substrate.

Here $k_Z = 2\pi/\lambda_0$ can be considered constant in the range of time scales we will experiment, i.e. $\Delta t \ll 1$ µs. Introducing the wavenumber $k = (k_Y^2 + k_Y^2)^{1/2}$ along the propagation direction r (Fig. 1), the dispersion relation $\omega = W(k)$ is :

$$\rho\omega^2 = C_{55}\, k_Z^2 + C_{44}(k^2 - k_Z^2) + 2C_{45}\, k_Z (k^2 - k_Z^2)^{1/2} \qquad (5)$$

It is plotted on Fig. 4. Note that the dispersion is maximum at the SSBW frequency (zero group velocity) and decreases gradually when the frequency is increased.

For a modulated input pulse centered on the wavenumber k_0, corresponding to the frequency $\omega_0 = W(k_0)$ (Fig. 5), we can expand the frequency in Taylor series [9]

$$\kappa = k - k_0$$
$$\omega = \omega_0 + \kappa\,\omega'_0 + \frac{1}{2}\,\kappa^2\omega''_0 \qquad (6)$$

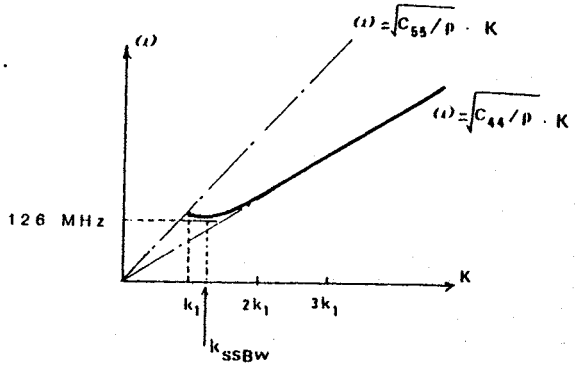


Fig. 4 : Dispersion curve for waves under a very long transducer on a Y-Z cut quartz substrate

where ω'_0 and ω''_0 are the group velocity and the concavity of the dispersion curve evaluated at the carrier wavenumber k_0.

The waveform $\Psi(r,t)$ is obtained by adding all monochromatic waves of wavenumber k and frequency $\omega = W(k)$ with the integral :

$$\Psi(r,t) = \int F(k)\, \exp\{i(kr - W(k)\,t)\}\, dk \qquad (7)$$

Extracting the carrier, we obtain from (6)-(7) :

$$\Psi(r,t) = \phi(r,t)\, \exp\{i(k_0 r - \omega_0 t)\} \qquad (8)$$

where $\phi(r,t)$ is the envelope of the pulse :

$$\phi = \int F(k_0 + \chi)\, \exp\left\{iKr - i\left(\kappa\omega'_0 + \frac{1}{2}K^2\omega''_0\right)t\right\} dK \qquad (9)$$

The interesting point of the analysis is that the envelope function obeys a Schrödinger type equation :

$$i(\phi_t + \omega'_0\phi_r) + \frac{1}{2}\,\omega''_0\,\phi_{rr} = 0 \qquad (10)$$

In the frame of the pulse eq. (10) is the conventional Schrödinger equation :

$$i\phi_t + \frac{1}{2}\,\omega''_0\,\phi_{rr} = 0 \quad \text{where} \quad r = r - \omega'_0 t \qquad (11)$$

Solutions of Eq. (11) are well known in quantum mechanics. During the propagation the pulse will be broadened and its frequency will be modulated. This is illustrated on Fig. 5.

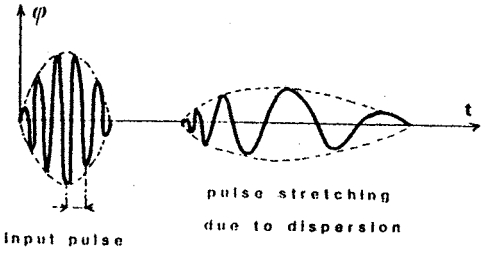


Fig. 5 : Propagation of a dispersion pulse

Nonlinear effects and pulse envelope for nonlinear wave propagation

An estimate of nonlinear propagation effects has been obtained from amplitude frequency measurements in resonators. For a travelling plane wave of amplitude a, the frequency shift is found to be [1]

$$\Delta\omega = k_o \Delta V = q a^2 \quad \text{where} \quad q = \frac{1}{16}\omega_o k_o^2 \frac{C^{(4)}}{C^{(2)}} \qquad (12)$$

where $C^{(2)}$ and $C^{(4)}$ are the effective second order and fourth order elastic constants. In our device the wave propagates in singly rotated orientations (rotation around X-axis) and third order constants vanishes by symmetry. The nonlinear coefficient $C = C^{(4)} / 12C^{(2)}$ is known to have an order of magnitude $C \simeq 10$ and is < 0 for AT cuts and > 0 for BT cuts [1, 10]. Referring to Fig. 2 it can be seen that orientations around the BT cut will be explored and then C is negative.

From the new dispersion relation

$$\omega = W(k) - q a^2 \qquad (13)$$

we can derive the pulse envelope for nonlinear wave propagation as follows [9]:

$$i(\phi_t + \omega'_o \phi_r) + \frac{1}{2}\omega''_o \phi_{rr} + q|\phi|^2\phi = 0 \qquad (14)$$

Solving Eq. (14) is a considerable task [11, 12] and only the nature of modulations predicted by Eq. (14) will be studied.

First case : $q\,\omega''_o < 0$

According to Witham's modulation theory [9], there are two real characteristics in this case. The system remains hyperbolic but the input pulse splits into two stable parts of two different group velocities. First experiments for this case were obtained recently in fiber optics by D. Krökel [13]. It has been shown that an input short pulse splits into two dark pulses (Fig. 6a).

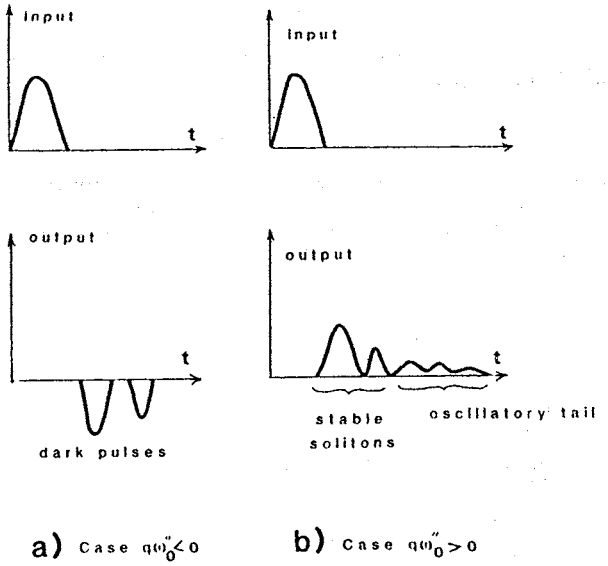


Fig. 6 : Schematic of solutions of the nonlinear Schrödinger equation

Second case : $q\,\omega''_o > 0$

There are two complex characteristics in this case. The system becomes elliptic and modulations are unstable. An input pulse generally splits in many stable parts called solitons and an oscillatory dispersive tail (Fig. 6b). Such behavior has been predicted in hydrodynamics as soon as 1965 by Sir J. Lighthill [14] and since then has been observed for water waves [2], electrical pulses [3,4], plasmas [5] and optical fibers [6,7].

For pulses in quartz BT cuts, we have $q > 0$ as just shown and $\omega''_o > 0$ from Fig. 4 ; then we will be restricted to this case.

An instability parameter and the fundamental soliton

For practical purposes two simple quantitative concepts to study the instability of systems ruled by the nonlinear Schrödinger equation can be settled.

First the modulational instability may be analyzed by assuming a solution with a uniform spatial state of amplitude a plus an infinitesimal propagating perturbation of frequency Ω and wavenumber K [15]. Substituting this form for the solution into (14) yields the following dispersion relation :

$$\Omega = K\left|\frac{1}{4}\omega''^2_o K^2 - q\,\omega''_o a^2\right|^{1/2} \qquad (15)$$

For $q\omega''_0 a^2 > \omega''^2_0 K^2/4 > 0$ the frequency Ω will become imaginary and the perturbation in the amplitude will grow exponentially in time.

The maximum growth rate occurs when $K = K_{max} = [2q a^2 / \omega''_0]^{1/2}$ with the value $\Omega = \Omega_{max} = qa^2$. Assuming that the wavepacket propagates approximately with the linear group velocity ω'_0 this allows to introduce the distance $x_{n\ell}$ for maximum instability and the corresponding modulating frequency f_m at which the growth rate is maximum [15]. We obtain :

$$x_{n\ell} = \omega'_o / qa^2 \quad \text{and} \quad f_m = \frac{\omega'_o K_{max}}{2\pi} \qquad (16)$$

A second glance to the modulational instability can be given by observing that Eq. (14) admits a family of localized solutions (called fundamental solitons) of the form [6,9]

$$\phi(r,t) = \phi_1 \operatorname{sech}\left(\frac{t - t_o - x/v_g}{t_o}\right) \exp i(rx - st) \qquad (17)$$

where τ_0 is the pulse half width, r and s are wavenumber and frequency shifts, $v_j = \omega'_0 + \omega''_0 r$ is the pulse velocity and ϕ_1 is the pulse amplitude given by

$$\phi_1^2 = \omega''_o / t_o^2 v_g^2 q \qquad (18)$$

A schematic of the shape of fundamental soliton (17) is shown on Fig. 7.

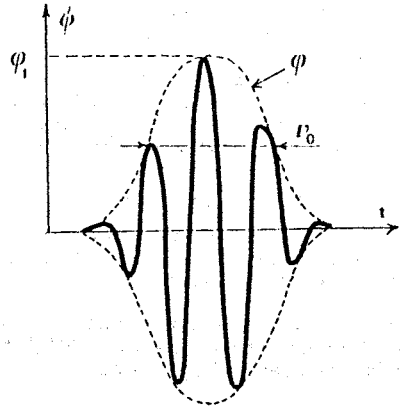


Fig. 7 : Shape of the fundamental soliton

Other effects due to modulational instability

A whole description of physical effects related to the nonlinear Schrödinger equation cannot be given here. We choosed to illustrate them by two examples borowed to recent experiments.

First, a perspective plot of the temporal pulse shape at various points along an optical fiber for the A = 3 soliton is shown on Fig. 8 [6,16]. The input soliton pulse first narrows, then splits and recurs as a function of propagation distance, as observed.

Second, the modulational instability in an electrical transmission line is shown on Fig. 9 for different values Ω/ω_0 between the envelope soliton frequency and the input frequency [17].

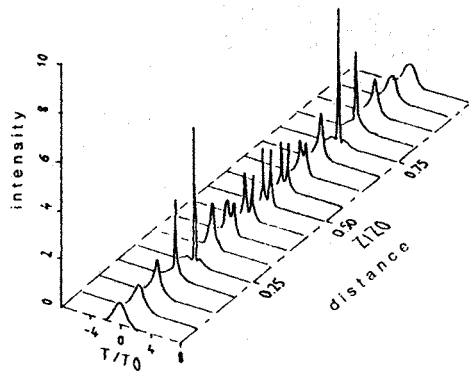


Fig. 8 : Perspective plot of the temporal pulse shape at various points along a fiber for the A = 3 soliton. The intensity variable is defined as $|V(z,t)/A|^2$. The parameters z_0 and t_0 are the spatial and temporal soliton periods (from Stolen, Mollenauer and Tomlinson [16])

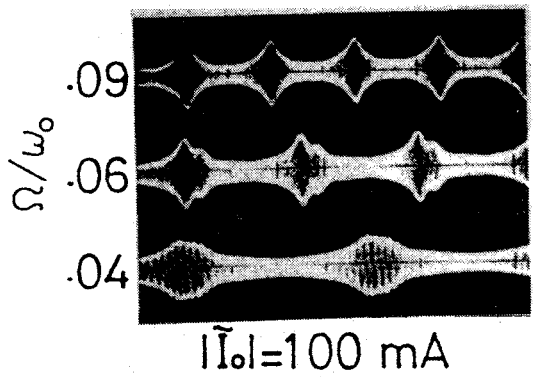


Fig. 9 : Observation of envelope solitons at a fixed point on the line. The ratio Ω/ω_0 is the ratio envelope soliton frequency/input frequency. (from Yagi and Noguchi [17])

Modulational instability and acoustic solitons in the Y-Z cut quartz plate

Theoretical expectations

The previous theory was used to estimate the possibility of observing nonlinear dispersion effects in Y-Z cut quartz plates.

Linear group velocity ω' and concavity of dispersion curves ω'' are shown on Fig. 10. In our setup the distance of maximum instability $x_{n\ell}$ and the amplitude of fundamental soliton are found from (12), (16) and (18)

$$x_{n\ell} = \omega'_0 / C\omega_0 k_0^2 a^2 \quad ; \quad \phi_1^2 = \omega''_0 / C\omega_0 k_0^2 v_g^2 \tau_0^2$$

These parameters have been plotted versus the carrier wavenumber k_0 for an input pulse amplitude a = 100 Å and an input pulse width τ_0 = 100 ns (Fig. 11).

In our device, since $\omega''_0 > 0$ and $q > 0$ then $\omega''_0 q > 0$ and modulations are unstable. Experiments will be performed in the range of carrier wave numbers between $k_1 = 2\pi/\lambda_0$ and $3k_1$, corresponding to frequencies between 100 and 300 MHz. At 126 MHz, $x_{n\ell}$ vanishes and strong modulational instability should be observed. At higher frequencies, $x_{n\ell}$ indreases, then decreases again towards approximately 5 mm at about 300 MHz. At this frequency, $\phi_1 \simeq 100$ Å (Fig. 11), which is in the range of applied excitations.

Using these estimations we can expect to observe soliton behavior at high frequencies ($\simeq$ 300 MHz), on the second or third bounce (corresponding to sufficient modulational instability) and at sufficiently high input lengths and amplitudes ($\tau_0 \simeq$ 100 ns and $a \simeq$ 100 Å).

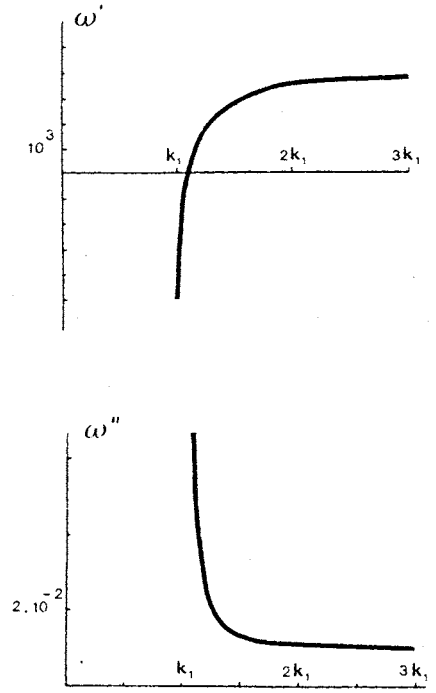


Fig. 10 : Linear group velocity and concacity

of the dispersion curve of Fig. 4

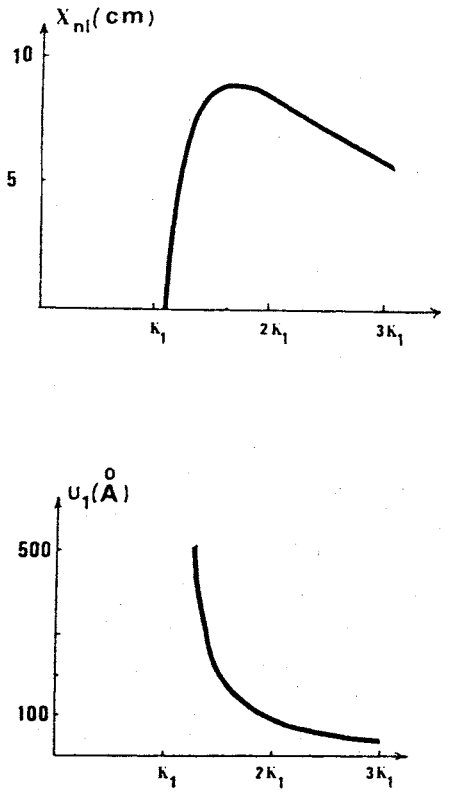


Fig. 11 : Distance of maximum instability $xn\ell$

and amplitude of the fundamental soliton

for a = 100 Å and τ = 100 ns

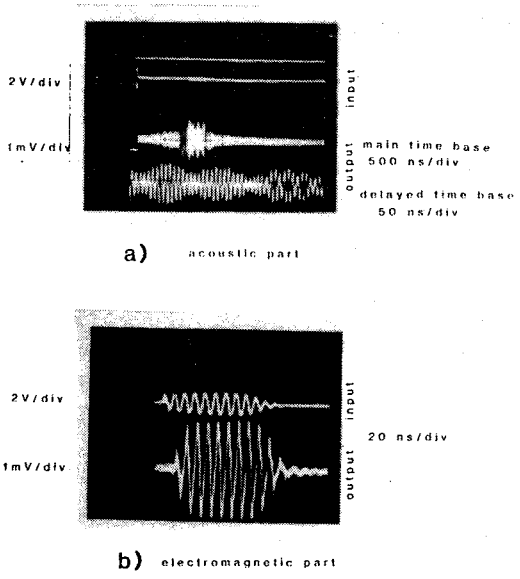


Fig. 12 : Modulational instability

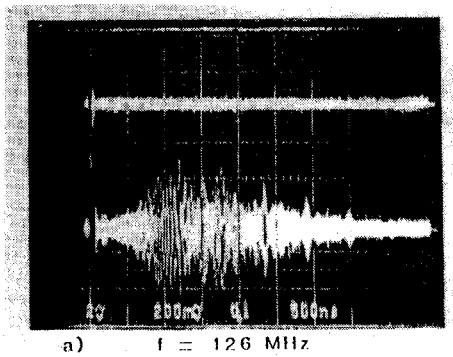
observed closed to the cut-off frequency

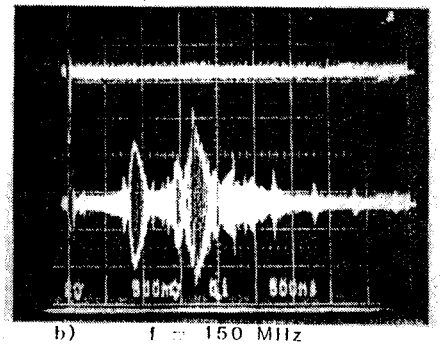
Experiments

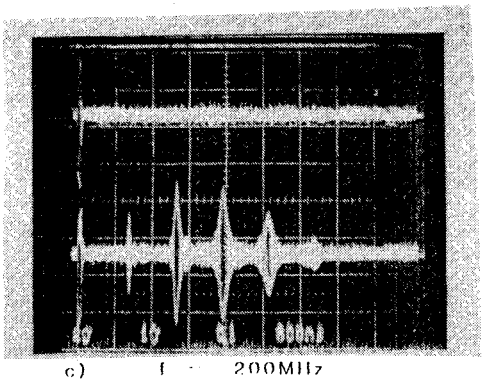
In a first set of experiments, rectangular pulses of 75 ns length and 2V amplitude were applied to the Y-Z cut device described at the beginning of this paper (§ 2). The applied modulational frequency 133 MHz was chosen to correspond to a strong instability (approximate zero group velocity). The output wavepacket (Fig. 12a) is observed to be very unstable (see the delayed time base) and extents over a 3 µs time window. The center wave packet delay 1.5 µs correspond to the velocity 4000 m/s of SSBW and the large time window reflects the finite length of the transducers.

In Fig. 12a the oscilloscope trace has been synchronized on the electromagnetic response of the delay line. An enlarged view of this response reveals a nonlinear structure which is correlated to the presence of the acoustic signal (this phenomenon will be analyzed in a future work).

In a second set of experiments input pulses of 50 ns duration and 7 V amplitude where fed into the input transducer at increasing frequencies. At the cut-off frequency 126 MHz) the response is broad and unstable as before (Fig. 13). At 150 MHz (Fig. 13b), the response has essentially split into two pulses corresponding to the two bounces illustrated on Fig. 1. At 200 MHz five bounces are observed. The time delay of each bounce decreases and then their apparent velocity increases with increasing frequency. Also the width of each bounce decreases with increasing velocity. It should be remarked that the notion of bounce has not to be used too far because there is not propagation in the conventional (hyperbolic) sense since the system is elliptic. On the other hand the amplitude dependence of bounce velocity and width is characteristic of nonlinear dispersive waves. The bounce propagation is reminiscent of that of a fundamental soliton. To confirm the hypothesis new experiments were performed.

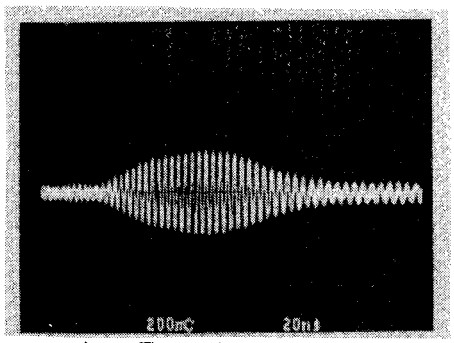




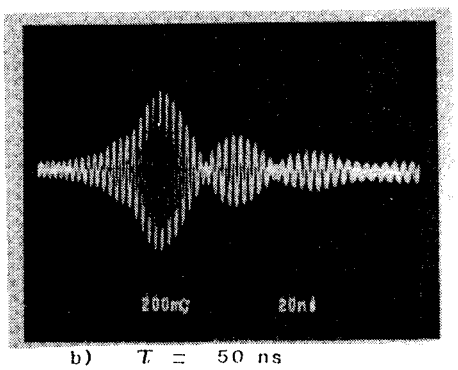


c) f = 200MHz

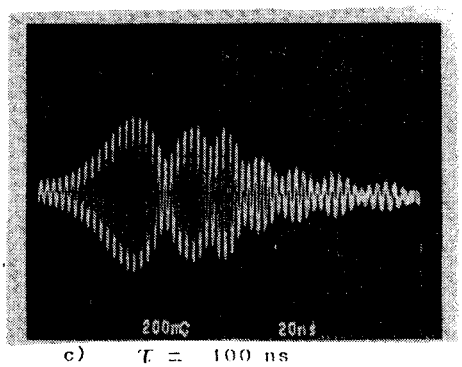
Fig. 13 : Impulse response of the quartz plate versus carrier frequency. Input pulse length : 50 ns, input amplitude : 7 V



a) T = 25 ns

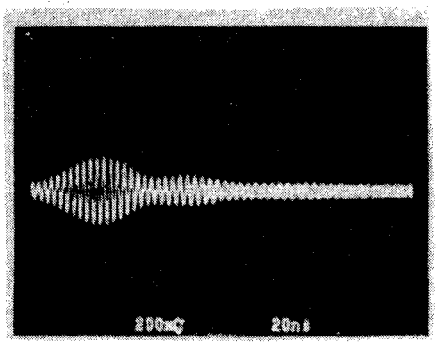


b) T = 50 ns

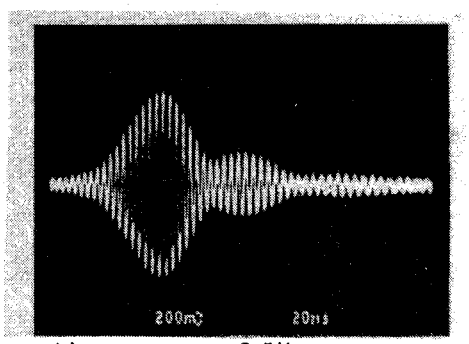


c) T = 100 ns

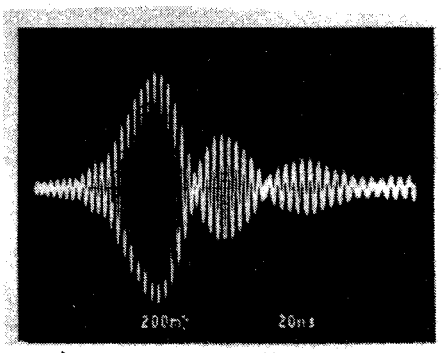
Fig. 14 : Impulse response of the quartz plate corresponding to the second bounce versus input length.
Carrier frequency : 280 MHz, input amplitude : 13 V



a) v = 2.5 V



b) v = 7.5 V



c) v = 15 V

Fig. 15 : Impulse response of the quartz plate corresponding to the second bounce versus input amplitude.
Carrier frequency : 280 MHz, input length : 40 ns

In Fig. 14 and 15 incident pulses with fixed frequency 280 MHz and variable length τ and amplitude v were considered. To make the oscilloscope record easier a 45 dB magnification of the output signal was used and all records were taken from the second bounce. In Fig. 14 the incident pulse had 13 V maximum amplitude. At 25 ns input length (Fig. 14a) the wave packet is very dispersive and the modulated frequency extents from 300 to 220 MHz. At 50 ns input length (Fig. 14b) the wave packet has split into three stationnary parts of frequencies 280 MHz, 256 MHz and 225 MHz. Dispersion is present only at the edges of the main signal. Then at 100 ns input length (Fig. 14c), modulational instability of the envelope is clearly seen and the frequency 280 MHz of the major part of the packet is that of the incident pulse. In Fig. 15, the incident pulse had 40 ns input length and variable amplitude. At 2.5 V (Fig. 15a) input amplitude the wavepacket is widely spread and is dispersive. At 7.5 V (Fig. 15b) input amplitude the wavepacket has split into two parts with frequencies 280 MHz and 244 MHz. At 15 V input amplitude (Fig. 15c) the wave packet has split into three parts with frequencies 280 MHz, 244 MHz and 225 MHz. Moreover an approximate 50 % narrowing has occured for each packet.

Conclusion

In this paper we have shown that bulk acoustic wave radiation from very long interdigital transducers is dispersive and nonlinear. At the SSBW cut-off frequency theory revealed a linear zero group velocity corresponding to a strong modulational instability. This instability was observed together with a nonlinear shape of the electromagnetic breakthrough.

At higher frequencies, pulse narrowing, pulse splitting and modulational instability corroborated by the existence of solitons was observed in rough agreement with quantitative estimate.

Applications of solitons shoudl be mainly toward the field of coding and decoding (soliton correlator). The instability concept encountered in nonlinear dispersion should also be a valuable tool to explain 1/f frequency instabilities in resonators. Lastly characterization of materials by solitons should be useful for basic physics.

References

[1] M. Planat, G. Théobald and J.J. Gagnepain, l'Onde Electrique 60(8-9) (1980) and 60(11) (1980).
[2] J.E. Feir, Proc. Roy. Soc. London, A299, 54 (1967).
[3] K.E. Lonngren, in "Solitons in Action", edited by K. Lonngren and A.Scott (Academic Press, New York, 1978), p. 127.
[4] H. Ikezi, S.S. Wojtowicz, R.E. Woltz, J.S. de Crassie and D.R. Baker, J. Appl. Phys. 54, 3277 (1988).
[5] H. Ikezi, in "Solitons in Actions", edited by K. Lonngren and A.Scott (Academic Press, New York, 1978), p. 153.
[6] L.F. Mollenauer, R.H. Stolen and J.P. Gordon, Phys. Rev. Lett. 45, 1095 (1980).
[7] S. Maneuf, R. Desailly and C. Froehly, Optics Communications 65, 193 (1988).
[8] Y. Zhang, M. Planat, Electronics Lett. 32, 68 (1987).
[9] G.B. Witham, "Linear and nonlinear waves", John Wiley & Sons, 1974, p. 601.
[10] J.J. Gagnepain and R. Besson, "Nonlinear effects in piezoelectric quartz crystals", Physical Acoustics, vol. XI, Academic Press, New York (1975).
[11] V.E. Zakharov, A.B. Shabat, Sov. Phys. JEPT 34, 62 (1972).
[12] J. Satsuma, N. Yajima, Suppl. Prog. Theor. Phys. 55, 284 (1974).
[13] D. Krökel, N.J. Halas, G. Giuliani, D. Grischkowsky, Phys. Rev. Lett. 60, 29 (1988).
[14] M.J. Lighthill, J. Inst. Math. Applic. 1, 269 (1965).
[15] J.F. Ewen, R.L. Gunshor and V.H. Weston, J. Appl. Phys. 53, 5682 (1982).
[16] R.H. Stolen, L.F. Mollenauer and W.J. Tomlinson, Opt. Lett. 8, 186 (1983).
[17] H.C. Yuen and B.M. Lake, in "Solitons in Action", edited by K. Lonngren and A.Scott (Academic Press, New York, 1978), p. 89

AUTEUR/AUTHOR	PAGE
ARNAUD R.	57
AUBRY J.P.	362, 387
AUDOIN C.	277
AULD B.A.	152
AVINENS C.	57
BACHHEIMER J.P.	61
BALLANDRAS S.	159
BARILLET R.	249
BATES A.G.	11
BAVA E.	259
BEAUSSIER J.	221, 238
BEETLEY D.E.	139
BERNIER L.G.	215
BIDART L.	181, 187
BIGLER E.	159
BIGNON O.	57
BLONDE D.	362, 387
BOURGEOIS C.	235
BOURQUIN R.	305
BRENDEL R.	28
BREUZET M.	221
BUISSON X.	57
BUSCA G.	215
BUZEK O.	204
CALDERA C.	259
CALISTI S.	165
CAPELLE B.	57, 301
CEREZ P.	81, 277
CERMAK J.	204
CLOEREN J.M.	11
CORDARA F.	204
CRACKNELL M.P.	169
CRAVEUR J.C.	221
DALY P.	121
DAUMET J.	356
DE CLERCQ E.	281
DEFRANOULD P.	165
DE HOND P.	293
DE JONG G.	198
DE LABACHELERIE M.	81
DE LA FOURNIERE P.	387
DELAITE R.	296
DESBOIS J.	381
DETAINT J.	227, 301
DIEULESAINT E.	293
DIMARCQ N.	274
DIOMANDE K.M.	81
DOLINO G.	61
DUFILIE P.	381
DULMET B.	305
ELCHEIKH R.E.	194
EL GHAZI A.	255
ENGEL G.	50
ERTMER W.	264
FASSL S.	94
FOISE J.W.	337
GAIGNEBET J.	220
GALLIANO P.G.	204
GALLIOU S.	147
GAUTIER H.	369, 381
GAY M.	221
GIORDANO V.	274
GIRARDET E.	181, 187
GODONE A.	259
GOIFFON A.	57
GRANVEAUD M.	104
GROSLAMBERT J.	173, 245
GUINOT B.	1
HADORN F.	269
HAMEL A.	277
HARRISON A.M.	169
HAUDEN D.	159, 165
HELLWIG H.	5
HOBDEN M.K.	180
HODE J.M.	381
HRUSKA C.K.	35
IMAE M.	89
IRVINE R.A.	337
JOHNSON G.R.	337
JOHNSON L.	73
JOHNSON W.A.	69
JOLY C.	227
KARUZA S.K.	69
KEMSSU P.	81
KHAYAR M.	356
KIRSHNER D.	94
KITCHING I.D.	121
KREMPL P.	50
KULMALA T.	326
LEBLOIS P.	173
LEBLOIS T.	41, 346
LE CORRE J.L.	387
LESCHIUTTA S.	220
LESTRADE J.F.	64
LEWANDOWSKI W.	89, 104
MAITRE P.	41
MANGIN P.	281
MARSHALL R.J.T.	133
MERIGOUX H.	356
MICHEL J.P.	372, 387
MIRANIAN M.	89
MORBIEU B.	293
MORLEY P.E.	133
MOUREY M.	147
MULLER J.H.	264
NELLESSEN J.	264
NORTON J.R.	11
OLIVIER M.	255
PEIER U.	11
PENAVAIRE L.	165, 369
PETIT P.	277
PETTITI V.	204
PHILIPPOT E.	57, 227
PLANAT M.	394
RENOULT P.	187
RESSLER H.	94
ROYER D.	293
RUBIOLA E.	313
SASSI M.P.	259
SAXENA G.M.	85
SCHLUETER B.	11
SCHUMACHER P.	319
SCHWARTZEL J.	227, 301
SEBASTIAN M.T.	57
SENGSTOCK K.	264
SHARPE D.J.	169
STADLER J.	50
STOERMER P.	143
SUTER J.J.	11
TAVELLA P.	204
TELLIER C.	41, 346
THEOBALD G.	274, 277
THOMANN P.	73, 269
THOMAS C.	89
TOURDE R.	104
UEBERSFELD J.	221
UHRICH P.	104
VALENTIN M.	194
VIALLE N.	362
VIARD B.	356
VIENNET J.	333
VOIT F.J.	69
VONDRAK J.	204
WILLIAMSON R.J.	22, 133
ZARKA A.	57, 301
ZHENG Y.	301

Imprimé en France

par l'Imprimerie du Conseil
Général du DOUBS
Hôtel du Département
Avenue de la Gare d'Eau
25031 BESANCON CEDEX

Printed in France

by the Printing Office of Conseil
Général du Doubs
Hôtel du Département
Avenue de la Gare d'Eau
25031 BESANCON CEDEX

Copyright by
3e Forum Européen Temps-Fréquence

Tous droits de reproduction et de traduction réservés
pour tous pays y compris l'U.R.S.S.

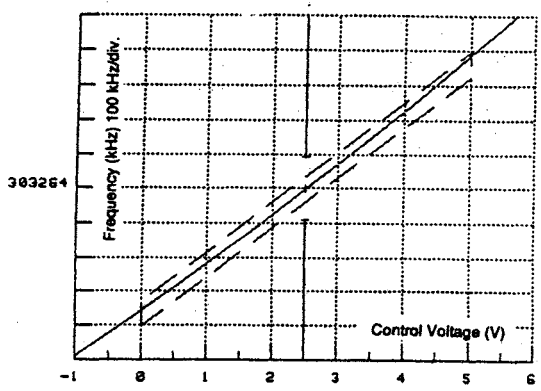


Fig.5 - Frequency control characteristic at 30°C of a narrowband VCO (F = 303.3 MHz, ΔF = 730 kHz).

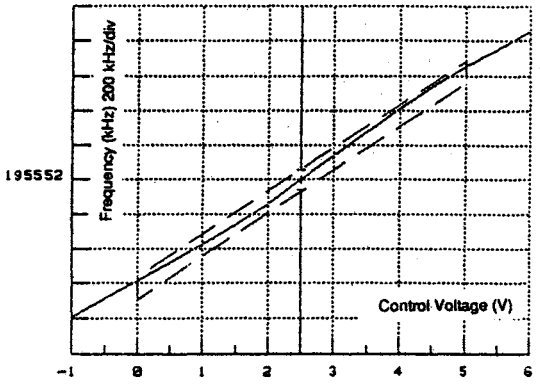


Fig.6 - Frequency control characteristic of a wideband VCO (F = 195.5 MHz, ΔF = 1250 kHz) at 30°C.

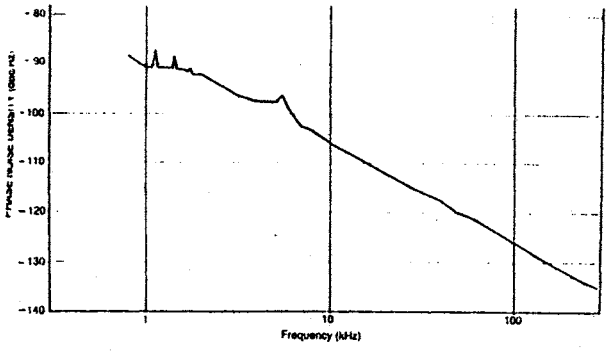


Fig.8 - Phase noise characteristics of VCO centered at 210 MHz with nominal ΔF = 700 kHz.

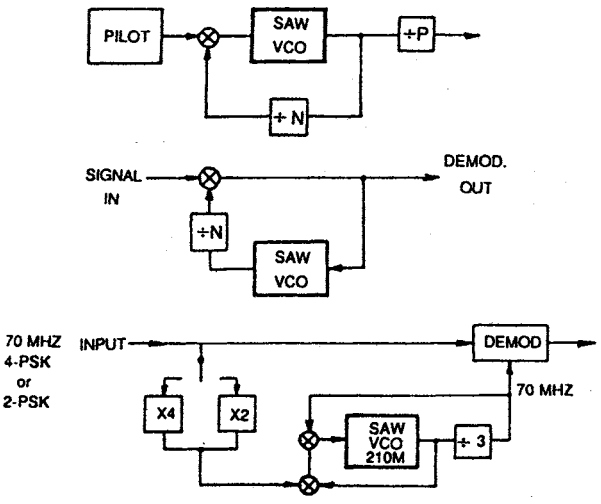


Fig.9 - Major applications of the compact SAW VCO

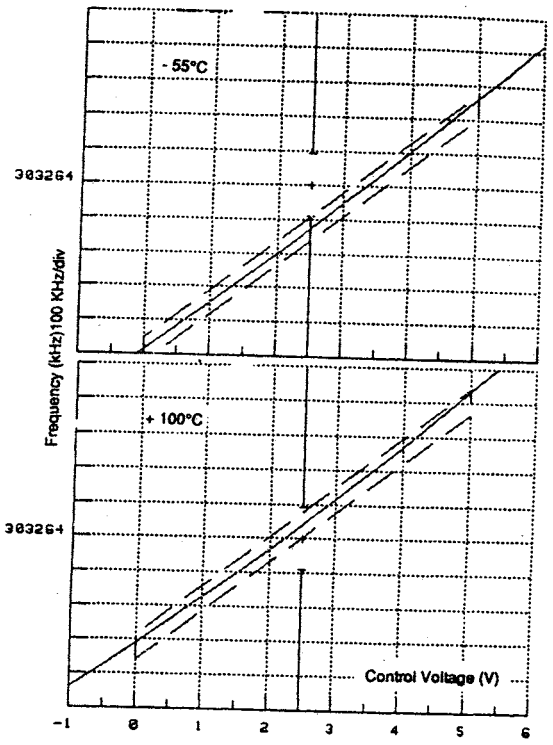


Fig.7 - Frequency control characteristics of a narrowband VCO (F = 303.3 MHz, ΔF = 730 kHz) at -55°C and +85°C.

4. - APPLICATIONS

These good electrical performances and the small size of the SAW VCO series SLC5000 make it well suited for ground, mobile and airborne applications under medium to severe environment.

Examples of actual implementations are schematized in Fig. 9. The oscillator is used in a phase-locked loop to benefit from its VCO capability.

(a) Locking the VCO on a low frequency pilot allows compact and high quality clocks to be built. The VCO bandwidth is chosen to exceed the sum of all instability causes.
(b) Frequency (or phase) demodulation of narrowband signals is simply performed through phase-locking of a VCO and reading of the control voltage signal. This application often requires large control bandwidth and values of 1-2 MHz are not uncommon.
(c) The last example relates to 70 MHz carrier recovery in a digital communication system employing 4-phase and 2-phase PSK modulation. A VCO operating around 210 MHz was developed and it now equips many modems in the French telecom network.

CONCLUSIONS

A family of Voltage Controlled Oscillators has been built using hybrid and SAW technologies. They operate in the VHF range and they offer low phase noise, good control linearity, together with small size and ruggedness. Present state-of-the-art in SAW design techniques and high frequency hybrid circuitry will allow the frequency range to be extended to 2.5 GHz. For higher frequency of operation, it is likely that the existing YIG and dielectric resonator technologies will continue to meet most requirements.

ACKNOWLEDGEMENTS

The authors wish to thank their colleague J. Elzière for his helpful collaboration in the definition of the custom hybrid circuits.

FILTRE FI FAIBLES PERTES A ONDE DE SURFACE

JP. MICHEL

C.E.P.E. 44 avenue de la Glacière 95100 ARGENTEUIL

La réalisation de filtres entre 30 et 120 MHz avec des bandes passantes voisines de 300 KHz (bandes relatives aux alentours de 0,5 %) pose des problèmes aux concepteurs car ils ont le choix entre deux technologies : 1/ les filtres à onde de volume 2/ les filtres à onde de surface.

Les deux technologies sous leurs formes traditionnelles sont distinctes, mais la progression de l'une et de l'autre les amènent à se recouper dans le domaine considéré. Elles présentent chacune leurs avantages et leurs inconvénients :

1/ Les filtres à onde de volume : ils sont traditionnellement limités à des fréquences inférieures à 35 MHz par des bandes passantes < 0,5 %.
Ils sont constitués autour de résonateurs à quartz usinés mécaniquement couplés entre eux à l'aide de capacités et de selfs.
Les limites sont repoussées grâce à des technologies spécifiques : a) l'emploi de matériaux à coefficient de couplage piézoélectrique supérieur au quartz permet d'élargir les bandes passantes. Ainsi, la berlinite autorise des bandes passantes voisines de 1 % et le tantalate de lithium s'adresse à des bandes de quelques pourcents. b) la fréquence centrale des résonateurs peut être élevée au delà de 35 MHz en prolongeant l'amincissement mécanique des lames de résonateurs par un usinage ionique ou un usinage chimique. On parvient ainsi à commercialiser des filtres à fréquence supérieure à 200 MHz.

2/ Les filtres à onde de surface : ils sont adaptés au domaine de fréquence et de bande passante, mais présentent traditionnellement une perte d'insertion élevée. Les techniques proposées dans la littérature pour surmonter cet inconvénient sont désormais nombreuses.

La nécessité de répondre à des demandes de filtres accrues dans le domaine fréquence x bande passante considéré et en particulier les demandes de filtres FI pour le radiotéléphone numérique cellulaire paneuropéen nous ont conduit à développer une technologie faible perte pour nos filtres à onde de surface afin de permettre une comparaison équitable de performances avec les filtres à onde de volume.

Les caractéristiques objectives que nous nous sommes fixées sont les suivantes

- fréquence centrale 30 à 120 MHz
- bande passante $\sim$ 300 KHz
- variation du temps de groupe < 300 ns
- raideur équivalente : 4 pôles
- réjection : 60 dB
- perte d'insertion < 6 dB sur 50 ohms
- encombrement : minimal
- quantité : > 100.000 pièces par an

I/ Choix d'une technologie faible perte

Les technologies faibles pertes pour les filtres à onde de surface peuvent être classées en 3 catégories 1/ les transducteurs à réflecteurs répartis 2/ les filtres à transducteurs d'entrée et de sortie imbriqués 3/ les transducteurs triphasés.

Ces trois principes ont été discutés selon les critères suivants 1/ performances de pertes faibles démontrées 2/ coûts de fabrication réduits compatibles avec des séries de 100.000 filtres 3/ développement réduit pour permettre une réponse rapide au marché.

a) Les transducteurs à réflecteurs imbriqués

Ils se présentent sous deux formes : 1/ le transducteur est composé de blocs de doigts séparés par des espaces occupés par des réflecteurs partiels [1]. 2/ le transducteur est continu et composé de doigts dédoublés. Un des deux doigts jumeaux est surchargé par une métallisation lourde donnant à ce doigt un pouvoir réflecteur supérieur [2].

Dans les deux cas, l'effet unidirectionnel est obtenu par un décalage de $\lambda/8$ entre le barycentre des émissions et le barycentre des réflexions. Cette disposition donne en théorie la perte faible d'un résonateur associée à la largeur de bande du transducteur.
L'avantage majeur de cette technologie est l'absence de réseau de déphasage pour alimenter le transducteur. Mais la technologie semble délicate : les deux niveaux de masquages doivent être réalisés

avec une grande précision de superposition et la performance du produit dépend de l'épaisseur de la surchage sur le réflecteur, difficile à maîtriser en pratique.

b) Les filtres à transducteurs d'entrée et de sortie imbriqués

Cette technique a été proposée par plusieurs auteurs ; la présentation la plus élaborée a été faite par Hitachi [3]. Cette technique semble très prometteuse tant par les performances annoncées que par la facilité de fabrication : pas de réseaux de déphasage et un seul niveau de masquage.

Cependant, les études de base ne nous ont pas semblé encore assez avancées pour permettre une mise en fabrication compatible avec le délai que nous prévoyons.

c) Les transducteurs triphasés

Ce principe est le plus traditionnel des principes faibles pertes. Il se présente sous deux formes : 1/ les doigts des transducteurs sont alternativement réliés à trois phases espacées de 120° ou de 90°. Dans ce cas, les pistes doivent se chevaucher ce qui impose une technologie à plusieurs niveaux de masquage [4]. 2/ les transducteurs sont formés d'une succession de blocs alimentés alternativement entre une phase et la masse puis entre l'autre phase et la masse [5]. Pour éviter les croisements de pistes, la piste de masse présente alors un tracé en forme de grecque. Il n'y a plus qu'un niveau de masquage, mais la périodicité des blocs apporte des réponses indésirables hors bande.

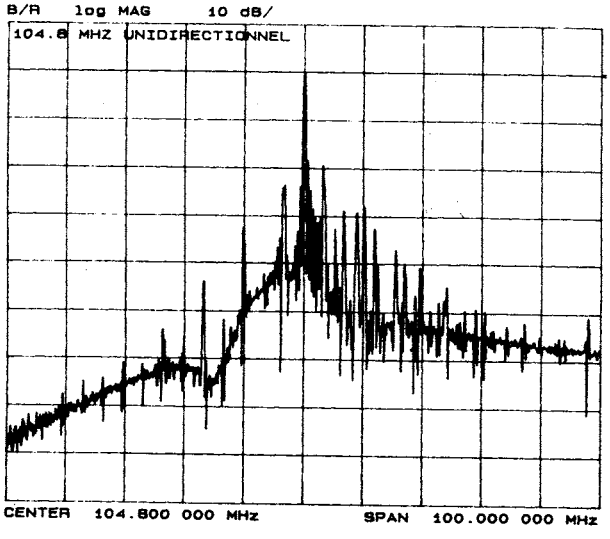
Les deux dispositions présentent en outre l'inconvénient de nécessiter des réseaux de déphasage. Cependant les performances en perte d'insertion sont bonnes, la synthèse est réalisable et les principes sont bien maîtrisés autorisant une mise en oeuvre industrielle rapide.

C'est donc cette dernière technologie que nous avons retenue.

Dans l'optique d'une minimisation des coûts de fabrication, la technologie à plusieurs niveaux de masquage nous est apparue comme une lourdeur. Nous avons donc opté pour la technologie des transducteurs à blocs.

II/ Optimisation du transducteur faibles pertes

Un premier filtre a été réalisé à 104,8 MHz. Chaque transducteur comportait 16 blocs de 32 doigts. L'alimentation en masse était réalisée par une grecque . Aucun transducteur n'est pondéré. Les réponses obtenues sont sur la courbe 1.Il est apparu -ce qui était prévisible- deux défauts majeurs vis à vis de l'objectif : 1/ la périodicité des blocs apporte de sérieuses réponses parasites hors bande. 2/ la grecque de masse, conducteur long et étroit était trop résistive ; il a fallu faire des rappels de masse par bondings en de nombreux points pour obtenir une bonne performance en perte d'insertion.



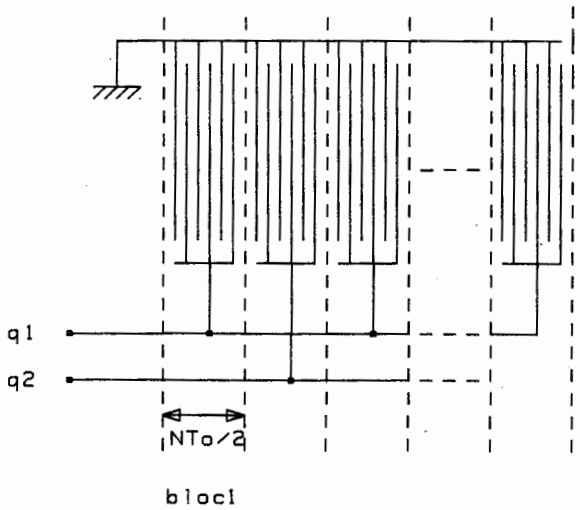
COURBE 1 : Filtre unidirectionnel à 104,8 MHz

Il est donc apparu nécessaire d'établir un certain nombre de règles de conception visant à minimiser ces inconvénients.

a) Analyse simplifiée du transducteur

Le transducteur à blocs peut être représenté dans le domaine temps par la combinaison de 3 fonctions : 1/ la fonction trame t(t) 2/ la fonction bloc b(t) 3/ la fonction enveloppe e(t)

La description du transducteur est la suivante : il comporte 2α blocs constitués chacun de N sources espacées de To/2 (fig.1).



q1 et q2 sont les 2 alimentations électriques à phases décalées.

FIGURE 1 : Description du transducteur

La fonction de transfert du transducteur h(t) s'écrit alors :

$$h(t) = (t(t) * b(t)).e(t)$$

* est le symbole du produit de convolution.

Les trois fonctions sont décrites dans le tableau 1.

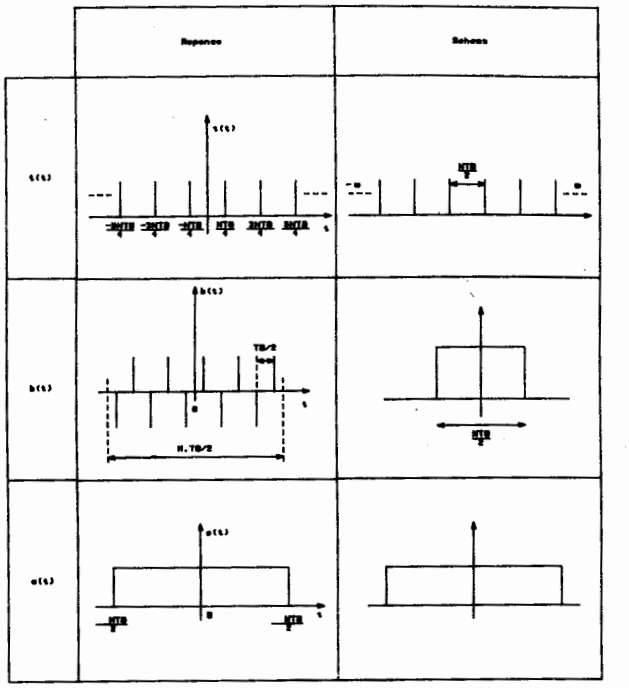


TABLEAU 1

Remarque :

- δ est le symbole de la fonction de Dirac
- e(t) dépend de la réponse en fréquence désirée. Dans le cas présent, nous supposons la réponse obtenue sans pondération

Ces fonctions se tranforment dans le domaine fréquence selon le tableau 2

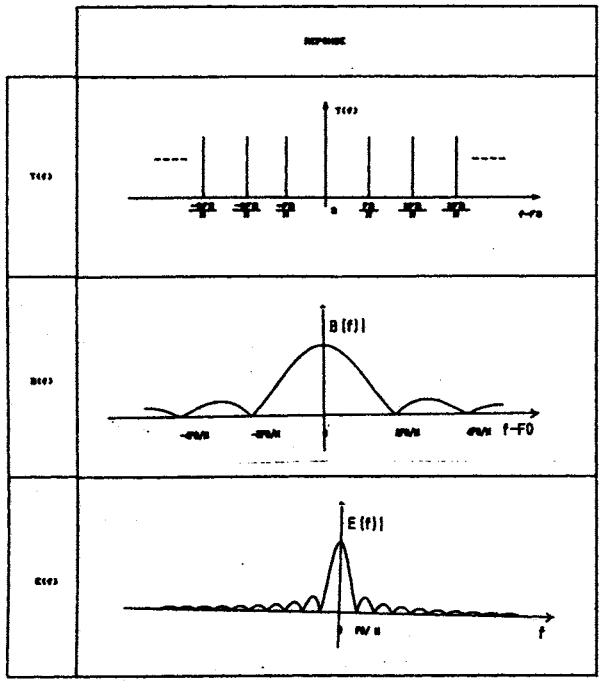


TABLEAU 2

La réponse globale en fréquence s'écrit

$$H(f) = (T(f).B(f)) * E(f)$$

La fonction T(f).B(f) dite fonction de structure, donne les différentes réponses de la structure. On convient d'appeler réponse parasite d'indice k les réponses pour lesquelles k # 0 et réponse principale celle pour k = 0.

On constate que la structure élimine les réponses parasites d'ordre pair (coïncidence d'un zéro de B(f) et d'un pic de T(f)) et ne concerne que les réponses parasites d'ordre impair d'amplitude décroissant en 1/k.

Par le biais de la convolution, la réponse d'enveloppe E(f) se manifeste sur chaque réponse de la structure selon la figure 2

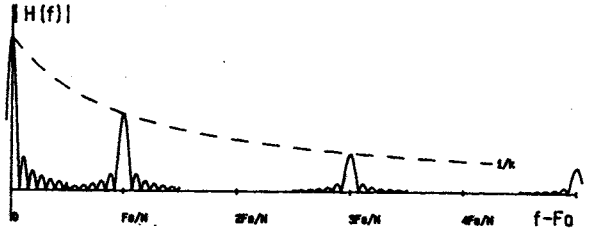


FIGURE 2

Cette allure de réponse est bien vérifiée sur la courbe réelle (courbe 1).

Les réponses supplémentaires obtenues pour f > Fo sont des réponses d'onde de volume facilement éliminables.

b) Optimisation de la dimension des blocs

La présence des réponses parasites étant due à la périodicité de la structure, une première hypothèse consiste à détruire cette périodicité. Des simulations ont été effectuées en donnant aux blocs des longueurs non plus constantes mais aléatoires : l'amélioration est effective mais insuffisante.

Dans la suite, les blocs sont à nouveau supposés de longueur constante. On examine l'influence de la longueur du bloc. Plus les blocs sont petits, plus la première réponse parasite est éloignée de la réponse principale. A priori, cette situation est favorable pour le filtre. Une analyse plus fine a montré le contraire, par suite de la prise en compte du nécessaire espacement entre les blocs.

Soit la structure s(t) définie par le schéma de la figure 3.

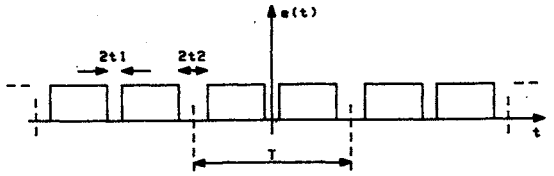


FIGURE 3

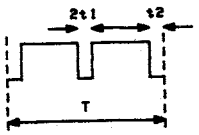
Cette structure comporte une succession infinie de blocs identiques espacés alternativement de $2t_1$ et $2t_2$.

$$s(t) = t'(t) * b'(t)$$

avec :

$$t'(t) = \sum_{k=-\infty}^{+\infty} \delta(t-kT)$$

$b'(t)$ =



La réponse de la structure s'écrit :

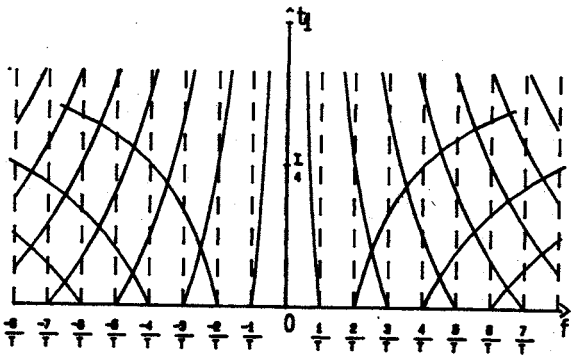
$$S(f) = T'(f)\, B'(f)$$

$$T'(f) = \sum_{k=-\infty}^{+\infty} \delta(f - 2\pi k/T)$$

$$B'(f) = \sin \pi f (t_1 + t_2 - T/2) \cos \pi f (T/2 + t_1 - t_2)$$

On recherche t_1 et t_2 pour lesquels un nombre maximum de pics de $T'(f)$ se trouve annulé par des zéros de $B'(f)$.

La figure 4 présente le lieu des zéros de $B'(f)$ pour t_1 variant de 0 à sa valeur maximale possible T/2 (en posant $t_2 = 0$)



En trait plein: lieu des zeros de $E_1(f)$
En pointilles: lieu des pics de $E_2'(f)$

FIGURE 4

On constate que plus t_1 augmente, moins les annulations de pics de trame sont fréquentes. L'optimum est pour $t_1 = 0$.

Le tracé des courbes comparables pour une valeur non nulle de t_2 ne modifie ni l'allure générale des courbes, ni la conclusion.

L'optimum est donc obtenu pour t_1/T et t_2/T voisins de zéro.

En pratique, il faut donc que l'espacement entre les blocs soit nul, ou que les blocs soient très grands devant l'espacement.

c) Dessin du transducteur

Pour obtenir l'effet d'unidirectionnalité, il est nécessaire que le barycentre des blocs alimentés par la phase 1 et le barycentre des blocs alimentés par la phase 2 soient déphasés dans le temps d'un angle opposé à l'angle de déphasage des alimentations (en pratique 90°).

Le nécessaire décalage entre les blocs en résultant empêche par suite de considérations géométriques l'annulation de

t_1 et t_2. Il est toutefois possible de les minimiser : 1/ en supprimant la grecque de masse, dont les trajets verticaux entre les blocs sont causes d'accroissement de t_1 et t_2 2/ en adoptant une disposition des doigts qui minimise t_1 et t_2. Cette disposition est décrite en figure 5 :

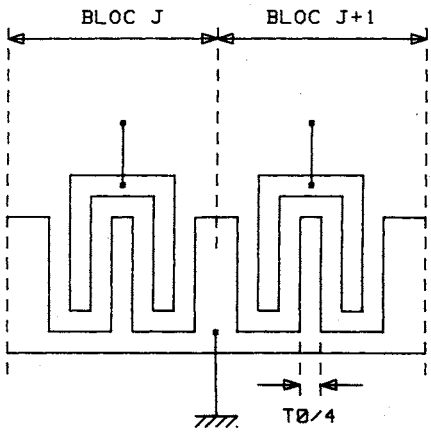


FIGURE 5

Elles est telle que $2t_2 = 3/4To$ et $2t_1 = 1/4\ To$

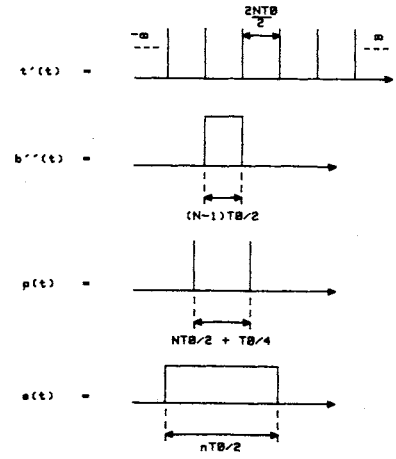
On a vérifié que $1/2\ (2t_2 - 2t_1) = To/4$ apporte bien le déphasage souhaité de 90°.

d) Analyse détaillée du transducteur

La règle de dessin définie pour le transducteur amène une nouvelle analyse du transducteur donnée par :

$$h(t) = (t'(t) * b''(t) * p(t)).e(t)$$

avec les schémas suivants :



où la fonction p(t) est la fonction de déphasage des blocs.

La réponse fréquentielle est alors :

$$H(f) = (T'(f).B''(f).P(f)) * E\ (f)$$

Elle est décrite par le haut du tableau 3

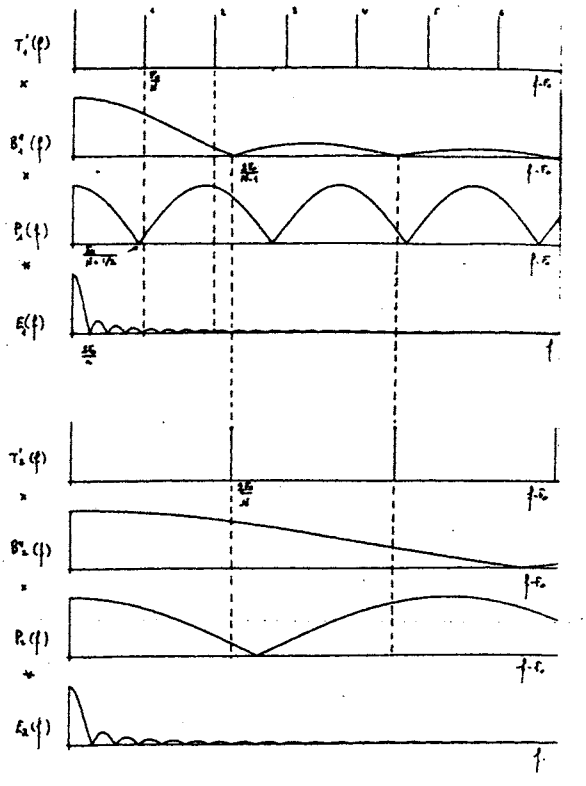


TABLEAU 3

On constate : 1/ que les réponses d'indices pairs de T'(f) ne sont plus exactement annulées par B"(f). Cependant, si N est grand (ie les blocs sont longs), les pics resteront dans des zones de fort affaiblissement de B"(f). Toutefois, on peut prévoir que le pic d'ordre 2, le moins atténué de tous, restera encore apparent. 2/ les zéros apportés par P(f) sont très rapprochés des pics impairs de T'(f).On peut prévoir qu'ils apporteront une atténuation pas parfaite mais tout de même très bonne des pics impairs de T'(f). Le pic 1 en particulier, le moins atténué de tous les impairs, sera certainement plus atténué que le pic 2

e) Analyse globale du filtre

Le filtre est composé de 2 transducteurs. Chaque transducteur a une réponse $H_1(f)$ et $H_2(f)$. Le transducteur 1 est non pondéré. Le transducteur 2 est pondéré. La réponse totale est :

$$H(f) = H_1(f) \cdot H_2(f)$$

On choisit la périodicité des blocs de $H_2(f)$ afin d'obtenir la situation décrite par l'ensemble du tableau 3.

On remarque que pour le transducteur 2, les blocs sont plus petits et donc le rapport espacement sur bloc est plus grand. L'annulation des pics pairs de $T_2'(f)$ par les zéros de $B_2"(f)$ est alors très médiocre. Pour remédier à ce défaut, les pics de $T_2'(f)$ ont été mis en correspondance avec les zéros de $B_1"(f)$. Ainsi tous les pics de $T_2'(f)$ sont parfaitement annulés.

Donc, sur l'ensemble du filtre le seul pic restant médiocrement annulé est le pic 2 de $T_1'(f)$ situé à 2Fo/N

f) Optimisation de l'enveloppe

La première application de l'enveloppe est de fournir la réponse correspondant au gabarit de filtrage souhaité. La définition de l'enveloppe est alors obtenue par des méthodes de synthèses classiques.

Dans le cas du transducteur unidirectionnel à blocs, il est intéressant d'imposer une condition supplémentaire sur la bande atténuée visant à accentuer encore l'atténuation des pics parasites, notamment les pics 1 et 2 de $T_1'(f)$

Cette optimisation n'a abouti que pour le transducteur non pondéré :

$$H_1'(f) = [T_1'(f)\, B_1"(f)\, P_1(f)] * E_1(f)$$

$$H_1'(f) = \int T_1'(f')\, B_1"(f')\, P_1(f') * E_1(f' - f)\, df'$$

$$H_1'(f) = \int \sum_k (f' - Fok/N)\, B_1"(f')\, P_1(f')\, \frac{\sin \pi n/2(f'-f)/Fo}{\pi n/2(f'-f)/Fo}\, df'$$

$$H_1'(f) = \sum_k B_1"(Fok/N)\, P_1(Fok/N)\, \frac{\sin\ n/2\ (k/N - f/Fo)}{n/2\ (k/N - f/Fo)}$$

Afin d'annuler les pics de $T_1'(f)$, on impose $H_1'(f) =$ pour

$f = mFo/N \qquad m = 1,2,\ldots$

$$\sum_k B_1"(Fok/N)\, P_1(Fok/N)\, \frac{\sin \pi n/2\ (k - m)/N}{\pi n/2\ (k - m)/N} = 0$$

Les solutions n = 2 α N (α = entier) satisfait cette équation pour toutes les valeurs de m.

III/ Application

L'étude décrite a défini un dessin du transducteur et induit les règles suivantes :

$N_1 = 4\beta + 3$
$N_2 = 2\beta + 1$
$n_1 = 2\alpha_1 N_1$
$n_2 = 2\alpha_2 N_2$
β = max

En outre :

est borné supérieurement par le fait que le nombre de blocs par transducteur doit être suffisant pour assuré l'unidirectionnalité. Ce nombre a été fixé arbitrairement à 4.

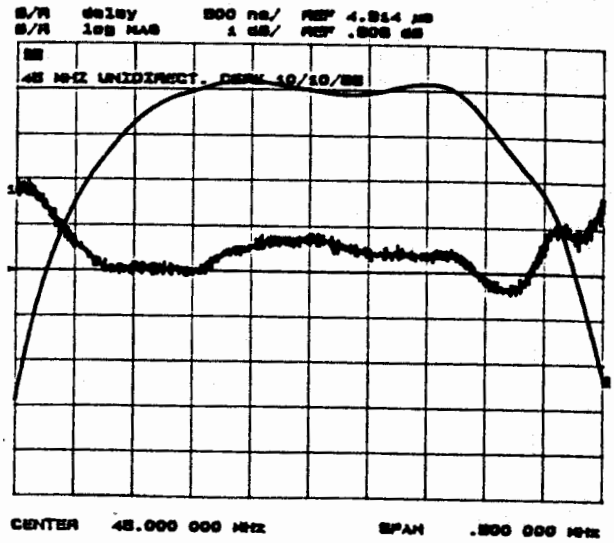
n1 et n2 sont définis à quelques unités près par la synthèse du filtre.

La définition finale de β a été obtenue grâce à une optimisation par simulation.

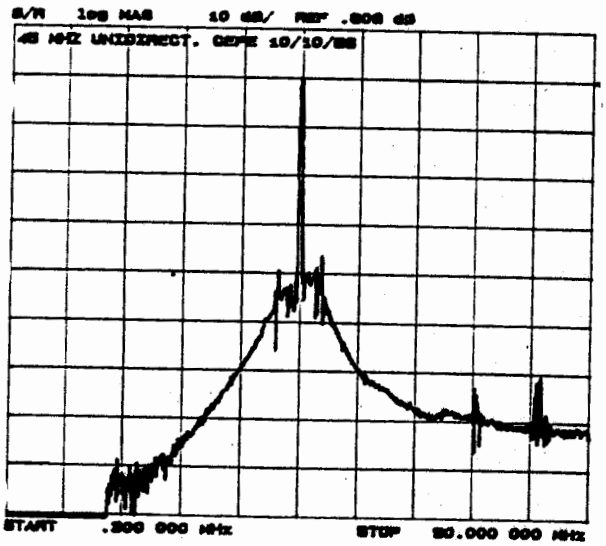
Les valeurs les plus satisfaisantes ont été :

$$\beta = 6 \quad \alpha_1 = 4 \quad \alpha_2 = 27$$

La réponse du filtre réalisé est donné par les courbes 2 et 3.



COURBE 2



COURBE 3

On vérifie bien une bonne réjection (de l'ordre de 40 dB) de toutes les raies parasites. Celle qui reste malgré tout la plus marquée est la raie 2 de $T_1'(f)$ située à $2Fo/N_1$ à 3,5 MHz.

C'est ce que prévoyait le raisonnement.

IV/ Conclusion

L'étude menée a eu pour but de rechercher les performances maximales pour des filtres à onde de surface faibles pertes applicables au filtrage FI dans le radiotéléphone numérique pan- européen.

Le choix du design a été orienté vers de technologies rapides à mettre en oeuvre et minimales en nombre d'opérations de fabrication afin d'être compatible en délai et en coût avec l'objectif visé.

En fonction de ces critères, la disposition à transducteur triphasé et un seul niveau de masquage a été retenue.

Les travaux menés ont montré qu'une optimisation de cette technologie permet de réduire les réponses parasites associées à cette disposition et d'obtenir des niveaux de réjection supérieurs à 40 dB.

L'optimisation de la fonction d'enveloppe du transducteur non pondéré et l'étude de blocs fractionnés permettraient d'améliorer encore ce résultat dans l'avenir.

BIBLIOGRAPHIE

[1] Low loss SAW device employing single stage fabrication
LEWIS
1983 Ultrasonics Symposium Proc. p 104-108

[2] Low loss SAW filter using internal reflection types of new single phase unidirectionnel transducer
YAMANOUCHI and Al
1984 Ultrasonics Symposium Proc. p 68-71

[3] High performance SAW filters with several new technologies for cellular radio
HIKITA and Al
1984 Ultrasonics Symposium Proc. p 82-92

[4] Low loss SAW band pass filter
DEFRANOULD
1981 Ultrasonics Symposium Proc. p 17-20

[5] Low insertion loss acoustic surface wawe filter using group type unidirectionnel interdigital transducers
YAMANOUCHI
1975 Ultrasonics Symposium Proc. p 317-321

SAW FILTERS FOR DIGITAL CELLULAR RADIO APPLICATIONS *

J.M.HODE, P.DUFILIE, J.DESBOIS, H.GAUTIER

THOMSON-SINTRA ASM / D.T.A.S
399, Route des Crêtes - Parc de Sophia-Antipolis
06561 VALBONNE CEDEX FRANCE

ABSTRACT

Two Surface Acoustic Wave (SAW) filters have been developed for both the front end and the IF sections of the future European digital cellular mobile telephone system.

The first one operates at 71 MHz with 200 kHz bandwidth; it exhibits insertion loss as low as 2.6 dB, 45 dB close-in rejection and 70 dB far-out rejection. This filter provides both channel filtering and blocking signal rejection.

The second one operates in the transmit or receive band (about 900 MHz). Very high performances were achieved : -1 dB bandwidth greater than 25 MHz with less than 0.5 dB and 20 ns peak to peak in-band ripple, close-in and far-out rejection greater than 50 dB, shape factor better than 3:1 and insertion loss as low as 7 dB.

Theoretical predictions for both filters and other results are also presented, showing the capability of the SAW technology to achieve high performance low loss RF and IF filters.

INTRODUCTION

Great efforts are being made all over Europe to develop a Pan-European mobile communication system. Many countries are involved in this ambitious project, such as the Scandinavian countries, Germany, U.K, Italy and France.

The architecture of this future cellular mobile radio system is rather different from the present 400 MHz and 900 MHz all-analog systems. Digital processing provides better S/N ratio (about 13 dB) and the TDMA technique allows for time separation of transmitted and received signals with the following consequences :

* suppression of the duplexer which was one of the largest circuits in the system

* the RF filter specifications can be relaxed

However such a complex mobile radio system requires various high performance technologies regarding both active and passive elements.

CIRCUIT ARCHITECTURE

A possible architecture of the RF and IF sections of both mobile and base station systems could be represented by the block diagram presented in fig. 1. The transmit and receive bands for the mobile are respectively 890 - 915 MHz and 935 - 960 MHz. Four RF filters can be identified, two in each section (transmission and reception). Insertion loss must be less than 2.5 dB for the two receive filters, as low as 1.5 dB for the transmit front end filter, and less than 8 dB for the transmit preamplifier filter.

For technological reasons the single IF scheme is the most intensively investigated, with a probable IF close to 71 MHz. The IF section processes a 270 kBit TDMA channel (containing 8 time-multiplexed voice channels) with a 90 dB sidelobe GMSK modulation. The corresponding bandwidth is 200 kHz allowing 1000 voice channels to be allocated.

The aim of the IF filter is to protect the IF section against both blocking signals and adjacent TDMA channels. Since channels n±1 and n±2 are not used while channel n is used, the ultimate rejection requirement is specified from IF ± 800 kHz (channel n±3) to IF ± 25 MHz (limit of the receiving band). Low insertion loss is also required for both power consumption, circuit size and cost considerations.

Surface Acoustic Wave (SAW) filters are known for miniaturization, mass production capability and high reproducibility. These points make the SAW technology a very attractive solution for both RF and IF filtering in mobile telephone systems.

This paper presents a 71 MHz and a 900 MHz SAW filter which were developed for digital radio applications.

LOW LOSS TECHNIQUES FOR SAW TECHNOLOGY

Generally conventional SAW filters show fairly large insertion loss (IL), typically 20 dB, which restrict their applications. Front end or first IF stage applications require lower losses, below 10 dB and even 5 dB.

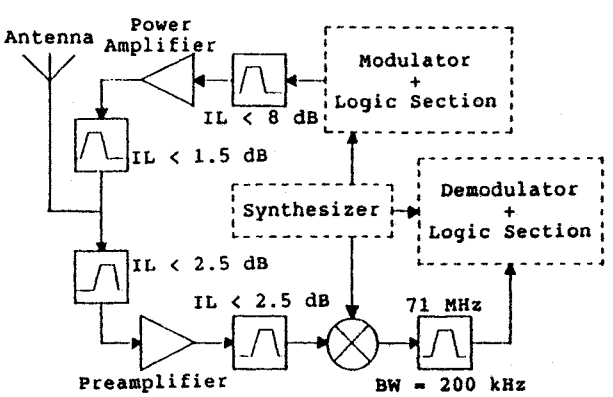


FIGURE 1 : Possible block diagram for Pan-European mobile communication system.
Receive band : 935 - 960 MHz
Transmit band : 890 - 915 MHz

* This work was partially sponsored by D.A.I.I (France)

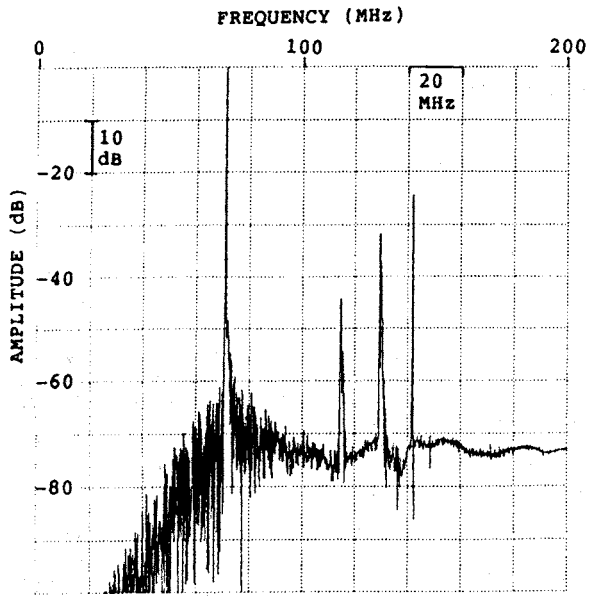


FIGURE 2 : Measured frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 100 MHz. Plot scales are 20 MHz/div horizontal and 10 dB/div vertical.

Numerous low loss filter structures have been developed during the past 15 years, such as 3-phase unidirectional transducers (3PUDT) [1,2], single phase unidirectional transducers (SPUDT), including both internal reflection transducers [3] and folded multistrip couplers transducers [4]), repetition structures [5,6] or track coupling via image impedance transducers [6-8].

For lower frequency applications the 3-phase technique is the most accurate for controlling in-band shaping and ripple. However air gap cross-over technology and minimum linewidth restrict the use of this technique to below 300 MHz.

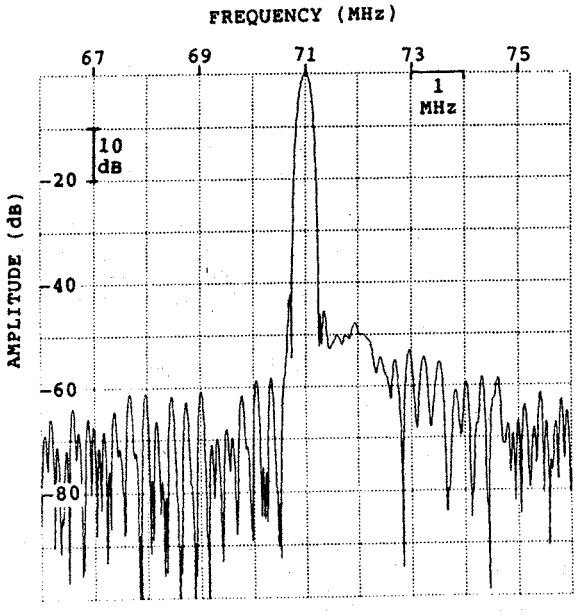


FIGURE 3 : Measured frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 1 MHz/div horizontal and 10 dB/div vertical.

High frequency applications require simple tuning, single phase and single layer technologies. These requirements are achievable at 1 GHz with a standard 1 µm lithography process by using a filter structure [8] derived from the repetition structure proposed by Hikita et al [6,7].

71 MHz IF FILTER

The center frequency of the IF filter was chosen to be 71 MHz. This choice represents a good compromise between the system requirements and the possibilities of the SAW technology.

Because of the narrow bandwidth requirement, quartz was chosen as the substrate material because of its high thermal stability ($3.5\ 10^{-2}$ ppm/°C²) and its low coupling efficiency (k^2 = 0.116 %).

The filter was designed and built using the 3-phase technology for easy in-band shaping and ripple control and for possibility of spurious signals compensation. In order to minimize insertion loss and diffraction distortion, withdrawal weighting was used.

The minimum cavity size required for this filter is 4.5 mm by 4 cm with a height of only 2 mm.

Figs. 2 to 4 give the measured frequency response of the filter. The insertion loss is only 2.7 dB, the close-in rejection is better than 45 dB and the far-out sidelobe level exceeds 70 dB from 0 to 200 MHz. Spurious bands occur at 115, 130 and 142 MHz with levels greater than 40, 30 and 20 dB from the main signal corresponding to SSBW (Surface Skimming Bulk Wave) radiation at 1.6 and 1.8 times f_0 and the 2nd harmonic of the SAW.
Figs. 3 to 4 can be compared with the analysis given in figs. 5 to 6. We can notice that, the agreement between experimental measurements and theoretical calculations is very good for both in-band and stop-band responses.

Group delay variation in the passband of the filter is parabolic in shape and is the result of the tuning. Such a shape is predicted by our simulations and can be approximated by a very simple relation, which is valid for many 3-phase filters :

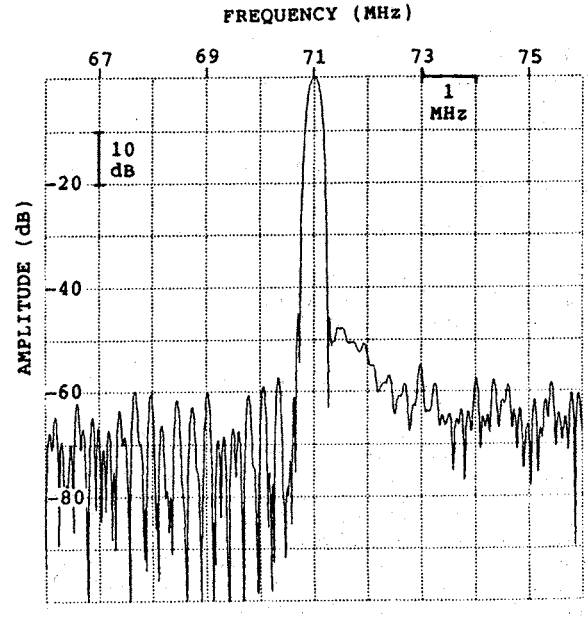


FIGURE 5 : Theoretical frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 1 MHz/div horizontal and 10 dB/div vertical.

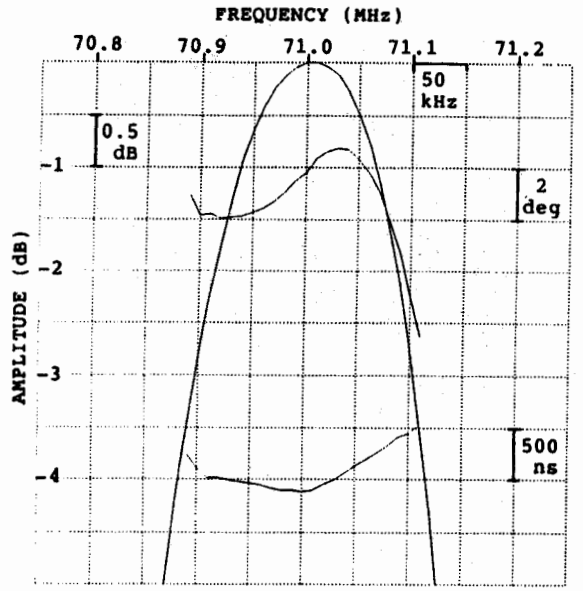


FIGURE 4 : Measured frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 50 kHz/div horizontal and 0.5 dB/div, 2 deg/div and 500 ns/div vertical.

$$\Delta\tau * B_{-3} \approx \frac{1}{10}\left(\frac{f - f_0}{0.5\, B_{-3}}\right)^2$$

This variation can be compensated for and reduced to ± 100 ns in the 3 dB band as shown in the simulation of fig. 7.

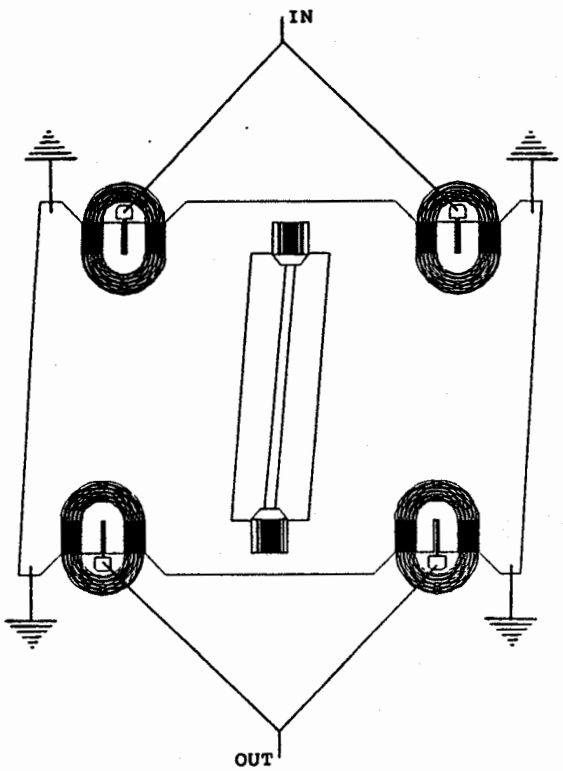


FIGURE 8 : 900 MHz RF filter configuration.

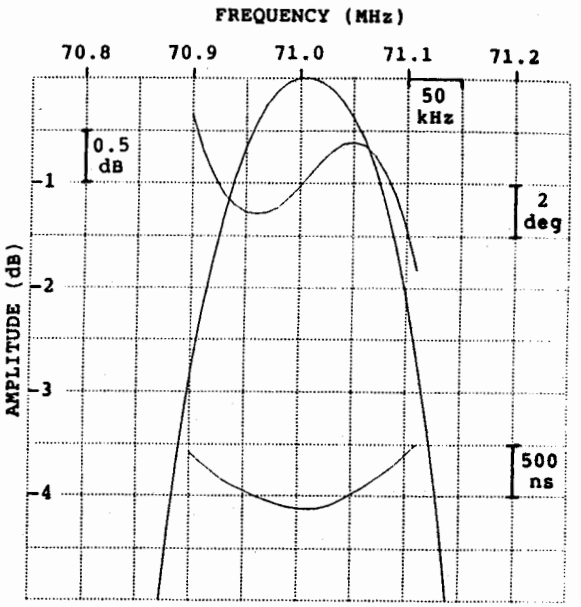


FIGURE 6 : Theoretical frequency response of the 71 MHz IF filter. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 50 kHz/div horizontal and 0.5 dB/div, 2 deg/div and 500 ns/div vertical.

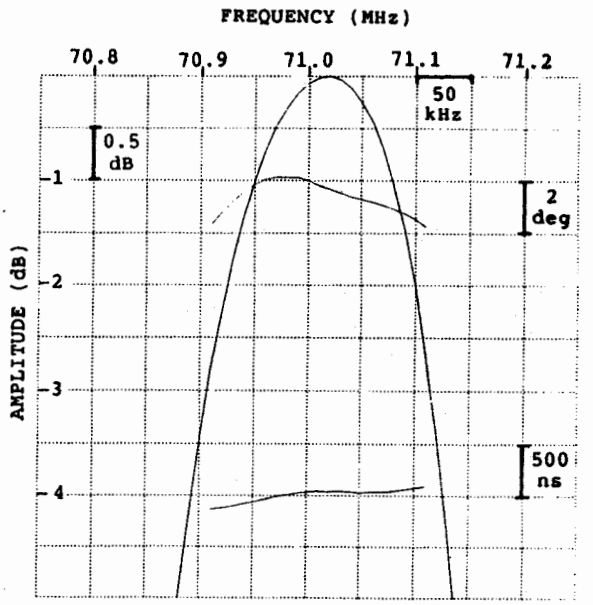


FIGURE 7 : Theoretical frequency response of the 71 MHz IF filter after phase error compensation. Amplitude reference is 2.7 dB. Plot center at 71 MHz. Plot scales are 50 kHz/div horizontal and 0.5 dB/div, 2 deg/div and 500 ns/div vertical.

902 MHz RF FILTER

We have already reported [8] an improved structure for low loss, sharp cutoff, high rejection filters which can operate up to 1.5 GHz with standard optical processing techniques. This structure, derived from the repetition structure proposed by Hikita et al [6,7], combines several low loss techniques to achieve a single phase, single layer and simple tuning low loss filter.

The basic configuration for this structure is shown in fig. 8. It consists of a pair of unidirectional

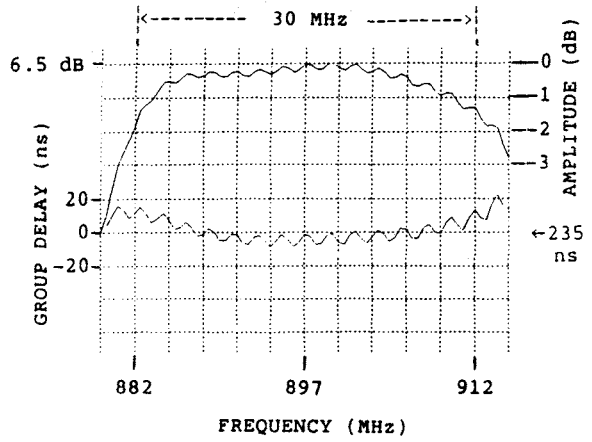


FIGURE 9 : Theoretical frequency response of the 900 MHz RF filter. Amplitude and group delay reference are 6.5 dB and 235 ns. Plot center at 897 MHz. Plot scales are 3 MHz/div horizontal and 1 dB/div and 20 ns/div vertical.

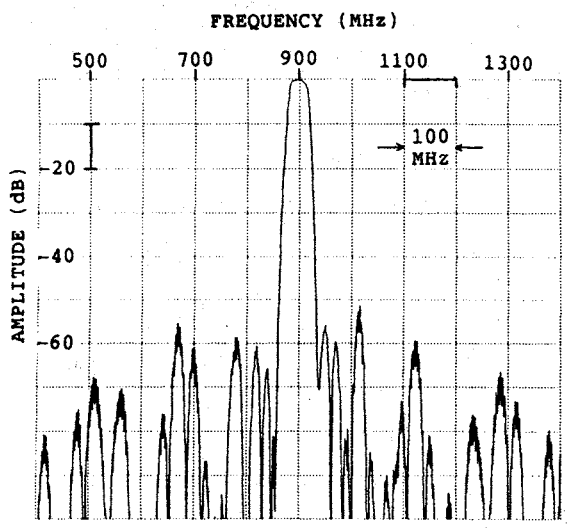


FIGURE 10 : Theoretical frequency response of the 900 MHz RF filter. Amplitude reference is 6.5 dB. Plot center at 900 MHz. Plot scales are 10 dB/div vertical and 100 MHz/div horizontal.

transducers (UDT) for both electrical input and output. These transducers are comprised of a bidirectional transducer inside a 3 dB multistrip coupler (MSC) which is folded in the form of a 0 (OMSC-UDT structure). Input and output are acoustically coupled via a pair of electrically coupled bidirectional transducers (TCs). Lossless coupling and passband shaping are obtained by a proper image impedance connection of the TCs as explained in reference [6]. Standard withdrawal weighting on the TCs is sufficient to obtain 50 dB rejection with the input/output transducer responses included as shown on figs. 9 and 10.

The filter was manufactured on 36° YX-$LiTaO_3$ because of its low temperature coefficient (35 ppm/°C) and high coupling coefficient (> 5.6 %). Figs. 11 and 12 show the measurement of its frequency response which appears to be in good agreement with the theoretical predictions shown in figs. 9 and 10.

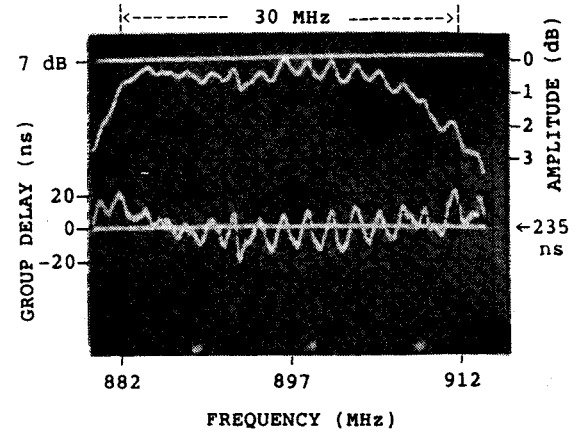


FIGURE 11 : Measured frequency response of the 900 MHz RF filter. Amplitude and group delay reference are 7 dB and 235 ns. Plot center at 897 MHz. Plot scales are 3 MHz/div horizontal and 1 dB/div and 20 ns/div vertical.

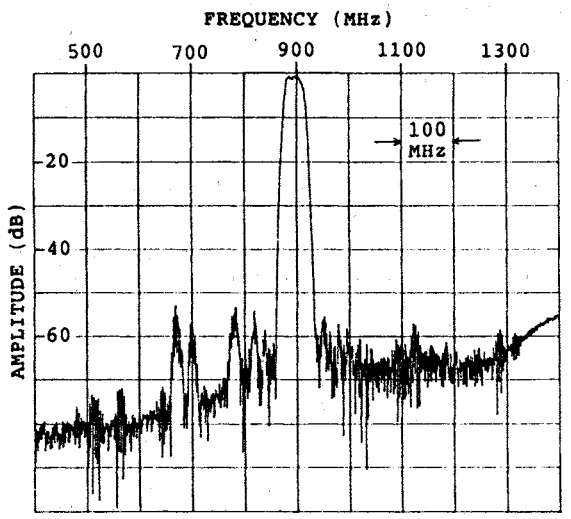


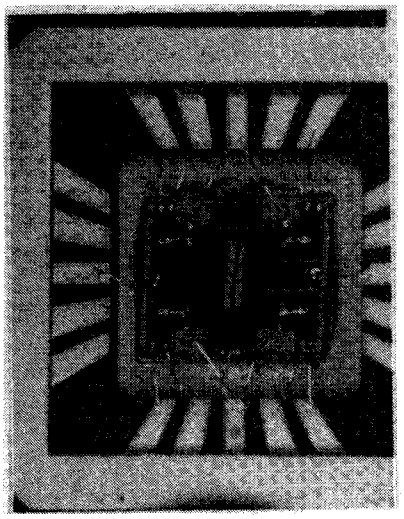
FIGURE 12 : Measured frequency response of the 900 MHz RF filter. Amplitude reference is 7 dB. Plot center at 900 MHz. Plot scales are 100 MHz/div horizontal and 10 dB/div vertical.

High performances are obtained since the achieved insertion loss is 7 dB with a 0.5 dB and 20 ns fast peak to peak amplitude and group delay in-band ripple. Both close-in and far-out rejection are greater than 50 dB over a 1 GHz bandwidth for a shape factor better than 3:1.

Chip size is only 2.2 x 2.2 mm^2 which allows mounting in a TO-5 can. Flat pack packaging affords improved electromagnetic feedthrough rejection and chip carrier mounting has also been investigated (fig. 13) for surface mount applications.

OTHERS LOW LOSS RESULTS

The "image impedance" technique described above and explained in reference [6], allows an achievable fractional bandwidths close to 0.67 times k^2. Using an improved transducer design technique we can vary both effective coupling and stopband width in order to extend the domain of achievable relative bandwidths on a given material.



4 mm

FIGURE 13 : Photograph of the 900 MHz RF filter mounted in a 9 x 9 mm² chip carrier.

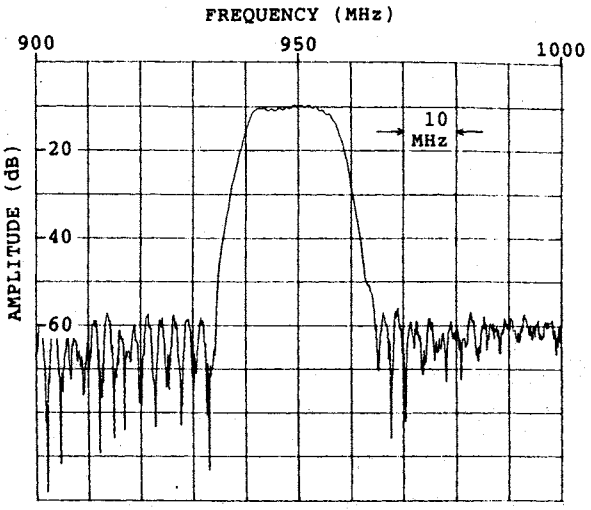


FIGURE 14 : Measured frequency response of the 950 MHz narrow band filter. Amplitude reference is 0 dB. Plot center at 950 MHz. Plot scales are 10 dB/div vertical and 10 MHz/div horizontal.

With such a technique we were able to design a narrow band filter having a relative 3 dB bandwidth of 1.7 %, operating at 950 MHz. Fig.14 shows the measured frequency response of this filter. The in-band and stop-band characteristics are very similar to those of the 900 MHz filter, except for the 10 dB isertion loss caused by a strong diffraction and propagation attenuation.

An improved weighting technique described in ref. [6] was used to design a 2 MHz bandwidth filter operating at 72 MHz with only 3 dB insertion loss (fig. 15 and 16). Proper design of the input/output transducers combined to the weighting technique results in both 70 dB far-out rejection (50 dB close-in with a shape factor of 3:1) and simple tuning (one inductor on each side of the filter). In addition the chip size is only 8.5 x 8.5 mm².

CONCLUSION

Two high performance low loss SAW filters operating at 71 MHz (IF) and 900 MHz (RF) have been developed for the future Pan-European digital cellular radio system.

Theoretical predictions for both filters are in perfect agreement with the experimental measurements.

The main specifications of the mobile radio system are satisfied. The results presented in this paper show the capability of the SAW technology to achieve high performance low loss RF and IF filters.

Future efforts are toward even lower losses and higher frequencies (up to 1.5 GHz).

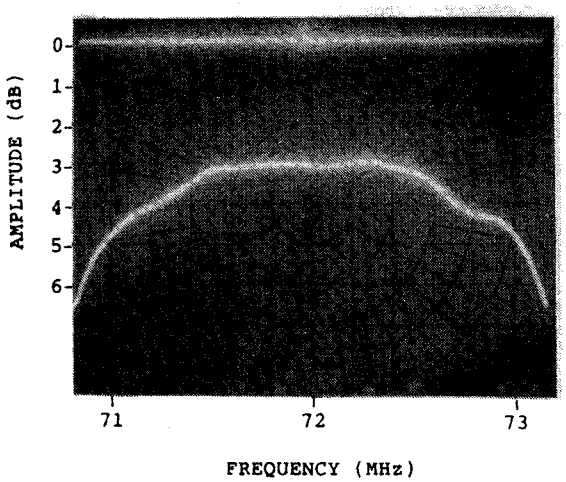


FIGURE 15 : Measured frequency response of the 72 MHz filter. Amplitude reference is 0 dB. Plot center at 72 MHz. Plot scales are 200 kHz/div horizontal and 1 dB/div vertical.

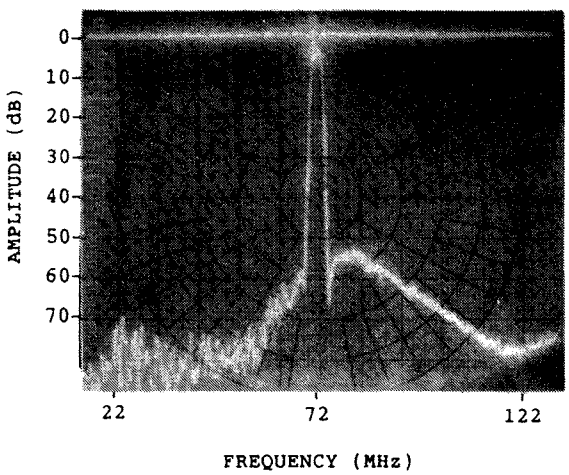


FIGURE 16 : Measured frequency response of the 72 MHz filter. Amplitude reference is 0 dB. Plot center at 72 MHz. Plot scales are 10 MHz/div horizontal and 10 dB/div vertical.

REFERENCES

[1] R.C ROSENFELD et al - "Low loss Unidirectional Acoustic Surface Wave Filters" Proceedings of 28th Annual Symposium on Frequency Control, New Jersey 1974

[2] Ph. DEFRANOULD, J. DESBOIS - "Low Loss SAW Bandpass Filters" 1981 IEEE Ultrasonics Symp. Proc. p. 17

[3] M. LEWIS - "Low Loss SAW Devices Employing Single Stage Fabrication" 1983 IEEE Ultrasonics Symp. Proc. p. 104

[4] F.G. MARSHALL et al - "New Unidirectional Transducer and Broadband Reflector of Acoustic Surface Waves" Elect. Lett., Vol. 7, p. 638

[5] M.LEWIS - "SAW Filter Employing Interdigitated Interdigital Transducers, IIDT" 1982 IEEE Ultrasonics Symp. Proc. p. 12

[6] M. HIKITA et al - " High Performance SAW Filters with Several New Technologies for Cellular Radio" 1984 IEEE Ultrasonics Symp. Proc. p. 82

[7] M. HIKITA et al - "New Low Loss Broadband SAW Filter using Unidirectional IDT's with U-Shaped MSCs" Elect. Lett., Vol. 20, p. 453

[8] P. DUFILIE, J.M. HODE, J. DESBOIS "A Low Loss High Performance Filter Structure" 1988 IEEE Ultrasonics Symp. Proc. (to be published)

A STRATEGY FOR A LOW COST IF FILTER FOR DIGITAL RADIOTELEPHONE

JP AUBRY, JP MICHEL, JL LE CORRE, D BLONDE & P de la FOURNIERE

C.E.P.E. 44 avenue de la Glacière 95100 ARGENTEUIL

SUMMARY

The next generation of digital radiotelephone which is expected to cover the entire EUROPE asks for many up to date technologies.

An intermediate IF filter will be required with the following basic specification :

- high center frequency (between 40 and 75 MHz)
- wide band (from 160 to 300 KHz)
- phase linearity < 5°
- low insertion loss (<6 dB)
- high out of band rejection
- low shape factor
- low ripple
- small size

The competition between SAW and BAW devices is still opened, even if SAW will have some inherent problems with rejection, size and cost.

BAW will be able to compete, but they have to overcome some limitation :

- high frequency fundamental
- piezoelectric substrate not in quartz
- small size
- low cost

The proposed strategy is based on four major points

- four poles on a single blank
- berlinite or lithium tantalate piezoelectric substrate
- automatique global filter adjustement
- low profile package

Two preliminary 4 poles designs have been fabricated and tested either with discrete or coupled resonators on a single wafer. Results meeting the specification are presented.

This is probably the first time that 4 poles single wafer are shown on $LiTaO_3$ or $AlPO_4$, and this is a very promising technology.

CLASSICAL TECHNOLOGY

Even if cost and size considerations are neglected, the IF filter for cellular digital radiotelephone is not an easy filter for most of crystal filters manufacturers.

The reasons are requirements for 1/ a wide relative bandwith 2/ a high center frequency 3/ a high out of band rejection 4/ a very good phase linearity.

1/ A wide relative bandwidth : Depending on the various system specifications, the required relative bandwidth is 0.25 to 0.7 %. This range of bandwidth is theoretically feasible with quartz resonators. But, practically, it will be a complicated structure involving such techniques as Jaumann design with two or more resonators per arm. This design needs complex transformers, is very hard to tune and has a poor rejection for a given number of resonators. In addition, the iterative impedance of the filter remains at a high level and matching the output of the IF mixer requires high quality factor and consequently large and costly inductors.

For wide bandwidth filters (above 0.3%) lithium tantalate is a well fitted piezoelectric crystal. Its high piezoelectric coupling coefficient makes possible a decrease of the iterative impedance compared to quartz (approximative ratio : 60).

So the required shape of the filter can be met with a very simple design like scales with serial only resonators. So it is a very promising material for the present filter. Nevertheless, for the lower bound of the bandwidth range there will be some limitations due to a slightly high insertion loss ($\sim$ 7 dB) caused by the limited Q factor of that material.

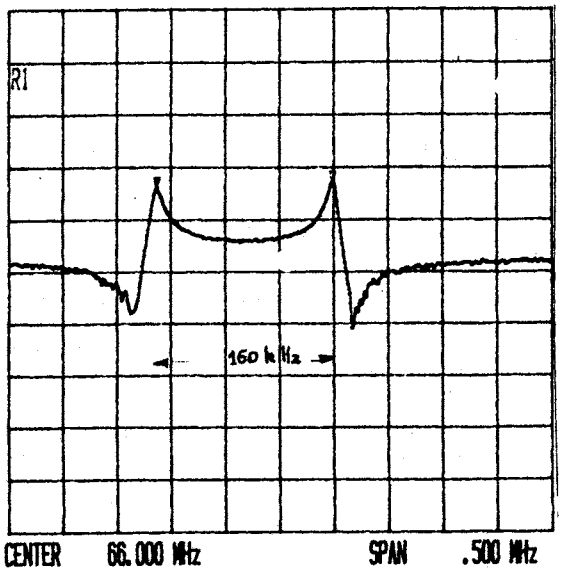


FIG. 1 : a 71 MHz - 100 KHz wide berlinite monolithic cell response

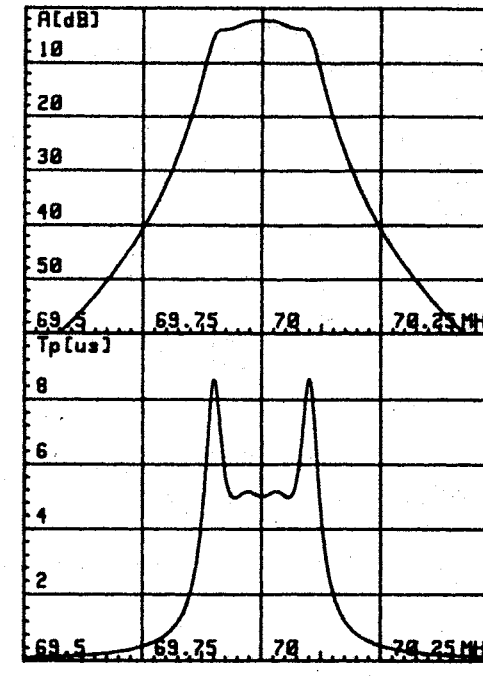
In order to cover the lower bound of the bandwith range we decided to investigate berlinite [1,2]. Berlinite is an aluminum phosphate with the same crystallographic structure as quartz (ie 32). But the piezoelectric coupling coefficient is approximately twice as compared to quartz and consequently achievable iterative impedances for filters are only half that of quartz. In addition monolithic cells up to 75 MHz proved to be feasible in the part (Fig.1). This makes possible a monolithic cells filter which is a very simple design (fig. 2).

So our choice has been lithium tantalate for the upper part of the bandwith range and berlinite for the lower part.

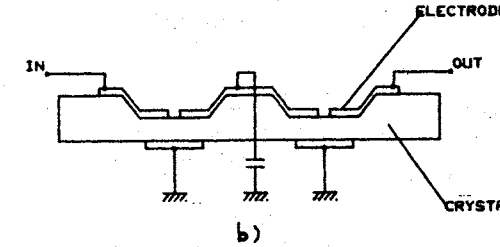
2/ A high center frequency : Either 40 MHz or 70 MHz are too high in frequency to make crystals by conventional mechanical etching. The current state of the art at CEPE would let them made by ionic etching. But we thought it is not convenient for mass production. We will see that there are other solutions.

3/ A high out of band rejection : The IF filter shall reject at a - 60 dB level. The shape factor required corresponds approximately to a four poles filter. But crystal resonators generally exhibit important spurious responses which compromise the high level of rejection. The classical mean to prevent this inconvenience is to increase the number of resonators and to achieve a higher order filter

4/ A good phase linearity : To get a good phase linearity is not properly speaking a great difficulty for crystal filters manufacturers. It acts more on the sensibility of the filter and it needs a more accurate trimming.



a)



b)

FIG. 2 : 4 poles berlinite filter at 70 MHz
a/ simulated response
b/ design

DEDICATED TECHNOLOGY FOR IF FILTER FOR DIGITAL RADIOTELEPHONE

The IF filter for radiotelephone will have to overcome the above difficulties. In addition it will need to be low cost and low volume. In order to achieve this, we made the following design choices :

1/ four poles design
2/ all the resonators on a single blank
3/ a maximally low iterative impedance
4/ a low profile package

1/ Four poles : As above stated, 6 poles would be better for spurii rejection. But the manufacturing costs for a filter are always more or less increasing with an increasing number of poles. Especially the curve tuning cost versus number of poles gets a strong upward bend around five poles. So it appeared necessary to limit at four (may be five) the number of poles. Fitting the required shape with only four poles however requires in certain cases a slightly bridged structure in order to produce attenuation maximum for a convenient closed rejection (fig.3).

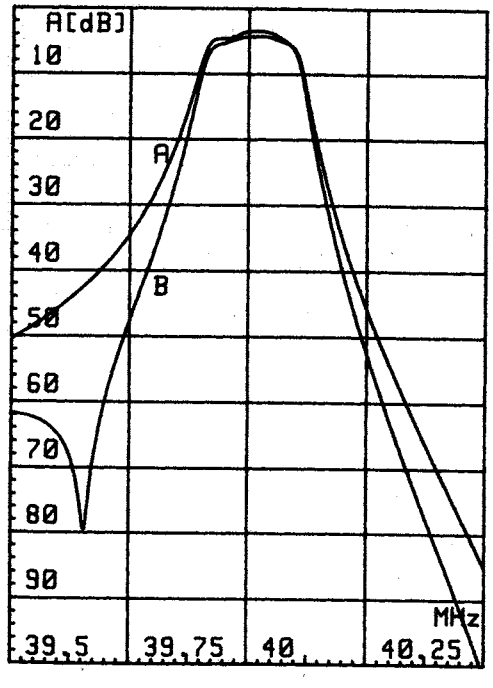


FIG.3 : Normal versus brigded filter design

When making filters this way with no special care as regards to spurious suppression, we got - what was expected - strong enough spurious responses whose rejection were only 30 dB.(fig. 4)

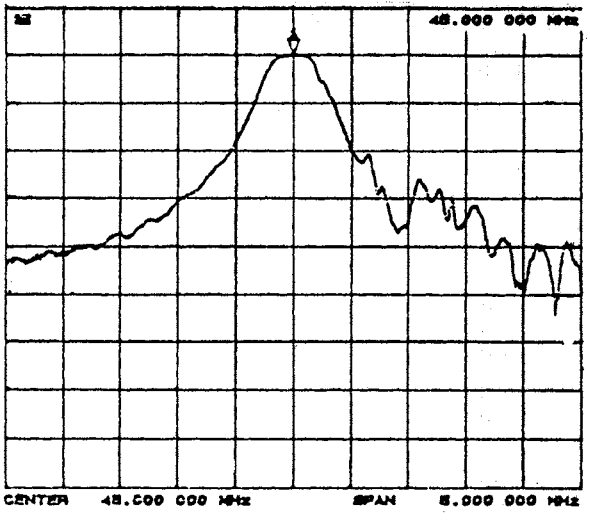


FIG. 4 : LiTa03 filter without spurii suppressor

To overcome this inconvenience we have studied resonators with special electrodes shape. On the resonator, the different mechanical modes of vibration are responsible for the main reponse and for the spurious ones too. One of them exhibits a strong coupling coefficient : it is the main and the useful one. Another specificity is the displacement figure caused by each different mode in the blank. These figures exhibit nodes and maxima in well defined places of the blank. A common method to avoid spurious is to place acoustical absorbant on the maxima of spurious modes provided they are not at the

same place as a main mode maximum. Results are good enough for low piezoelectric coupling materials such as quartz or berlinite. But even for these materials, and for materials with higher coupling coefficients, it can be improved by replacing the acoustical absorbant by an appropriate metallic load. The absorbing effect of spurious is got from mass loading added to a piezoelectric effect, especially strong for high coupling coefficient materials. Good results have been obtained with rejection up to 60 dB on spurii(see fig.10). Ease of implementation is also very good, because the metallic loads are deposited in the same operation as main electrodes.

2/ All the resonators on a single blank : Crystal costs are the main contribution of the cost of a crystal filter. So it is a great deal to reduce the cost of a resonator.

Multipoles resonators have already been shown [3].

The flow chart for a resonator's manufacture is given on figure 5. We made the following analysis : there are only two operations needing an individual action :

Materials

Sawing

Etching

Mounting

Tuning

Enclose

Final check

FIG. 5 : Flow chart for a crystal resonator

1- **Tuning** : as regard to achievable manufacturing process, one cannot hope that the resonator may reach the required precision without any individual tuning.

2- **Control** : each resonator is individually measured.

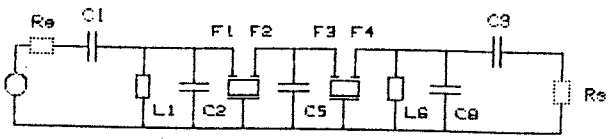
All other operations may find a solution in a collective or semi-collective production.

So the best should be done to individually process only for above both operations and to avoid individual manufacturing for all of the others. One solution for that last point was to study a device with all of the resonators on a same plate.

This disposal actually allows a collective processing on most of the manufacturing operations. In addition, it is cost saving on materials : obviously on mechanics, but also on crystal : the reduced needed quantity of crystal material is a very interesting point for lithium tantalate resonators, because this material is very dependant of the tantalum oxyde cost.

For lithium tantalate filters the design we have studied is a design with four resonators on a same plate. The diameter of the plate is approximatly 5 mm. The resonators are placed at a given distance of each other in order to avoid any unwanted coupling effects on the main response as well as on spurii. Results are very promizing.

For berlinite, we are studying a design based on a double monolithic cell. The frequency of the filter is > 60 MHz, the thickness of the plate is < 20 um. The two cells are placed in a depression made by ion etching in order to achieve the required thickness (see figure 6).



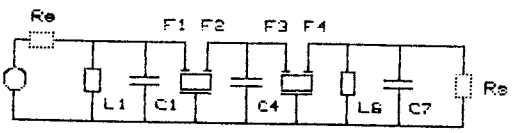


FIG. 6 : Monolithic two cells filter

3/ A maximally low iterative impedance : A high impedance filter may be a disadvantage for the manufacture of the filter but also for the user : the user generally needs an impedance value of 50 to 200 ohms. If the filter exhibits a much higher impedance, matching components between filter and user need to be high Q. The practical limitation for low cost low size inductors are Q = 30. Thus, we choose to use materials for crystal leading to the lowest impedance. The other end of the compromise is given by the filter insertion loss which shall not exceed 6 dB.

4/ A low profile package : The radiotelephone will be a portable device and even a portative one. Size will be a critical parameter. The volume for a classical filter meeting the required shape is approximately 35 x 12 x 10 mm. Our objective for this work was to reduce the volume in order to reach small imensions. This aim is fully consistant with our previously described choices. Firstly, placing the four resonators on the same crystal blank is a great advantage as regard to size reduction. The resonators contribution to the whole size is only one fourth of the usual size. Secondly, the maximally low iterative impedance for the filter leads to a small size for the coils, due to the limited Q value required. Thirdly, the design - either in the scale, either in the monolithic form - avoids the need for such size consuming components as transformer.

But even so, the filter requires a special internal design to allow an easy tuning. This is yet under study. Nevertheless, benches models have yet been produced.

For the benches models we choose to use a T08 package. Nevertheless, the technology we developped may be easily arranged in a SMD package.

RESULTS

Some response curves for encapsulated filters are shown on figure 7 to 10. They are very closed to our target specifications. The pass-band is conform for amplitude as well as for group delay.

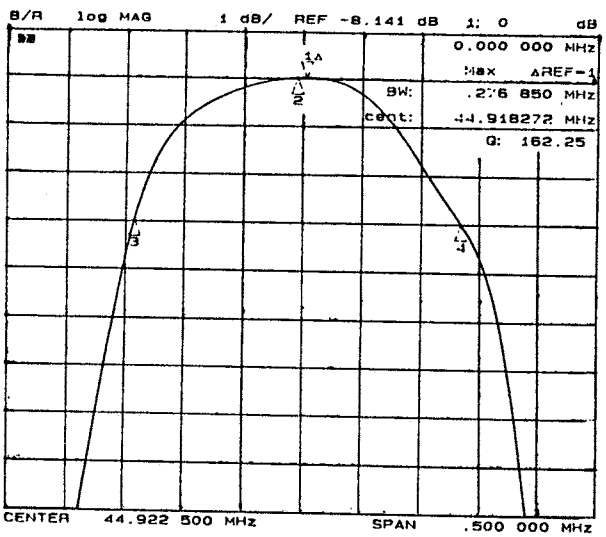


FIG. 7 : Pass band response for a lithium tantalate filter/ 4 poles per plates

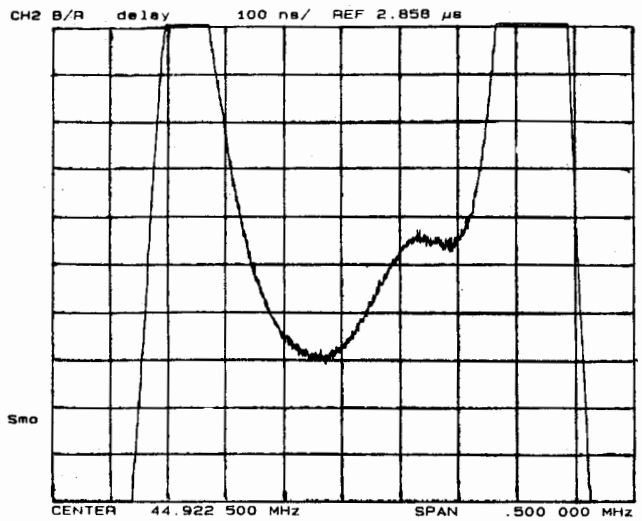


FIG.8 : Phase response for a lithium tantalate filter 4 poles per plate

The - 60dB shape factor is good. Compared to the specification generally required by customers, the response is still slightly too broad at - 20 dB and - 30 dB. This will be improved by a better filter synthesis.

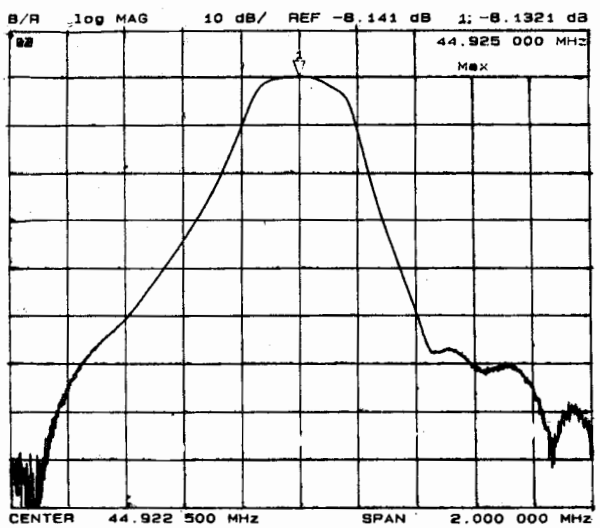


FIG 9 : Closed to band rejection for a lithium tantalate filter - 4 poles per plate

Rejection is - 60 dB. This is a good result because one may fear that the reduced distance between resonators or between resonators and components due to the small volume inside the package may have lead to a poor ultimate rejection.

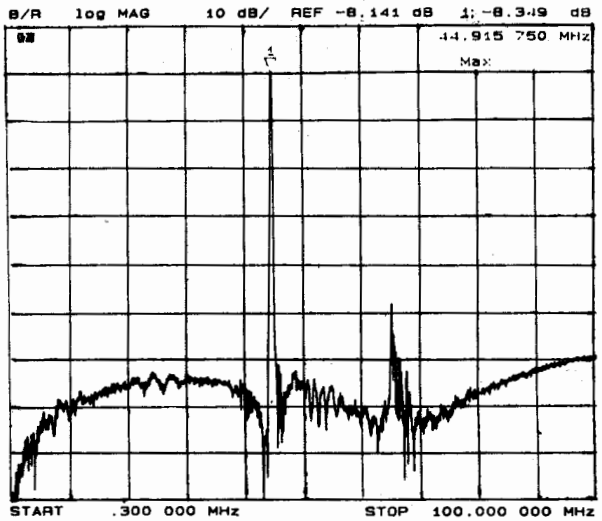


FIG.10 : Out of band rejection for a lithium tantalate filter / 4 poles plate

Spurious responses have been drastically reduced in amplitude as well as in number. Only one response remains at a level of - 45 dB at a distance of 20 MHz frome the center frequency. This response corresponds to a longitudinal propagation mode which is well known on lithium tantalate. We continue studies to reduce this response by modifying the physical arrangement of the resonators on the crystal blank. Compares the resulting filter to a classical equivalent filter.

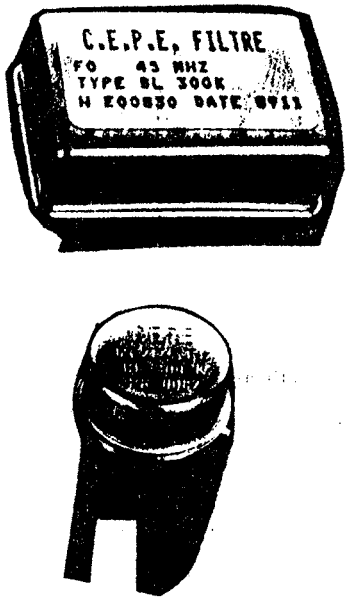


FIG. 11 : Compared dimension classical filter versus new filter

INDUSTRIALISATION

We have shown that numerous choices have been made in order to make the filter easily produceable at a low cost without jeopardizing the small overall size. We are also doing work on the manufacturing process in order to lead to reduced manufacturing costs.

The main lines for that work are :

1- materials saving, especially crystal saving by improved row material processing

2- etching automatic control

3- internal assembly design and mounting technology

4- automatic metallization of the resonators and tuning of the filter

Package is not yet concerned by this work because we don't know if customers will require a SMD package or a plug-in package.

COMPETITIVE TECHNOLOGIES

We have been led to the conclusion that crystal filter is a well fitted technology for the digital radiotelephone IF filter. But we needed before to confirm this opinion by looking at potential competitors, for example SAW filter. In order to make up our mind we have made a low loss SAW filter with such a design we have condidered as cost saving. This realization has been described in a previous paper during this conference [4].

The results are quite interesting : insertion loss is low enough, amplitude and group delay in the pass band are very good. Nevertheless it seems that rejection will hardly reach lower values than 50 dB which is, according to to-day customer specifications, not yet enough. Another main inconvenient is size, which is about 30 mm long.

Manufacturing cost is quite the same as for crystal filters mainly due to the need of a large substrate and of a 3 phases signal generation to achieve a low loss effect in the SAW filter.

CONCLUSION

After having compared the two main technologies seen to meet the specification for an IF filter for digital radiotelephone, we made our mind that crystal filter was a well fitted technology. For that purpose we have made choices to improve cost and size for the filter and to reduce manufacturing costs.

The results are benches models which are very closed to the required electrical specification and target size.

Work is continuing to develop the options above described in order to improve manufacturing process.

BIBLIOGRAPHY

[1] - A new Approach to Berlinite Crystal Growth
E. PHILIPPOT and Al
2nd EFTF 1988 - P. 843

[2] - Evaluation and Application of high purity Berlinite Crystals
J. DETAINT and Al
2nd EFTF 1988 - P. 873

[3] - An LCC Monolithic crystal filter
R.C SMYTHE
2nd EFTF 1988 - P. 637

[4] - Filtre RF faibles pertes à ondes élastiques de surface
JP. MICHEL
These Proceedings

OBSERVATION OF PIEZOELECTRIC ENVELOPE SOLITONS GENERATED IN THE BULK ACOUSTIC WAVE RADIATION FROM INTERDIGITAL TRANSDUCERS DEPOSITED ON QUARTZ PLATES

M. Planat

Laboratoire de Physique et Métrologie des Oscillateurs du C.N.R.S.
associé à l'Université de Franche-Comté-Besançon
32, avenue de l'Observatoire - 25000 Besançon - France

Abstract

Solitons can be roughly thought as stable propagating wave packets resulting from a balance between nonlinear distortion and dispersive spreading. They have been observed recently in numerous fields like hydrodynamics, electrical transmission lines, plasma physics and optical fibers, but until now attempt to excite solitons in solid state acoustics have failed. In a previous work we showed that the bulk acoustic wave radiation from interdigital transducers deposited on a quartz crystal is angularly focused and that the direction of emission is scanned by the applied excitation frequency [1]. Using pulse excitation, it is found that wave packets propagate dispersively under the transducer lattice. On the other hand it is shown that the wave propagation is also highly nonlinear. Pulse instability, pulse narrowing, pulse splitting and dispersion compensation are demonstrated for 1 to 15 V amplitude and 25 ns to 200 ns duration pulses from a transducer with a 17.2 µm spatial period on a Y cut - Z propagation quartz substrate. Main features of experiments are interpreted by establishing the nonlinear Schrödinger equation which governs the pulse envelope function for the mechanical displacement. A reasonable agreement with experiments is found.

Introduction

Most physical phenomena in solid state electronics can be explained within the frame of a linear wave propagation theory. Nonlinear effects are generally introduced to explain the sensitivity of devices to temperature, pressure, acceleration, humidity ... Nonlinear effects are also introduced to explain propagation of harmonics of the main signal, amplitude-frequency effect in resonators and intermodulation products in the bandwidth of filters. In all cases nonlinearities are small and the problem can be satisfactory solved by a perturbation approach [1].

In this paper we will study a novel effect which cannot be explained with the perturbation approach because involved nonlinearities are strong : this is the process of nonlinear dispersion and of solitons. Solitons can be roughly thought as stable propagating wavepackets resulting from a balance between the broadening due to dispersion and the distortion due to nonlinearities. They have been already introduced in numerous fields like hydrodynamics [2], electrical transmission lines [3,4], plasma physics [5] and optical wave guides [6,7] to explain the observed space-time structures. Our goal is to show that nonlinear dispersion should be the key concept to explain space-time in stabilities in acousto-electronics as well. In this paper observation of modulational instability and soliton propagation of shear horizontal waves in quartz is reported for the first time.

First remarks on the experimental setup

A schematic of the experimental setup is shown on Fig. 1. We used a Y-cut quartz substrate (10 × 9 mm2 and thickness 1 mm). The bottom was unpolished and two interdigital transducers were deposited on the top normally to Z-axis. Each transducer had 100 periods, a 4.5 mm width, a spatial periodicity $\lambda_0/2$ = 17.2 µm and 800 Å thick aluminum.

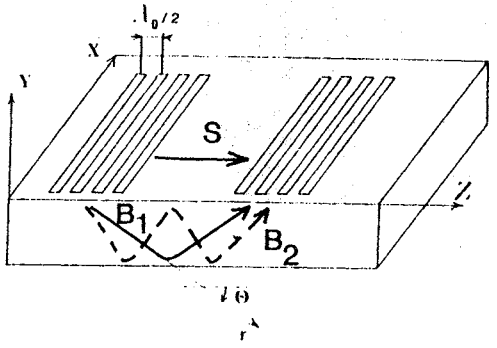


Fig. 1 : Schematic of the experimental setup

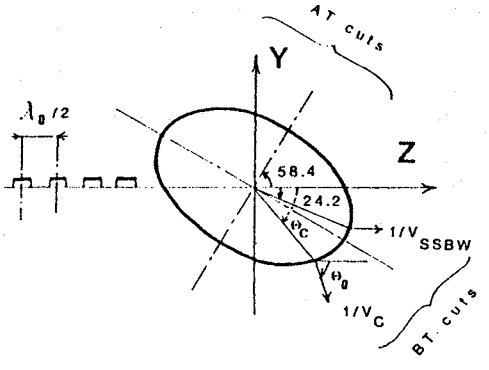


Fig. 2 : Slowness curve for singly rotated Y-cuts of quartz

In the linear regime such a cut supports only shear horizontal motion. Its slowness curve is shown on Fig. 2.

Since the excitation transducer is very long, propagation takes place with the wavenumber k along the direction of constructive interference θ_c given by the relation

$$k_Z = 2\pi/\lambda_o = k\cos\theta_c \qquad (1)$$

In other words the propagation angle θ_c can be scanned by the applied frequency f according to the relation

$$V_c = f\lambda_o \cos\theta_c \qquad (2)$$

The angle θ_0 for the propagation of energy is normal to the slowness curve (Fig. 2) and consequently there will be a cut-off in the device response when θ_o = 0. This angle corresponds to the propagation of the so-called surface skimming bulk wave (SSBW). In our device this happens for θ_c = 24.2° and the corresponding cut-off frequency is f_o = 126 MHz.

The radiation efficiency for the shear wave in a Y-Z cut quartz half space was computed at the ray approximation [8] (Fig. 3). At the cut-off frequency the main lobe is broad and parallel to the surface ; with increasing applied frequency the lobe rotates and gets thinner.

The lobe broadening due to the finite length ΔL of the transducer can be estimated from the uncertainty relation

$$\Delta k \, \Delta L = 2\pi \tag{3}$$

In our device, this leads to an uncertainty $\Delta f = V_c/\Delta L$ in frequency of the order 1 MHz. Broadening and secondary lobes due to the finite length of transducer are then to be expected for times larger than $\Delta t = 1/\Delta f = 1$ µs.

To avoid such modulations we will perform pulse experiments in a range of time scales between 10 ns and 100 ns. This means that these input pulses will propagate under a transducer which can be assumed to be infinite. To explain experiments both dispersion due to the strip lattice (transducer) and nonlinearity have to be taken into account.

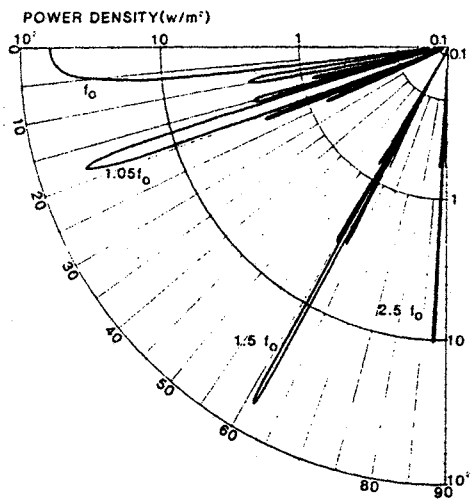


Fig. 3 : Power distribution for a 50-electrode-pair IDT on a Y-cut Z-propagation quartz crystal
f_0 = 126 187 MHz, input voltage ϕ_0 = 1 V

Derivation of dispersion relation and of the linear propagation equation for the pulse

We start with the equation for the slowness curve :

$$\rho\omega^2 = C_{55} k_Z^2 + C_{44} k_Y^2 + 2 C_{45} k_Y k_Z \tag{4}$$

with $C_{44} = 4 \times 10^{10}$ N/m2 ; $C_{55} = 5.8 \times 10^{10}$ N/m2 and $C_{45} = -1.8 \times 10^{10}$ N/m² for the Y-Z cut substrate.

Here $k_Z = 2\pi/\lambda_0$ can be considered constant in the range of time scales we will experiment, i.e. $\Delta t \ll 1$ µs. Introducing the wavenumber $k = (k_Y^2 + k_Y^2)^{1/2}$ along the propagation direction r (Fig. 1), the dispersion relation $\omega = W(k)$ is :

$$\rho\omega^2 = C_{55} k_Z^2 + C_{44}(k^2 - k_Z^2) + 2C_{45} k_Z (k^2 - k_Z^2)^{1/2} \tag{5}$$

It is plotted on Fig. 4. Note that the dispersion is maximum at the SSBW frequency (zero group velocity) and decreases gradually when the frequency is increased.

For a modulated input pulse centered on the wavenumber k_0, corresponding to the frequency $\omega_0 = W(k_0)$ (Fig. 5), we can expand the frequency in Taylor series [9]

$$\begin{aligned} \kappa &= k - k_0 \\ \omega &= \omega_0 + \kappa \omega'_0 + \frac{1}{2} \kappa^2 \omega''_0 \end{aligned} \tag{6}$$

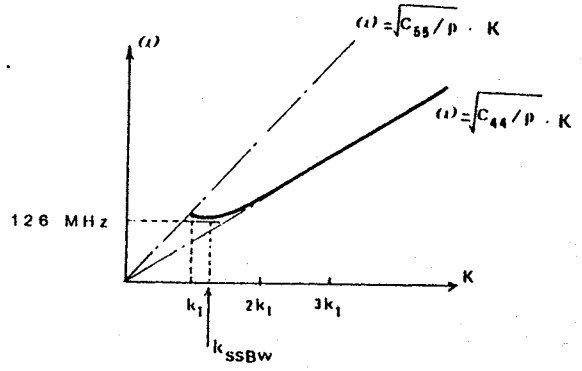


Fig. 4 : Dispersion curve for waves under a very long transducer on a Y-Z cut quartz substrate

where ω'_0 and ω''_0 are the group velocity and the concavity of the dispersion curve evaluated at the carrier wavenumber k_0.

The waveform $\Psi(r,t)$ is obtained by adding all monochromatic waves of wavenumber k and frequency $\omega = W(k)$ with the integral :

$$\Psi(r,t) = \int F(k) \exp\{i(kr - W(k)t)\} dk \tag{7}$$

Extracting the carrier, we obtain from (6)-(7) :

$$\Psi(r,t) = \phi(r,t) \exp\{i(k_0 r - \omega_0 t)\} \tag{8}$$

where $\phi(r,t)$ is the envelope of the pulse :

$$\phi = \int F(k_0 + \chi) \exp\left\{iKr - i(\kappa\omega'_0 + \frac{1}{2} K^2 \omega''_0)t\right\} dK \tag{9}$$

The interesting point of the analysis is that the envelope function obeys a Schrödinger type equation :

$$i(\phi_t + \omega'_0 \phi_r) + \frac{1}{2} \omega''_0 \phi_{rr} = 0 \tag{10}$$

In the frame of the pulse eq. (10) is the conventional Schrödinger equation :

$$i\phi_t + \frac{1}{2} \omega''_0 \phi_{rr} = 0 \quad \text{where} \quad r = r - \omega'_0 t \tag{11}$$

Solutions of Eq. (11) are well known in quantum mechanics. During the propagation the pulse will be broadened and its frequency will be modulated. This is illustrated on Fig. 5.

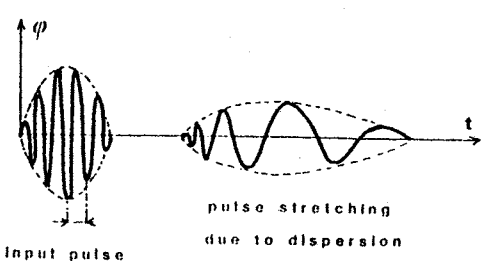


Fig. 5 : Propagation of a dispersion pulse

Nonlinear effects and pulse envelope for nonlinear wave propagation

An estimate of nonlinear propagation effects has been obtained from amplitude frequency measurements in resonators. For a travelling plane wave of amplitude a, the frequency shift is found to be [1]

$$\Delta\omega = k_o \Delta V = q a^2 \quad \text{where} \quad q = \frac{1}{16}\omega_o k_o^2 \frac{C^{(4)}}{C^{(2)}} \qquad (12)$$

where $C^{(2)}$ and $C^{(4)}$ are the effective second order and fourth order elastic constants. In our device the wave propagates in singly rotated orientations (rotation around X-axis) and third order constants vanishes by symmetry. The nonlinear coefficient $C = C^{(4)}/12C^{(2)}$ is known to have an order of magnitude $C \simeq 10$ and is < 0 for AT cuts and > 0 for BT cuts [1, 10]. Referring to Fig. 2 it can be seen that orientations around the BT cut will be explored and then C is negative.

From the new dispersion relation

$$\omega = W(k) - q a^2 \qquad (13)$$

we can derive the pulse envelope for nonlinear wave propagation as follows [9]:

$$i(\phi_t + \omega'_o \phi_r) + \frac{1}{2}\omega''_o \phi_{rr} + q|\phi|^2\phi = 0 \qquad (14)$$

Solving Eq. (14) is a considerable task [11, 12] and only the nature of modulations predicted by Eq. (14) will be studied.

First case : $q\,\omega''_o < 0$

According to Witham's modulation theory [9], there are two real characteristics in this case. The system remains hyperbolic but the input pulse splits into two stable parts of two different group velocities. First experiments for this case were obtained recently in fiber optics by D. Krökel [13]. It has been shown that an input short pulse splits into two dark pulses (Fig. 6a).

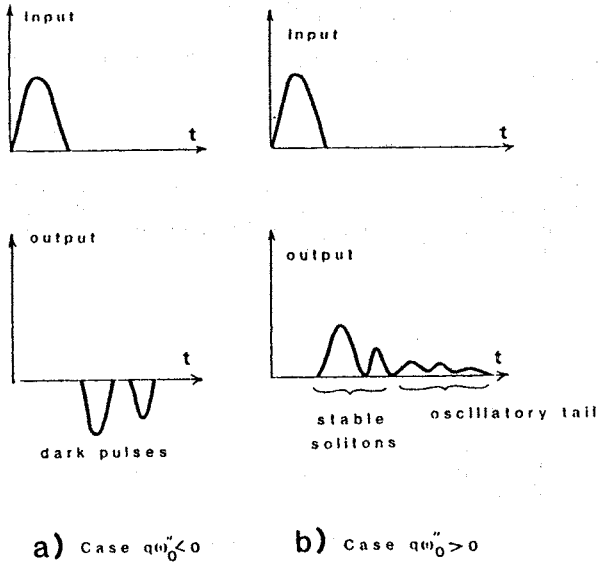


Fig. 6 : Schematic of solutions of the nonlinear Schrödinger equation

Second case : $q\,\omega''_o > 0$

There are two complex characteristics in this case. The system becomes elliptic and modulations are unstable. An input pulse generally splits in many stable parts called solitons and an oscillatory dispersive tail (Fig. 6b). Such behavior has been predicted in hydrodynamics as soon as 1965 by Sir J. Lighthill [14] and since then has been observed for water waves [2], electrical pulses [3,4], plasmas [5] and optical fibers [6,7].

For pulses in quartz BT cuts, we have $q > 0$ as just shown and $\omega''_o > 0$ from Fig. 4 ; then we will be restricted to this case.

An instability parameter and the fundamental soliton

For practical purposes two simple quantitative concepts to study the instability of systems ruled by the nonlinear Schrödinger equation can be settled.

First the modulational instability may be analyzed by assuming a solution with a uniform spatial state of amplitude a plus an infinitesimal propagating perturbation of frequency Ω and wavenumber K [15]. Substituting this form for the solution into (14) yields the following dispersion relation :

$$\Omega = K\left|\frac{1}{4}\omega''^2_o K^2 - q\,\omega''_o a^2\right|^{1/2} \qquad (15)$$

For $q\omega''_o a^2 > \omega''^2_o K^2/4 > 0$ the frequency Ω will become imaginary and the perturbation in the amplitude will grow exponentially in time.

The maximum growth rate occurs when $K = K_{max} = [2q a^2/\omega''_o]^{1/2}$ with the value $\Omega = \Omega_{max} = qa^2$. Assuming that the wavepacket propagates approximately with the linear group velocity ω'_o this allows to introduce the distance $x_{n\ell}$ for maximum instability and the corresponding modulating frequency f_m at which the growth rate is maximum [15]. We obtain :

$$x_{n\ell} = \omega'_o/qa^2 \quad \text{and} \quad f_m = \frac{\omega'_o K_{max}}{2\pi} \qquad (16)$$

A second glance to the modulational instability can be given by observing that Eq. (14) admits a family of localized solutions (called fundamental solitons) of the form [6,9]

$$\phi(r,t) = \phi_1 \operatorname{sech}\left(\frac{t - t_o - x/v_g}{t_o}\right) \exp i(rx - st) \qquad (17)$$

where t_0 is the pulse half width, r and s are wavenumber and frequency shifts, $v_j = \omega'_o + \omega''_o r$ is the pulse velocity and ϕ_1 is the pulse amplitude given by

$$\phi_1^2 = \omega''_o/t_o^2 v_g^2 q \qquad (18)$$

A schematic of the shape of fundamental soliton (17) is shown on Fig. 7.

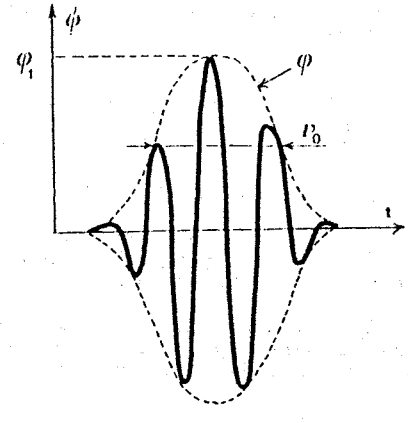


Fig. 7 : Shape of the fundamental soliton

Other effects due to modulational instability

A whole description of physical effects related to the nonlinear Schrödinger equation cannot be given here. We choosed to illustrate them by two examples borowed to recent experiments.

First, a perspective plot of the temporal pulse shape at various points along an optical fiber for the A = 3 soliton is shown on Fig. 8 [6,16]. The input soliton pulse first narrows, then splits and recurs as a function of propagation distance, as observed.

Second, the modulational instability in an electrical transmission line is shown on Fig. 9 for different values Ω/ω_0 between the envelope soliton frequency and the input frequency [17].

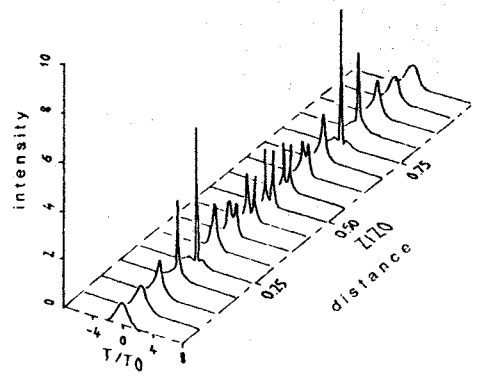


Fig. 8 : Perspective plot of the temporal pulse shape at various points along a fiber for the A = 3 soliton. The intensity variable is defined as $|V(z,t)/A|^2$. The parameters z_0 and t_0 are the spatial and temporal soliton periods (from Stolen, Mollenauer and Tomlinson [16])

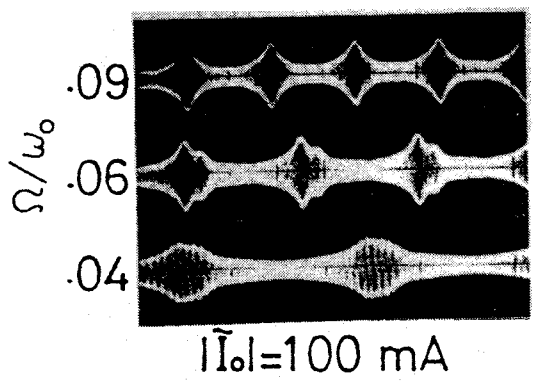


Fig. 9 : Observation of envelope solitons at a fixed point on the line. The ratio Ω/ω_0 is the ratio envelope soliton frequency/input frequency. (from Yagi and Noguchi [17])

Modulational instability and acoustic solitons in the Y-Z cut quartz plate

Theoretical expectations

The previous theory was used to estimate the possibility of observing nonlinear dispersion effects in Y-Z cut quartz plates.

Linear group velocity ω' and concavity of dispersion curves ω'' are shown on Fig. 10. In our setup the distance of maximum instability $x_{n\ell}$ and the amplitude of fundamental soliton are found from (12), (16) and (18)

$$x_{n\ell} = \omega'_0 / C\omega_0 k_0^2 a^2 \quad ; \quad \phi_1^2 = \omega''_0 / C\omega_0 k_0^2 e_R^2 \tau_0^2$$

These parameters have been plotted versus the carrier wavenumber k_0 for an input pulse amplitude a = 100 Å and an input pulse width τ_0 = 100 ns (Fig. 11).

In our device, since $\omega''_0 > 0$ and $q > 0$ then $\omega''_0 q > 0$ and modulations are unstable. Experiments will be performed in the range of carrier wave numbers between $k_1 = 2\pi/\lambda_0$ and $3k_1$, corresponding to frequencies between 100 and 300 MHz. At 126 MHz, $x_{n\ell}$ vanishes and strong modulational instability should be observed. At higher frequencies, $x_{n\ell}$ indreases, then decreases again towards approximately 5 mm at about 300 MHz. At this frequency, $\phi_1 \simeq$ 100 Å (Fig. 11), which is in the range of applied excitations.

Using these estimations we can expect to observe soliton behavior at high frequencies ($\simeq$ 300 MHz), on the second or third bounce (corresponding to sufficient modulational instability) and at sufficiently high input lengths and amplitudes ($\tau_0 \simeq$ 100 ns and $a \simeq$ 100 Å).

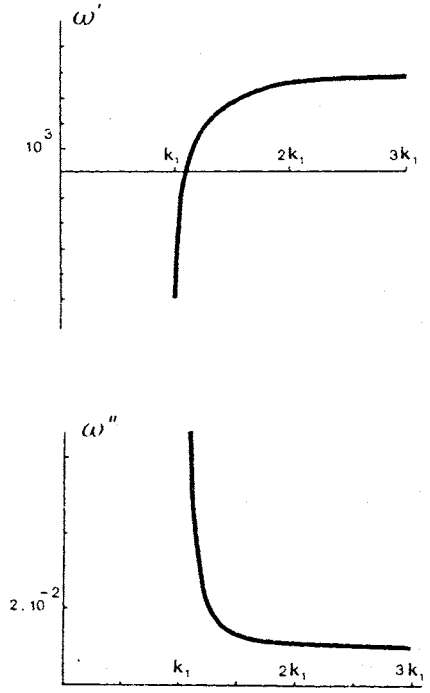


Fig. 10 : Linear group velocity and concacity of the dispersion curve of Fig. 4

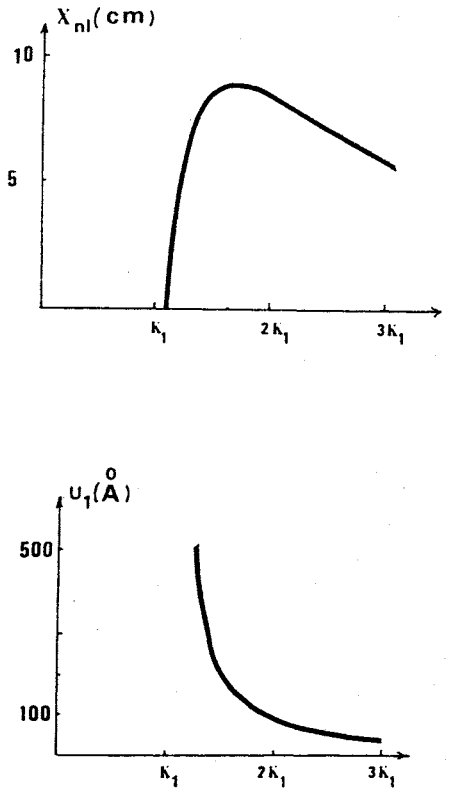


Fig. 11 : Distance of maximum instability $xn\ell$

and amplitude of the fundamental soliton

for a = 100 Å and τ = 100 ns

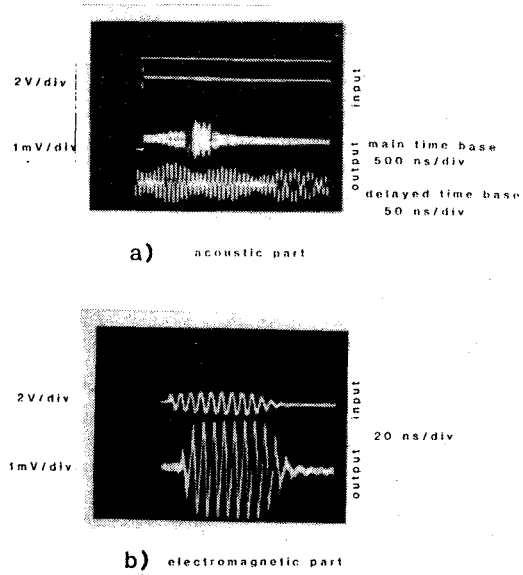


Fig. 12 : Modulational instability

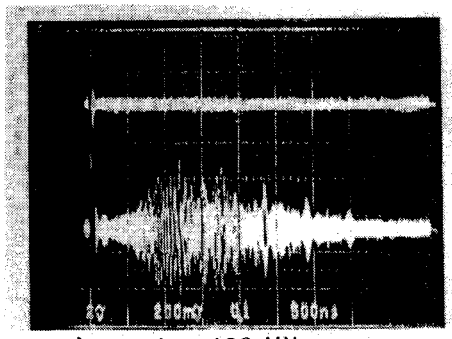
observed closed to the cut-off frequency

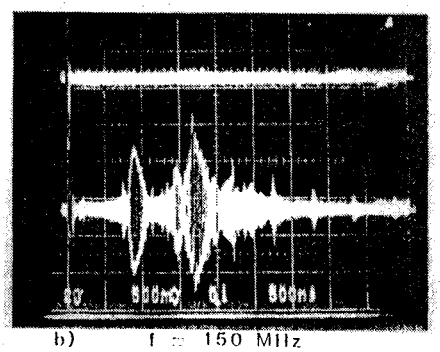
Experiments

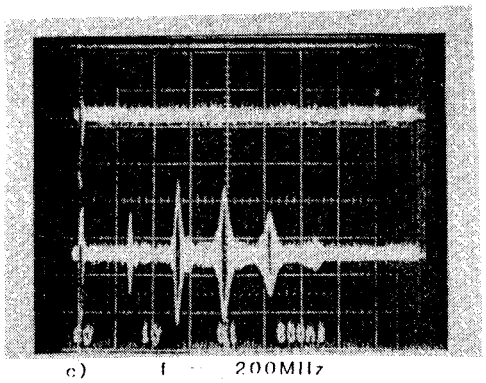
In a first set of experiments, rectangular pulses of 75 ns length and 2V amplitude were applied to the Y-Z cut device described at the beginning of this paper (§ 2). The applied modulational frequency 133 MHz was chosen to correspond to a strong instability (approximate zero group velocity). The output wavepacket (Fig. 12a) is observed to be very unstable (see the delayed time base) and extents over a 3 µs time window. The center wave packet delay 1.5 µs correspond to the velocity 4000 m/s of SSBW and the large time window reflects the finite length of the transducers.

In Fig. 12a the oscilloscope trace has been synchronized on the electromagnetic response of the delay line. An enlarged view of this response reveals a nonlinear structure which is correlated to the presence of the acoustic signal (this phenomenon will be analyzed in a future work).

In a second set of experiments input pulses of 50 ns duration and 7 V amplitude where fed into the input transducer at increasing frequencies. At the cut-off frequency 126 MHz) the response is broad and unstable as before (Fig. 13). At 150 MHz (Fig. 13b), the response has essentially split into two pulses corresponding to the two bounces illustrated on Fig. 1. At 200 MHz five bounces are observed. The time delay of each bounce decreases and then their apparent velocity increases with increasing frequency. Also the width of each bounce decreases with increasing velocity. It should be remarked that the notion of bounce has not to be used too far because there is not propagation in the conventional (hyperbolic) sense since the system is elliptic. On the other hand the amplitude dependence of bounce velocity and width is characteristic of nonlinear dispersive waves. The bounce propagation is reminiscent of that of a fundamental soliton. To confirm the hypothesis new experiments were performed.

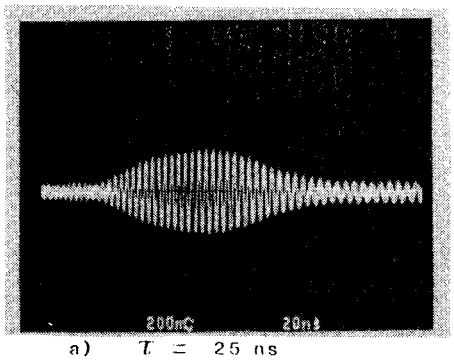




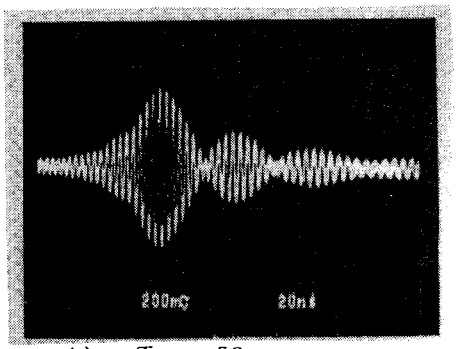


c) f = 200MHz

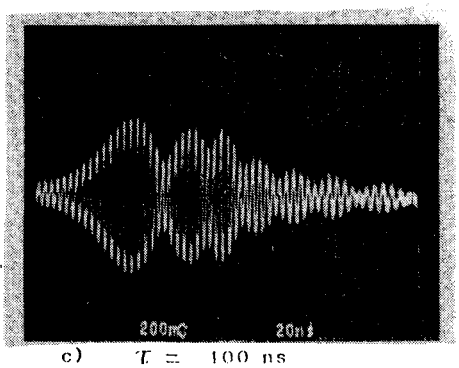
Fig. 13 : Impulse response of the quartz plate versus carrier frequency. Input pulse length : 50 ns, input amplitude : 7 V



a) τ = 25 ns

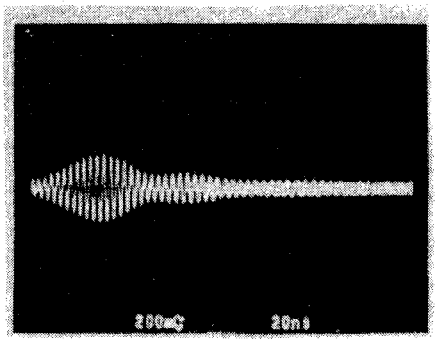


b) τ = 50 ns

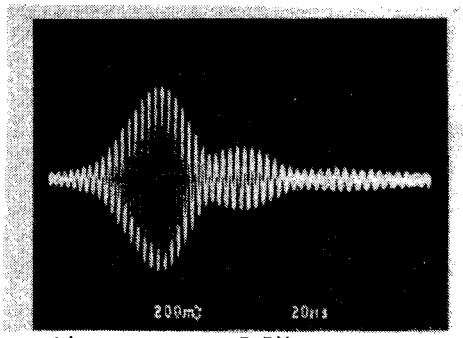


c) τ = 100 ns

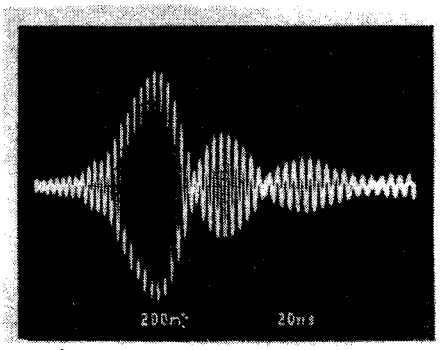
Fig. 14 : Impulse response of the quartz plate corresponding to the second bounce versus input length.
Carrier frequency : 280 MHz, input amplitude : 13 V



a) v = 2.5 V



b) v = 7.5 V



c) v = 15 V

Fig. 15 : Impulse response of the quartz plate corresponding to the second bounce versus input amplitude.
Carrier frequency : 280 MHz, input length : 40 ns

In Fig. 14 and 15 incident pulses with fixed frequency 280 MHz and variable length τ and amplitude v were considered. To make the oscilloscope record easier a 45 dB magnification of the output signal was used and all records were taken from the second bounce. In Fig. 14 the incident pulse had 13 V maximum amplitude. At 25 ns input length (Fig. 14a) the wave packet is very dispersive and the modulated frequency extents from 300 to 220 MHz. At 50 ns input length (Fig. 14b) the wave packet has split into three stationnary parts of frequencies 280 MHz, 256 MHz and 225 MHz. Dispersion is present only at the edges of the main signal. Then at 100 ns input length (Fig. 14c), modulational instability of the envelope is clearly seen and the frequency 280 MHz of the major part of the packet is that of the incident pulse. In Fig. 15, the incident pulse had 40 ns input length and variable amplitude. At 2.5 V (Fig. 15a) input amplitude the wavepacket is widely spread and is dispersive. At 7.5 V (Fig. 15b) input amplitude the wavepacket has split into two parts with frequencies 280 MHz and 244 MHz. At 15 V input amplitude (Fig. 15c) the wave packet has split into three parts with frequencies 280 MHz, 244 MHz and 225 MHz. Moreover an approximate 50 % narrowing has occured for each packet.

Conclusion

In this paper we have shown that bulk acoustic wave radiation from very long interdigital transducers is dispersive and nonlinear. At the SSBW cut-off frequency theory revealed a linear zero group velocity corresponding to a strong modulational instability. This instability was observed together with a nonlinear shape of the electromagnetic breakthrough.

At higher frequencies, pulse narrowing, pulse splitting and modulational instability corroborated by the existence of solitons was observed in rough agreement with quantitative estimate.

Applications of solitons shoudl be mainly toward the field of coding and decoding (soliton correlator). The instability concept encountered in nonlinear dispersion should also be a valuable tool to explain 1/f frequency instabilities in resonators. Lastly characterization of materials by solitons should be useful for basic physics.

References

[1] M. Planat, G. Théobald and J.J. Gagnepain, l'Onde Electrique 60(8-9) (1980) and 60(11) (1980).
[2] J.E. Feir, Proc. Roy. Soc. London, A299, 54 (1967).
[3] K.E. Lonngren, in "Solitons in Action", edited by K. Lonngren and A.Scott (Academic Press, New York, 1978), p. 127.
[4] H. Ikezi, S.S. Wojtowicz, R.E. Woltz, J.S. de Crassie and D.R. Baker, J. Appl. Phys. 54, 3277 (1988).
[5] H. Ikezi, in "Solitons in Actions", edited by K. Lonngren and A.Scott (Academic Press, New York, 1978), p. 153.
[6] L.F. Mollenauer, R.H. Stolen and J.P. Gordon, Phys. Rev. Lett. 45, 1095 (1980).
[7] S. Maneuf, R. Desailly and C. Froehly, Optics Communications 65, 193 (1988).
[8] Y. Zhang, M. Planat, Electronics Lett. 32, 68 (1987).
[9] G.B. Witham, "Linear and nonlinear waves", John Wiley & Sons, 1974, p. 601.
[10] J.J. Gagnepain and R. Besson, "Nonlinear effects in piezoelectric quartz crystals", Physical Acoustics, vol. XI, Academic Press, New York (1975).
[11] V.E. Zakharov, A.B. Shabat, Sov. Phys. JEPT 34, 62 (1972).
[12] J. Satsuma, N. Yajima, Suppl. Prog. Theor. Phys. 55, 284 (1974).
[13] D. Krökel, N.J. Halas, G. Giuliani, D. Grischkowsky, Phys. Rev. Lett. 60, 29 (1988).
[14] M.J. Lighthill, J. Inst. Math. Applic. 1, 269 (1965).
[15] J.F. Ewen, R.L. Gunshor and V.H. Weston, J. Appl. Phys. 53, 5682 (1982).
[16] R.H. Stolen, L.F. Mollenauer and W.J. Tomlinson, Opt. Lett. 8, 186 (1983).
[17] H.C. Yuen and B.M. Lake, in "Solitons in Action", edited by K. Lonngren and A.Scott (Academic Press, New York, 1978), p. 89

INDEX DES AUTEURS - AUTHOR INDEX

AUTEUR/AUTHOR	PAGE
ARNAUD R.	57
AUBRY J.P.	362, 387
AUDOIN C.	277
AULD B.A.	152
AVINENS C.	57
BACHHEIMER J.P.	61
BALLANDRAS S.	159
BARILLET R.	249
BATES A.G.	11
BAVA E.	259
BEAUSSIER J.	221, 238
BEETLEY D.E.	139
BERNIER L.G.	215
BIDART L.	181, 187
BIGLER E.	159
BIGNON O.	57
BLONDE D.	362, 387
BOURGEOIS C.	235
BOURQUIN R.	305
BRENDEL R.	28
BREUZET M.	221
BUISSON X.	57
BUSCA G.	215
BUZEK O.	204
CALDERA C.	259
CALISTI S.	165
CAPELLE B.	57, 301
CEREZ P.	81, 277
CERMAK J.	204
CLOEREN J.M.	11
CORDARA F.	204
CRACKNELL M.P.	169
CRAVEUR J.C.	221
DALY P.	121
DAUMET J.	356
DE CLERCQ E.	281
DEFRANOULD P.	165
DE HOND P.	293
DE JONG G.	198
DE LABACHELERIE M.	81
DE LA FOURNIERE P.	387
DELAITE R.	296
DESBOIS J.	381
DETAINT J.	227, 301
DIEULESAINT E.	293
DIMARCQ N.	274
DIOMANDE K.M.	81
DOLINO G.	61
DUFILIE P.	381
DULMET B.	305
ELCHEIKH R.E.	194
EL GHAZI A.	255
ENGEL G.	50
ERTMER W.	264
FASSL S.	94
FOISE J.W.	337
GAIGNEBET J.	220
GALLIANO P.G.	204
GALLIOU S.	147
GAUTIER H.	369, 381
GAY M.	221
GIORDANO V.	274
GIRARDET E.	181, 187
GODONE A.	259
GOIFFON A.	57
GRANVEAUD M.	104
GROSLAMBERT J.	173, 245
GUINOT B.	1
HADORN F.	269
HAMEL A.	277
HARRISON A.M.	169
HAUDEN D.	159, 165
HELLWIG H.	5
HOBDEN M.K.	180
HODE J.M.	381
HRUSKA C.K.	35
IMAE M.	89
IRVINE R.A.	337
JOHNSON G.R.	337
JOHNSON L.	73
JOHNSON W.A.	69
JOLY C.	227
KARUZA S.K.	69
KEMSSU P.	81
KHAYAR M.	356
KIRSHNER D.	94
KITCHING I.D.	121
KREMPL P.	50
KULMALA T.	326
LEBLOIS P.	173
LEBLOIS T.	41, 346
LE CORRE J.L.	387
LESCHIUTTA S.	220
LESTRADE J.F.	64
LEWANDOWSKI W.	89, 104
MAITRE P.	41
MANGIN P.	281
MARSHALL R.J.T.	133
MERIGOUX H.	356
MICHEL J.P.	372, 387
MIRANIAN M.	89
MORBIEU B.	293
MORLEY P.E.	133
MOUREY M.	147
MULLER J.H.	264
NELLESSEN J.	264
NORTON J.R.	11
OLIVIER M.	255
PEIER U.	11
PENAVAIRE L.	165, 369
PETIT P.	277
PETTITI V.	204
PHILIPPOT E.	57, 227
PLANAT M.	394
RENOULT P.	187
RESSLER H.	94
ROYER D.	293
RUBIOLA E.	313
SASSI M.P.	259
SAXENA G.M.	85
SCHLUETER B.	11
SCHUMACHER P.	319
SCHWARTZEL J.	227, 301
SEBASTIAN M.T.	57
SENGSTOCK K.	264
SHARPE D.J.	169
STADLER J.	50
STOERMER P.	143
SUTER J.J.	11
TAVELLA P.	204
TELLIER C.	41, 346
THEOBALD G.	274, 277
THOMANN P.	73, 269
THOMAS C.	89
TOURDE R.	104
UEBERSFELD J.	221
UHRICH P.	104
VALENTIN M.	194
VIALLE N.	362
VIARD B.	356
VIENNET J.	333
VOIT F.J.	69
VONDRAK J.	204
WILLIAMSON R.J.	22, 133
ZARKA A.	57, 301
ZHENG Y.	301

Imprimé en France

par l'Imprimerie du Conseil
Général du DOUBS
Hôtel du Département
Avenue de la Gare d'Eau
25031 BESANCON CEDEX

Printed in France

by the Printing Office of Conseil
Général du Doubs
Hôtel du Département
Avenue de la Gare d'Eau
25031 BESANCON CEDEX

Copyright by
3e Forum Européen Temps-Fréquence

Tous droits de reproduction et de traduction réservés
pour tous pays y compris l'U.R.S.S.