(2)

AD-A217 275

# PROCEEDINGS

# OF THE

# 42ND ANNUAL

# FREQUENCY CONTROL SYMPOSIUM

# 1988

Co-Sponsored by the

U.S. ARMY
ELECTRONICS TECHNOLOGY AND
DEVICES LABORATORY

and

THE INSTITUTE OF ELECTRICAL AND
ELECTRONICS ENGINEERS, INC.
ULTRASONICS, FERROELECTRICS AND
FREQUENCY CONTROL SOCIETY

IEEE Catalog No. **88CH2588-2**

Library of Congress No. 58-60781

DTIC
ELECTE
JAN 11 1990
S B D

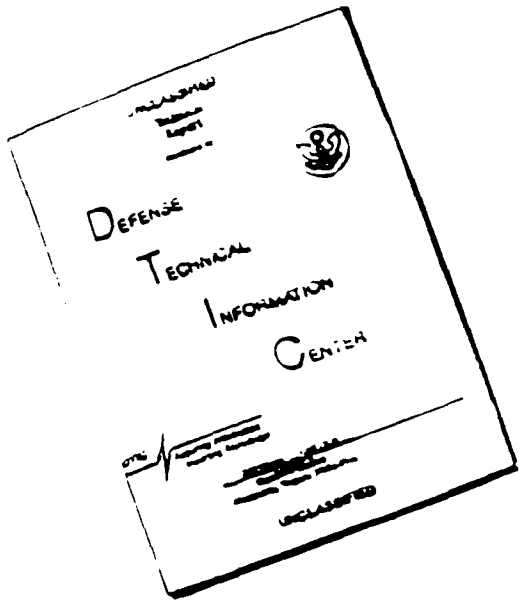
Stouffer Harborplace Hotel

Baltimore, Maryland

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

90 01 11 029

# DISCLAIMER NOTICE




THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

Abstracting is permitted with the credit to the source. Libraries are permitted to photocopy beyond the limits of US copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through the Copyright Clearance Center, 21 Congress Street, Salem, MA 01970. Instructors are permitted to photocopy isolated articles for commercial classroom use without fee. For other copying, reprint or republication permission, write to Director, Publishing Services, IEEE, 345 East 47th Street, New York, NY 10017. All rights reserved. Copyright © 1987 by the Institute of Electrical and Electronics Engineers, Inc.

## SYMPOSIUM CHAIRMEN

General Chairman.....Dr. John R. Vig, U.S. Army LABCOM
Technical Program Chairman.....Dr. Thrygve R. Meeker, Consultant
Finance Chairman.....Dr. Thomas E. Parker, Raytheon
Publicity Chairman.....Dr. Raymond L. Filler, U.S. Army LABCOM
Local Arrangements Chairman.....Michael Driscoll, Westinghouse Systems Center

## VICE CHAIRMEN

Michael Driscoll.....Westinghouse Systems Center
Larry Halliburton.....Oklahoma State University
Donald Malocha.....University of Central Florida
Lauren Rueger.....Applied Physics Laboratory

## TECHNICAL PROGRAM COMMITTEE

David Allan
National Bureau of Standards

Arthur Ballato
U.S. Army LABCOM

Martin Bloch
Frequency Electronics, Inc.

Jean-Simon Boulanger
National Research Council of Canada

Janet Brown
Fischer Controls International, Inc.

Leonard Cutler
Hewlett Packard Company

Michael Driscoll
Westinghouse Systems Center

Errol EerNisse
Quartztronics, Inc.

Raymond Filler
U.S. Army LABCOM

Marvin Frerking
Rockwell International

Satoru Fujishima
Murata Manufacturing Co., Ltd.

Jean-Jacques Gagnepain
CNRS-LPMO

Edward Garber
TRW

Michael Garvey
Frequency & Time Systems

Larry Halliburton
Oklahoma State University

Helmut Hellwig
National Bureau of Standards

James Pon-Wei Hou
Quartz Frequency Tech., Ltd.

Charles Jensik
Piezo Crystal Company

Robert Kinsman
Motorola, Inc.

William Klepczynski
U.S. Naval Observatory

John Kusters
Ball Corporation

Kei-Fung Lau
TRW

Nicholas Lias
Motorola Quartz Operations

Theodore Lukaszek
U.S. Army LABCOM

Donald Malocha
University of Central Florida

Thrygve Meeker
Consultant

Y. Nakazawa
Toyo Communication Equipment Co., Ltd.

Thomas Parker
Raytheon Research Division

Robert Peach
GEC Research Ltd.

Vincent Rosati
U.S. Army LABCOM

Lauren Rueger
Applied Physics Laboratory

Stanley Schodowski
U.S. Army LABCOM

Robert Smythe
Piezo Technology, Inc.

Samuel Stein
Ball Corporation

Richard Sydnor
Jet Propulsion Laboratory

Jacques Vanier
National Research Council of Canada

John Vig
U.S. Army LABCOM

Rolf Weglein
Hughes Aircraft Company

Joseph White
Naval Research Laboratory

Bernot Winkler
U.S. Naval Observatory

Nicholas Yannoni
U.S. Air Force/RADC

DTIC
COPY
INSPECTED

ion For
GRA&I
AB
unced
ication
per letter
ution/
ability Codes
Avail and/or
Special
A-1

## TECHNICAL SESSION CHAIRMEN

**SUPERCONDUCTIVITY AND FREQUENCY CONTROL**
L. Cutler, Hewlett-Packard Company

**SURFACE ACOUSTIC WAVE DEVICES/MAGNETO STATIC WAVES**
W. Skudera, Jr., U.S. Army LABCOM

**OSCILLATORS I**
M. Driscoll, Westinghouse

**I.I. RABI MEMORIAL SESSION, ATOMIC FREQUENCY STANDARDS I**
J. Boulanger, National Research Council, Canada

**R.D. MINDLIN MEMORIAL SESSION, PIEZOELECTRIC RESONATORS I**
H. Tiersten, Rensselaer Polytechnic Institute

**PIEZOELECTRIC RESONATORS II**
E. EerNisse, Quartztronics, Inc.

**OSCILLATORS II**
V. Rosati, U.S. Army LABCOM

**ATOMIC FREQUENCY STANDARDS II**
L. Rueger, Johns Hopkins University

**TIME COORDINATION**
W. Klepczynski, U.S. Naval Observatory

**QUARTZ MATERIAL I**
J. Martin, Oklahoma State University

**QUARTZ MATERIAL II**
J. King, Sandia Laboratories

**DIELECTRIC RESONATORS**
A. Ballato, U.S. Army LABCOM

**PRECISION MEASUREMENTS**
R. Sydnor, Jet Propulsion Laboratory

**TUTORIALS: NOISE**
T. Meeker, Consultant

**FREQUENCY STABILITY**
F. Walls, National Bureau of Standards

**RESONATOR PROCESSING**
J. Kusters, Ball Corporation

**PIEZOELECTRIC FILTERS**
W. Smith, Bell Laboratories (retired)

**FREQUENCY SYNTHESIZERS**
T. Parker, Raytheon Research Division

## TABLE OF CONTENTS

**PLENARY SESSION**

Award Presentations..... 1

In Memoriam:

R.D. Mindlin..... 2

I.I. Rabi..... 4

**R.D. MINDLIN MEMORIAL SESSION**
**PIEZOELECTRIC RESONATORS AND FILTERS**

Mindlin's Elastoelectrodynamics Problem..... 6
A. Ballato, U.S. Army LABCOM

Acceleration Insensitivity of Thickness Frequencies of Doubly Rotated Quartz Disks..... 14
P. Lee & M. Tang, Princeton University

Linear Model of the Contoured Resonators..... 19
J. Detaint, H. Carru & J. Schwartzel, Centre National d'Etudes des Telecommunications, &
B. Capelle & A. Zarka, Universite Pierre et Marie Curie, France

Vibrations of Z-Cut Resonator-Structure by Finite Element Analysis..... 29
Y. Yong; Rutgers University, P. Lee, Princeton University & S. Chuang, Statek Corporation

Rigorous Two-Dimensional Equations for the Analysis of Contoured Crystal Resonators..... 38
R. Peach, GEC Hirst Research Centre, England

Vibration Analysis of Coupled Flexural Torsional Mode Tuning Fork Type Quartz Crystal Resonator..... 45
H. Kawashima, Seiko Electronic Components, Ltd., Japan

Temperature Derivatives of the Dynamic Permittivity and Permeability of the Simple Thickness Modes of Quartz Plates..... 53
J. Kosinski, A. Ballato, T. Lukaszek, M. Mizan, R. McGowan & K. Klohn, U.S. Army LABCOM

A Variational Analysis of a New Shape Face Shear Mode Quartz Crystal Resonator Formed by an Etching Method..... 65
H. Kawashima, M. Matsuyama & M. Nakazato, Seiko Electronic Components, Ltd., Japan

Chemically-Milled UHF SC-Cut Resonators..... 73
R. Smythe & R. Angove, Piezo Technology, Inc.

A Low Cost Force Sensing Crystal Resonator Applied to Weighing..... 78
W. Albert, Crystal Gage, Inc.

Stroboscopic X-Ray Topography of Quartz Resonators..... 85
A. Zarka, B. Capelle & Y. Zheng, Universite P & M Curie & J. Detaint & J. Schwartzel,
Centre National d'Etudes des Telecommunications, France

Probe Examination of Thickness-Shear Vibrations of AT-Cut Natural Quartz Crystals: Some New Results..... 93
R. Parshad & A. Sharma, National Physical Laboratory, India

New Monolithic Crystal Filter with Wide Tabs..... 95
Y. Okamoto, Nihon Dempa Kogyo Co., & H. Sekimoto, Tokyo Metropolitan University, Japan

Semi-Analytical Analysis of Coupled Thickness-Shear and Thickness-Twist in Monolithic Crystal Filters.... 101
H. Carru & R. Lefevre, CNET & J. Aubry & S. LeChopier, CEPE, France

Performance of TFR Filters Under Elevated Power Conditions..... 106
R. Ketcham, G. Kline & K. Lakin, Iowa State University

Frequency Sorting Based on Planar YIG Resonators..... 112
S. Hanna & S. Zeroug, Weber Research Institute, Polytechnic University

**QUARTZ MATERIAL AND RESONATOR PROCESSING**

High Performance Quartz (Invited Paper)..... 116
R. Laudise, R. Barns, D. Stevens & H. Brown, AT&T Bell Laboratories & E. Simpson,
AT&T Network Systems

Recent Experiments in a Silver Lined Autoclave..... 127
R. Irvine, J. Foise, E. Leeson & G. Johnson, Sawyer Research Products, Inc.

New Advances in Crystal Growth of High Purity Berlinite: A New Solvent the Sulfuric Acid..... 138
E. Philippot, A. Goiffon, J. Jumas & C. Avinens, Universite de Montpellier &
J. Detaint & J. Schwartzel, CNET & A. Zarka, Universite de Paris, France

A New Measurement of the Basic Elastic and Dielectric Constants of Quartz (Invited Paper)..... 146
B. James, GEC Hirst Research Centre, England

Possible Mechanisms for the Introduction of Hydrogen into Alpha-Quartz During Sweeping..... 155
J. Gualtieri, U.S. Army LABCOM

Evaluation of Resonators Fabricated from High Quality Quartz..... 162
J. Martin & A. Lopez, Oklahoma State University; A. Armington, U.S. Air Force RADC &
J. Balascio, Motorola, Inc.

Dose Dependence of Radiation-Induced Defect Changes in Quartz..... 169
H. Lipson & F. Euler, U.S. Air Force RADC

A Production Study of Acoustic Loss Related Defects in Quartz..... 176
A. Lopez, H. Hwang & J. Martin, Oklahoma State University

The Thermoluminescence (TSL) of Lithium - and Sodium-Swept Quartz Crystals..... 184
A. Halperin & S. Katz, The Hebrew University, Israel

New Technology for Detection and Removal of Surface Contamination Involving Particulates or Water/Organic Materials (Invited Paper)..... 189
S. Hoenig, University of Arizona

Measurement of Plano-Convex SC Quartz Blanks Using Lateral Field Excitation..... 202
A. Warner, Frequency Electronics, Inc.

Use of Annular SAW for the Cutting of Quartz Resonator Blanks and Substrates..... 205
J. Dowsett, R. Spencer & P. Morley, STC Components, England

X-Ray Topographic Study of Vacuum Swept Quartz Crystals..... 208
A. Zarka, M. Sebastian & B. Capelle, Universite de Pierre et Marie Curie, France

A New Shutter System for Fine-Tuning Coupled-Dual Crystals..... 211
G. Roberts, General Electric Company

**SURFACE ACOUSTIC WAVE DEVICES/MAGNETO STATIC WAVES AND DIELECTRIC RESONATORS**

Analysis of Anharmonics in Surface Skimming Bulk Wave Devices..... 217
J. Hou, Allied-Signal, Inc.

Experimental Temperature and Stress Sensitivities of Surface Acoustic Wave Quartz Cuts..... 224
E. Bigler, R. Coquerel & D. Hauden, Laboratoire de Physique et Metrologie des Oscillateurs du CNRS, France

An Analysis of the In-Plane Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators with the Substrate Extending Beyond the Supports..... 230
D. Shick & H. Tiersten, Rensselaer Polytechnic Institute

Improved Vibration Sensitivity of the All Quartz Package Surface Acoustic Wave Resonator..... 239
J. Greer & T. Parker, Raytheon Research Division

SAW Tapped Delay Lines for New Potential Circuit Applications..... 252
W. Skudera, Jr., U.S. Army LABCOM

Dielectric Resonators (Invited Paper)..... 259
P. Guillon, I.R.C.O.M., France

Temperature Sensitivity of Dielectric Resonators and Dielectric Resonator Oscillators..... 263
M. Loboda, T. Parker & G. Montress, Raytheon Research Division

Non-Linear Modeling and Performance of Oscillators Using Thin-Film Bulk-Acoustic Wave Devices..... 272
S. Burns, P. Thompson & G. Kline, Iowa State University

**OSCILLATORS AND SYNTHESIZERS**

A Wide Frequency Range, Surface Mountable, Voltage Controllable Crystal Oscillator Family..... 276
S. Logan, D. Embree, R. Sheehey & D. Stevens, AT&T Bell Labs

The Influence of Pressure and Humidity on the Medium and Long-Term Frequency Stability of Quartz Oscillators..... 279
F. Walls, National Bureau of Standards

Temperature Compensation of SAW Oscillators..... 284
M. Cracknell, A. Harrison & D. Sharpe, STC Components, England

Frequency Stability of SAW Oscillators at High Temperatures..... 288
B. Sinha, Y. Sudo, S. Sato, Nippon Schlumberger, Japan & J. Groves, Schlumberger-Doll Research

Theory and Design of Low Phase Noise Microwave Oscillators..... 301
R. Rogers, California Microwave Inc.

A Simple Way of Characterizing High Q Oscillators..... 304
J. Goldberg, National Semiconductor

A New Digitally Temperature Compensated Crystal Oscillator for a Mobile Telephone System..... 327
T. Miyayama, Y. Ikeda & S. Okano, Toyo Communications, Japan

An Acceleration Compensated Precision Quartz Oscillator..... 334
B. Milliren, D. Martin & D. Emmons, Frequency & Time Systems

The Relaxation Quartz Crystal Oscillators - Design & Performances..... 342
D. Vasiljevic, University of Belgrade, Yugoslavia

Clarification and a Generalized Restatement of Leeson's Oscillator Noise Model..... 348
B. Parzen, Consultant

Phase Noise in Direct Digital Synthesizers..... 352
E. Mattison & L. Coyle, Smithsonian Astrophysical Observatory

The Optimization of Direct Digital Frequency Synthesizer Performance in the Presence of Finite Word Length Effects..... 357
H. Nicholas, TRW, H. Samueli, University of California & B. Kim, Intel Corporation

A Phase-Locked 4-6 GHz Local Oscillator Using Microwave Prescaler..... 364
P. Kandpal, Rockwell International

Low Noise, Microwave Signal Generation Using Bulk and Surface Acoustic Wave Resonators..... 369
M. Driscoll, Westinghouse Electric Corporation

Low Phase Noise Oscillators Above 1 GHz Utilizing Hermetically Sealed All Quartz Resonators and Harmonic Extraction..... 378
C. Chase, R. Laton & C. Yuen, Raytheon Company

**FREQUENCY STABILITY, NOISE AND MEASUREMENTS**

Measurement and Analysis of Thermal Hysteresis in Resonators and TCXO's..... 380
R. Filler, U.S. Army LABCOM

Burst Noise and 1/f Noise in Quartz Crystals and Oscillators..... 389
G. Moulton, Hewlett-Packard Company

Resonator Surface Contamination - A Cause of Frequency Fluctuations?..... 397
Y. Yong, Rutgers University & J. Vig, U.S. Army LABCOM

Aging Prediction of Quartz Crystal Units..... 404
M. Miljkovic, G. Trifunovic & V. Brajovic, Institute Mihajlo Pupin, Yugoslavia

Experimental Results on Aging of AT-Cut Strip Resonators..... 412
J. Gehrke & R. Klawitter, Motorola, Inc.

Standard Terminology for Fundamental Frequency & Time Metrology..... 419
D. Allan & H. Hellwig, National Bureau of Standards, P. Kartaschaff, Swiss PTT.
J. Vanier, National Research Council of Canada, J. Vig, U.S. Army LABCOM,
G. Winkler, U.S. Naval Observatory & N. Yannoni, U.S. Air Force RADC

Noise and Time and Frequency --- A Potpourri..... 426
J. Barnes, Austron, Inc.

Extending the Range and Accuracy of Phase Noise Measurements..... 432
F. Walls, A. Clements, C. Felton, M. Lombardi & M. Vanek, National Bureau of Standards

Technique for Measuring the Acceleration Sensitivity of SC-Cut Quartz Resonators..... 442
M. Watts, E. EerNisse, R. Ward & R. Wiggins, Quartztronics, Inc.

Kalman Filter Analysis for Real Time Applications of Clocks and Oscillators..... 447
S. Stein, Ball Corporation and R. Filler, U.S. Army LABCOM

Radar Measurement Applications of Fiber Optic Links..... 453
I. Newberg, C. Gee, G. Thurmond & H. Yen, Hughes Aircraft

Precise Measurements of Quartz Crystal Units by Network Analyzer Technique Applied to Two Different Types of Test Jig ..... 456
V. Popovic-Milovanovic, B. Dobnikar & V. Popovic, Institute Mihajlo Pupin, Yugoslavia

**TIME COORDINATION**

Ensemble Time and Frequency Stability of GPS Satellite Clocks ..... 465
D. Allan & T. Peppler, National Bureau of Standards

Preliminary Comparison Between GPS & Two-Way Satellite Time Transfer ..... 472
W. Klepczynski, P. Wheeler, W. Powell, J. Jeffries, A. Myers & R. Clarke, U.S. Naval Observatory & W. Hanson, J. Jespersen & D. Howe, National Bureau of Standards

Fiber Optic Frequency Transfer Link ..... 478
L. Primas, G. Lutes & R. Sydnor, Jet Propulsion Laboratory

Israel's New Synchronized Time Scale, UTC (INPL) ..... 485
A. Shenhar & W. Litman, The National Physical Laboratory of Israel, A. Lepek & A. Citrinovitch, TFL, Israel and D. Allan & T. Peppler, National Bureau of Standards

**I.I. RABI MEMORIAL SESSION**
**ATOMIC FREQUENCY STANDARDS**

Performance of the PTB's Primary Clocks CS2 and CS1 ..... 490
A. Bauch, H. deBoer, B. Fischer, T. Heindorff & R. Schroeder, Physikalisch-Technische Bundesanstalt, W. Germany

Research on the Optically Pumped Cesium Beam Frequency Standards ..... 496
G. Theobald, V. Giordano, M. de Labachelerie, A. Hamel, N. Dimarcq, P. Cerez & C. Audoin, Universite Paris, France

Calculations on the Efficiency of Optical Pumping of a Cesium Atomic Beam by Lasers of Finite Linewidth ..... 505
C. Jacques, National Research Council & P. Tremblay, Universite Laval, Canada

Cesium Gettering by Graphite and Colloidal Graphite (Aquadag) ..... 510
N. Bhaskar & C. Kahla, The Aerospace Corporation

The Effect of Humidity on Commercial Cesium Beam Atomic Clocks ..... 514
J. Gray, H. Machlan & D. Allan, National Bureau of Standards

A Novel Compact Rubidium Frequency Standard with a Low Sensitivity to Magnetic and Vibrational Disturbances ..... 519
A. Stern, A. Hertz, Y. Zarfaty & A. Lepek, TFL, Israel

An Improved Rubidium Consumption Model for Discharge Lamps Used in Rubidium Frequency Standards ..... 525
R. Cook & R. Frueholz, The Aerospace Corporation

Neutron Hardness of Photodiodes for Use in Passive Rubidium Frequency Standards ..... 532
T. English & G. Malley, Ball Corporation & R. Korde, United Detector Technology

**SUPERCONDUCTIVITY AND FREQUENCY CONTROL**

High-Temperature Superconductivity: Novel Concepts, Fundamental Issues, and Frequency Control Aspects (Invited Paper) ..... 540
G. Iafrate, U.S. Army LABCOM & T. Parker, Raytheon Research Division

Superconducting Resonators and High-$T_c$ Materials ..... 545
D. Oates, A. Anderson & J. Steinbeck, M.I.T. Lincoln Laboratory

High Temperature Superconductors and Their Promise for Better Microwave Circuits (Invited Paper) ..... 550
P. Carr & J. Derov, U.S. Air Force RADC

Evaluating Superconducting Resonator Materials ..... 556
B. McAvoy, J. Adam & G. Wagner, Westinghouse R&D Center

**AUTHOR INDEX** ..... 559

**SPECIFICATIONS AND STANDARDS GERMANE TO FREQUENCY CONTROL** ..... 560

**PROCEEDINGS AVAILABILITY INFORMATION** ..... 562

# 1988 AWARD WINNERS

## The Cady Award

The Cady Award was presented to Baldwin Sawyer "for his work leading to the development of improved cultured quartz crystals, improved qualification techniques, and his tireless contributions to the frequency control industry." The award was presented by Arthur Ballato, U.S. Army LABCOM.

## The Rabi Award

The Rabi Award was presented to Gernot M.R. Winkler "for early development of worldwide clock synchronization through use of portable clocks; encouragement and support for the development of atomic frequency standards from their earliest days; and international leadership in the time and frequency community." The award was presented by Helmut Hellwig, National Bureau of Standards.

## The Sawyer Award

The Sawyer Award was presented to Charles A. Adams "for contributions to the development of unique devices and manufacturing technology." The award was presented by John A. Kusters, Efratom Div., Ball Corp.

Charles Adams, Sawyer Award winner; Gernot Winkler, Rabi Award winner; and Baldwin Sawyer, Cady Award winner; after the award presentation.

## IN MEMORIAM

# RAYMOND D. MINDLIN

Raymond D. Mindlin, Finch Professor Emeritus of Applied Science at Columbia University died on November 22, 1987 at the Dartmouth-Hitchcock Medical Center in Hanover, NH after a prolonged illness. He was 81 years old and resided in Grantham, NH at the time of his death. Professor Mindlin's active and prolific scientific career spanned more than half a century, during which he made significant contributions in solid mechanics and the peripheral fields of acoustics and optics. As a consequence of this work he was a world renowned authority in both engineering and scientific fields and a consultant to government and industry. In particular, his very important work on the mathematical theory of the vibration of elastic plates and its impact on the field of frequency control is the reason for the inclusion of this remembrance in these proceedings.

Raymond David Mindlin was born in New York City on September 17, 1906. He received four degrees from Columbia University, culminating in the Ph.D. degree in 1936. In the thesis Mindlin addressed and solved the problem of a force at a point in the interior of a semi-infinite solid, which is now and was shortly thereafter regarded as a classic. Mindlin began teaching at Columbia while still a graduate student and started graduate education in mechanics at Columbia in the thirties. Since his mentors were elsewhere, i.e., at Michigan and Harvard, he was in a very real sense the father of the graduate program in mechanics at Columbia. After receiving his degree he continued on at Columbia until the war when he left to serve the government at the Applied Physics Laboratory of Johns Hopkins University. It is of interest to know that during that period Professor Mindlin was instrumental in getting an Italian Jewish scientist out of Italy and that subsequently this scientist was instrumental in getting others in the same boat out.

At the Applied Physics Laboratory Mindlin led a team which designed the firing element of the first proximity fuse, a very important device in World War II. For this he was awarded the Presidential Medal of Merit in 1946. After the work on the proximity fuse Mindlin consulted for Bell Telephone Laboratories, where he worked on the design of spring systems to reinforce containers used for packaging electronic tubes so that they would not break when the packages were dropped from a certain height. This formed the basis of a publication entitled "Dynamics of Package Cushioning," which spawned an entire field.

After the war Mindlin returned to Columbia and became a full professor in 1947. However, he still consulted regularly at Bell Labs. During this period he was made aware of some of the problems associated with the proper functioning of the carbon microphone. This led to his pioneering work on the response of bodies in contact when subject to pressure and shear as a result of elasticity and friction. The results of this work have been applied by others in the design of relay contacts and ball bearings. Mindlin himself used the results in his work on the description of the behavior of granular media.

During this same period Mindlin was introduced to the field of frequency control by being asked some questions by people at Bell Labs about the loss of activity in AT-cut quartz thickness-shear resonators for certain ranges of aspect ratio of the plates. Mindlin was asked these questions even though he knew nothing about thickness-shear because he was an expert in the theory of elasticity and the flexure of thin plates. Although Mindlin had some difficulty communicating with most of the people, he found one, namely Irv Fair, who was able to educate him on what he did not know. This led to his work on the vibrations of crystal plates, in which he included the coupling of thickness-shear to flexure, which quantitatively described the behavior of the activity with aspect ratio and enabled AT and T to relax its tolerances in the fabrication of the plates. But, more importantly, this work led, in one way or another, to all work on the accurate analytical description of the three-dimensional modal behavior of quartz plates to this day.

Once having entered the field of crystal plate vibrations, Mindlin never left it. To be sure, he digressed into other fields at times but he always returned. For example, in the fall of 1960 Rudolph Bechmann asked Professor Mindlin whether a new theory of elasticity with an asymmetric stress tensor and forty-five independent elastic constants that was proposed by two independent groups of scientists should be included in a revision of the IEEE Standards that was currently underway. After a brief investigation Mindlin recommended that the committee ignore the forty-five constant theory. His recommendation was based on the fact that both groups had neglected to satisfy the law of the conservation of angular momentum. This investigation motivated Professor Mindlin to think of generalizing the theory of elasticity by including a couple-traction in addition to the usual force-traction. This led to the work on couple-stress

theory, which led further to work on even more general elastic continua, including microstructure theory. However, even though Mindlin worked avidly on these generalized descriptions during this period, he continued his work on crystal plate vibrations, to which his many publications in this field during the same period testify.

After originating and working in the area of microstructure theory for a number of years, Mindlin investigated ionic lattice theory in order to compare the long wavelength limit with microstructure theory. When he found that the systems of equations differed in certain respects, Mindlin halted his effort on microstructure theory and devoted his time to other productive research. From this it seems clear that just as specific physical considerations led Mindlin to investigate both couple-stress theory and microstructure theory, they caused him to abandon his effort on microstructure theory. This thread of intense concern with the description of physical reality in mathematical terms seems to pervade all of Mindlin's research.

Mindlin had remarkable physical insight and an uncanny ability to get to the heart of a problem, extract the essentials and construct a simple useful model that could relatively quickly yield the desired information. Partly for this reason he was extremely helpful to students he was advising when difficulties arose in their research. He was an excellent, clear and inspiring teacher who was sincerely appreciated by the students. This kind, mild-mannered, gentle man was admired by students and colleagues alike. All held him in great esteem and valued his counsel highly. In fact, a book dedicated to Professor Mindlin *encompassing eight* distinct subject areas to which Mindlin had made significant contributions was written by a group of his former students and published in 1974.

Mindlin was a member of the National Academy of Sciences and the National Academy of Engineering. He was a fellow of the American Academy of Arts and Sciences, the American Society of Mechanical Engineers and the Acoustical Society of America. Honors and awards that are too numerous to mention were showered on him by such organizations as the American Society of Mechanical Engineers, the American Society of Civil Engineers, the Acoustical Society of America, the Society for Experimental Stress Analysis and, of course, his alma mater and the school where he spent his career, Columbia University. Among these was the C. B. Sawyer memorial award, which was presented to him at the 21st Annual Symposium on Frequency Control in 1967. The citation stated "For fundamental contributions to the theory of vibration in piezoelectric resonators leading directly to advancements in the art." In recognition of his life's work, Mindlin received the National Medal of Science in 1979. Among the industrial organizations he served as a consultant to were Bell Telephone Laboratories, General Electric, General Motors and IBM. He served as an adviser to several government agencies, including the National Defense Research Committee.

Mindlin's first wife, the former Elizabeth Roth, died in 1950; his second wife, the former Patricia Kaveney, died in 1976. A memorial service for Professor Mindlin was held on March 23, 1988 in St. Paul's Chapel on the Columbia campus.

During the long illness prior to his death, Professor Mindlin was struggling to complete the third section of a three section book on the vibrations of crystal plates, but he was very weak and could not find the energy to get the work done. Fortunately, Peter Lee is working to complete the third section of the book, which, thankfully, will now be finished.

## IN MEMORIAM

# I. I. RABI

I. I. Rabi died on January 11, 1988. An obituary which appeared on the front page of The New York Times on January 12, 1988 is reprinted below, with the permission of The New York Times. The photograph is from Rabi's last appearance at the Frequency Control Symposium, in 1986.

# Isidor Isaac Rabi, a Pioneer In Atomic Physics, Dies at 89

By MARILYN BERGER

Isidor Isaac Rabi, a pioneer in exploring the atom and a major force in 20th-century physics, died yesterday at his home on Riverside Drive after a long illness. He was 89 years old.

Dr. Rabi, who won the 1944 Nobel Prize in Physics, was a strong intellectual influence on succeeding generations of physicists and a moral influence in the debate over controlling the power of the atom. The center for physics that he established at Columbia University became a training ground for scientists who went on to establish other major centers.

He was awarded the Nobel Prize for developing a method of measuring the magnetic properties of atoms, molecules and atomic nuclei. His work in turn made possible the precise measurements necessary for the development of the atomic clock, the laser and diagnostic scanning of the human body by nuclear magnetic resonance.

I. I. Rabi

### Tailor or Physicist

The man who was known as I. I. Rabi to his students at Columbia, and as Rab to his wife and friends, was born July 29, 1898, in the town of Rymanow, Austria-Hungary. He was a baby when his parents brought him to New York's Lower East Side. When his father worked, he worked as a tailor. When he did not work, the family went hungry.

"It's a miracle," Dr. Rabi said years later, "how a sickly child from a Lower East Side poverty-stricken family moved in one generation to where I did. Had we stayed in Europe, I probably would have become a tailor."

Fellow scientists admired what they called his "street smarts" and came to regard him as a conscience of their community. Dr. Rabi (pronounced RAH-bee) devoted much of his life to channeling man's knowledge of the

*Continued on Page A24, Column 1*



# *Isidor Isaac Rabi, Leader in*

*Continued From Page A1*

forces of the atom to peaceful purposes. He was the originator of the idea for the CERN nuclear research center in Geneva, which was instrumental in the rebirth of science in post-war Europe.

In this country he played a major role in the creation and style of the Brookhaven National Laboratory on Long Island. He devoted a great deal of his energy to reforming science education in the United States and ending the isolation of science from the humanities.

## Physics as a Discipline: Noble and Fundamental

Physics, to Dr. Rabi, was the most ennobling of disciplines, as well as the most fundamental. He said he had always tried to impart to his students a certain point of view of physics, its greatness. "You're wrestling with a champ," he recalled telling them. "You're trying to find out how God made the world, just like Jacob wrestling with the angel."

Although most of Dr. Rabi's pure scientific research was over by the 1940's, he was a major force in the world of 20th-century science. Sidney Drell, the theoretical physicist, described Dr. Rabi as "a very great giant on the scientific scene."

Not only was Dr. Rabi a great research scientist, Mr. Drell said, but he was also the founder of a great center of research in physics at Columbia. He also cited Dr. Rabi's contributions to science policy in World War II, his role in reviving European science through the creation of CERN and "his clear speaking out on the deep moral issues" involving nuclear weapons.

In World War II, Dr. Rabi was a leader of the research team in Cambridge, Mass., that helped in the development of radar. He also served as a senior adviser on the Manhattan Project, which developed the atomic bomb, and later as a science adviser to President Eisenhower.

As a professor, Dr. Rabi did not enjoy a reputation as a great lecturer and was feared by students as a tough taskmaster. But he is remembered for his moral integrity and for an impeccable taste that set a style for the study of physics in the United States.

One former Columbia faculty member said: "The most spectacular thing about Rabi was that during a 15-year period there were four Nobel Prizes all in different fields of physics at Columbia. Although Rabi wasn't directly involved in the specific work, he was the key motivating person. He built a great physics department so far beyond anything else in the world at the time that nobody could hold a candle to it."

Many of the teachers in most of the great centers of physics in the United States studied in Dr. Rabi's physics department. Of his many honors, Dr. Rabi was especially proud of one from the American Association of Physics Teachers, the Hans Christian Oersted award for his notable contribution to the teaching of physics.

Dr. Rabi actually stumbled into physics after he had already graduated from college with a degree in chemistry but with an abiding interest in the structure of matter.

"It was a time," he remembered, "of great revolution in fundamental physical ideas of time and space and causality. It was a little like landing on Ellis Island and going west and discovering a whole continent."

As a young man, Dr. Rabi won a scholarship to Cornell University, after graduating from Manual Training High School in Brooklyn.

"I turned away from the Old World," he said one day as he reminisced about his life. "I realized I had to be an American, not a Jewish-American. In all of my reading, I tried to become an American. I read a tremendous amount of colonial history. It takes a person like me to really understand what a wonderful country America is."

"I'm an omnivorous reader," he said, looking back at his discovery of the local branch of the Brooklyn Public Library after moving to the Brownsville section of Brooklyn when he was a child. He was about 10 years old when he started reading through the shelves of books, finally coming upon one about Copernicus that changed his idea of the world. "I was small for my age; I still am," the 5-foot 4-inch tall physicist said in his 88th year. "The librarian made me read to her from one of the books I was borrowing to prove I could read before she let me leave."

He said Jack London's writings made him a Marxist by the time he was 13 years old. "What appealed to me was the democratic idea that anybody

> **His work led to the development of the laser and the atomic clock.**

could become anything," he said. But he later decided that Marxism could not work. "The ideas are very appealing," he said, "but they're not meant for humans."

At Cornell, he signed up for engineering and chemistry on a scholarship so meager that he lost several teeth because of malnutrition. "I was living on a dollar a day," he said, "and I didn't see wasting the money on vegetables."

After graduating from Cornell, he got a job as a chemist analyzing furniture polish and mothers' milk, put out a local newspaper in Brooklyn and worked for a time in a lawyer's office dealing with accounts receivable. When he was 25 years old, he said, he decided it was time "to stop horsing around." He went back to Cornell for graduate work in chemistry.

"Now comes the celebration," he said, looking back with relish at what happened to him. "I applied for a fellowship and I didn't get it. So I decided to study physics, and it was then that I realized that this was my field."

Dr. Rabi transferred to Columbia, not only because a fellowship was available to him there, but also, as he put it many years later, because he had met "the most beautiful girl in the world," who lived in New York. Three years later, on Aug. 17, 1926, that girl,

# Atomic Physics, Dies at 89

Helen Newmark, became his wife. She survives him, and do their two daughters, Nancy Lichtenstein of Princeton, N.J., and Margaret Beels of New York, and four grandchildren.

After receiving his doctorate from Columbia in 1927, Dr. Rabi obtained a small fellowship for postgraduate study in Europe, where the major work in physics was being done. He went from Munich to Copenhagen to Hamburg to Leipzig to Zurich. Arriving in Copenhagen, he presented himself at the institute of Niels Bohr, was given a key and allowed to go to work on his own.

## Making First Contact With Top Physicists

In Europe he met and worked with some of the outstanding physicists of the time: Bohr, Wolfgang Pauli, Pascual Jordan, Arnold Sommerfeld and Werner Heisenberg. Perhaps the most influential, in Dr. Rabi's work, was Otto Stern, who won the Nobel Prize in Physics the year before Dr. Rabi.

Among the scholars traveling in Europe were other scientists from the United States who were to become lifelong friends: Hans Bethe, Linus Pauling, E. U. Condon and J. Robert Oppenheimer. With Dr. Rabi, these men became the link between European and American science at a time of tremendous ferment.

With his fellowship money running out, Dr. Rabi faced returning to America with no prospect of a job. He said he knew that anti-Semitism in the universities made it unlikely that he would be accepted anywhere, and he was thunderstruck when in 1929 he received a *cablegram from Columbia offering* him a job at $3,000 a year, a princely sum on the eve of the Depression. He believed he was the first Jew on the faculty of the Columbia physics department, advancing from lecturer to assistant professor to Higgins Professor of Physics t professor emeritus. In 1964 he was named Columbia's first University Professor and in 1985 a chair in physics was named for him.

At the beginning he had to fight to keep his research appropriation and even offered to take a cut in salary rather than a cut in his research funds. When Harold Urey, a professor at Columbia, won the Nobel Prize in Chemistry in 1934, he astonished Dr. Rabi by giving him half the prize money so he could continue his research.

By 1937 that research had led him to the technique for which he won his Nobel Prize. It followed directly from his association with Otto Stern at the University of Hamburg, where Dr. Stern applied magnetic fields to beams of molecules. He confirmed that the proton, a particle found in the nucleus of all atoms, generated a weak magnetic field by its spin.

This method for measuring the spin of the core of the atom, or nuclear magnetic moments, became the central technique of all modern molecular and atomic beam experiments, and that won for him the Nobel Prize.

Dr. Rabi's research led to the ability to make the most precise measurements that today are necessary for the guidance systems of missiles and satellites. It also led to the advanced medical diagnostic technique of magnetic resonance imaging.

It was in 1945 that Dr. Rabi proposed the construction of an atomic clock, capable of tuning in on atomic frequencies.

When word of the Nobel Prize reached him, Dr. Rabi was working at the Radiation Laboratory of the Massachusetts Institute of Technology, developing radar. He had refused Oppenheimer's invitation to become deputy director of the Manhattan Project, although he occasionally went to the atomic bomb laboratory at Los Alamos, N.M., to serve as a troubleshooter and consultant.

Dr. Rabi said he believed it might be possible to win the war without the bomb, but that without radar the war would be lost. Furthermore, he reportedly had reservations about a weapon of mass destruction becoming the culmination of three centuries of physics.

*Nevertheless, Dr. Rabi said that in* the war he was ready to support any idea that would contribute to the defeat of Hitler. "Some person would come along with a bright idea and I'd say, 'How many Germans will it kill?' " he *recounted. With that in mind, he sup*ported the push to develop the atomic bomb and overcame his doubts about using it.

He witnessed the first nuclear explosion on July 16, 1945, and later told an interviewer that for the first moment he was thrilled. "Then, a few minutes *afterward," he said, "I had gooseflesh* all over me when I realized what this meant for the future of humanity."

Dr. Rabi joined with Enrico Fermi to oppose the next step in the arms race, the "Super," a weapon proposed by Dr. Edward Teller and others that eventually was developed as the hydrogen bomb. Oppenheimer had also opposed it as a potential "weapon of genocide."

Dr. Rabi and Fermi wrote: "The fact that no limits exist to the destructiveness of this weapon makes its very existence and the knowledge of its construction a danger to humanity as a whole. It is necessarily an evil thing. We think it is wrong on fundamental ethical principles to initiate the development of such a weapon."

Speaking to his last class at Columbia in 1967, Dr. Rabi said he had realized early on that "just because we got there first doesn't mean that we should have the power of life and death over the whole world."

## Years of Devotion To Controlling Arms

Immediately after the war, Dr. Rabi started his efforts to control the atom, working first with Oppenheimer to devise what became known as the Baruch Plan for international control of atomic energy.

Forty years later, reminiscing about the ideas behind the Baruch Plan, Dr. Rabi said: "It had to be self-enforcing and universal. I think it was our finest moment in America. People wouldn't believe it today. The Russians turned it down, but that was expected, it was such a new and revolutionary idea. We wanted to internationalize everything. We should have gone back to the Russians when they said no."

In his later years, his eyesight failing, his face still unlined, Dr. Rabi grew increasingly impatient with the ever-increasing stockpiles of weapons, convinced that each new weapon would inevitably produce a countermeasure. He wanted nothing less than a re-education of the American people.

"The more you talk about arms control, the more you become morally obtuse," he said. "You have to recoil from the horror of it. You have to arouse the imagination of people building bombs in peacetime to use against an ally in two world wars.

"Americans are a moral people. They have respect for human life even where there are differences of opinion. Most Americans have enjoyed Russian novels, Russian music. We've yet to produce novels or music as great as theirs. What you're talking about is killing individuals because you don't like their government. I say that's un-American. That hatred is un-American."

Dr. Rabi did not speak out often. He usually confined his opinions to the inner councils of government. He succeeded Oppenheimer as head of the General Advisory Committee of the Atomic Energy Commission and was a member of what became the President's Science Advisory Committee from its 1952 birth in the Office of Defense Mobilization until 1968.

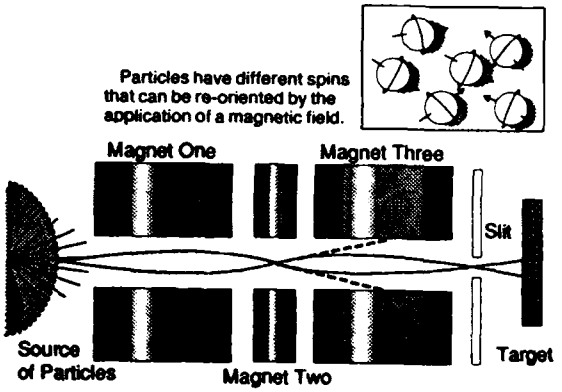
But he won the respect of many scientists for his forthright and spirited defense of Oppenheimer at hearings in 1954 of the Atomic Energy Commission, which ultimately stripped the former head of the Manhattan Project of his security clearance over questions raised in the McCarthy era about his loyalty to the country. "I was sore at pygmies trying a great man like Oppenheimer," Dr. Rabi said years later.

Dr. Rabi told the committee that because of the leadership of Oppenheimer, the United States had the atomic bomb, a whole series of them. In a burst exasperation, he asked, "What more do you want — mermaids?"

In a life filled with honors, Dr. Rabi gave so little thought to making money *that he even promised himself never to* patent any of his ideas, including the atomic clock. "In the late 1930's," he *once recalled, "I and my friends sat* around and talked about what we'd do if we had a million dollars. I thought and thought and finally I said, 'I think *I'd buy a new hat.*' "

Mrs. Rabi said funeral services for her husband would be private and that a memorial service would be held at a date to be announced later.

## Precise Measurement of Atoms: A Major Contribution From Rabi




An extremely sensitive method of revealing the magnetic properties of atomic particles was one of I. I. Rabi's central contributions to physics. It led to scientific and practical applications that include the atomic clock and the medical imaging technique known as magnetic resonance imaging.

Dr. Rabi created a device that guides a beam of particles through a series of magnets. Depending on the direction of each particle's spin, the magnets push the particle up or down.

The first and third magnets are aligned to oppose each other, so their effects cancel out. Some atoms curve up and then down; others down and then up — and either way they pass through a slit to a detector.

The key to the device is the middle magnet, which can be precisely tuned to different frequencies. When the frequency matches the exact frequency of the particles in the beam — when they are in "resonance" — the spins flip to the opposite direction. As a result, the particles that curved *upward under the first magnet* will now curve upward again, and the particles that curved downward will curve downward (dotted lines), and they fail to reach the detector. Thus when the frequency of the magnetic is set correctly, the detector shows a sudden dip in the number of particles reaching it. That signal allows exact measurement of the particles' magnetic properties.

*Source: Richard P. Feynman, "The Feynman Lectures on Physics"*

42nd Annual Frequency Control Symposium - 1988

# MINDLIN'S ELASTOELECTRODYNAMICS PROBLEM

Arthur Ballato

US Army Electronics Technology & Devices Laboratory
LABCOM, Fort Monmouth, NJ 07703-5000

## Abstract

Professor Mindlin's vast contributions to crystal plate vibration theory and associated areas of mechanics could be the subject of an entire symposium. The present paper honors his memory by focusing on a very specific, but conceptually very enlightening discovery that he made in 1972: an exact solution of the three dimensional equations of piezoelectromagnetism leading to the prediction of electromagnetic radiation from a rotated-Y-cut quartz plate vibrating in simple thickness-shear. After Mindlin's paper was published it was recognized how the solution resolved a contradiction that arose in connection with lateral field excitation of plates and their representation by electrical equivalent networks.

We sketch Mindlin's solution to the problem of coupled elastic and electromagnetic fields for rotated Y-cuts of quartz, and present some additional results.

## Introduction

Crystal plate vibrations were studied first in their purely mechanical aspects. Incorporation of the interaction with electric fields then gradually developed within the framework of the quasistatic (QS) approximation; see, e.g., [1]-[6]. In the quasistatic approximation [7], [8], the curl of the electric field is taken to vanish; this is equivalent to allowing the speed of light to increase without bound.

The full coupling of the elastic and electromagnetic fields in crystals without boundaries has been treated by a number of authors [9]-[13]; conductivity and viscosity effects have likewise been considered [14], [15].

Professor Mindlin gave an exact solution of the fully coupled equations for the case of a bounded crystal, namely, an AT-cut quartz plate [16], [17]. This solution resolved a contradiction that arose when the QS approximation was used to describe lateral-field excitation of crystal plates [18]-[20].

Mindlin's solution is briefly sketched below. We begin first with developments that lead up to the case he treated.

## Plate Vibrations in the Quasistatic Approximation

Figure 1 summarizes our situation. We consider a laterally unbounded plate of quartz having orientation (YXl)θ; $X_1$ is the digonal axis. Wave propagation is assumed to take place along $X_2'$, the thickness direction; no variations take place along $X_1$ or $X_3'$. Under these conditions the plate modes are referred to as simple thickness modes.

The pure shear mode can be driven piezoelectrically by an electric field along $X_2'$; this is thickness excitation (TE). It can also be driven by an $X_3'$-directed field; this is lateral excitation (LE).

Professor Raymond D. Mindlin
September 17, 1906 - November 22, 1987

The other predominantly shear mode and the predominantly extensional mode can each be driven by lateral excitation only.

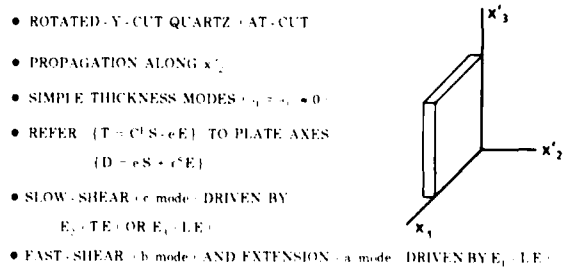



Figure 1. Conditions for Mindlin's problem.

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

Figure 2 gives the constitutive matrix for the QS case considered. The pure shear mode is governed by the elastic stiffness $c_{66}$ which is modified by piezoelectricity [2] to yield a "piezoelectrically stiffened" value $\bar{c}_{66}$. This is shown as function of rotation angle theta in Fig. 3, along with the piezo contribution ($\bar{c}_{66}$ - $c_{66}$). In Fig. 4 is given the exact network representation of the QS case for thickness excitation (TE). The piezoelectric transformer turns ratio $n_T$ is directly proportional to piezoelectric coupling factor k.

| | $S_2$ | $S_4$ | $S_6$ | $E_1$ | $E_2$ | $E_3$ |
|---|---|---|---|---|---|---|
| $T_2$ | $c_{22}$ | $c_{24}$ | 0 | $-e_{12}$ | 0 | 0 |
| $T_4$ | $c_{24}$ | $c_{44}$ | 0 | $-e_{14}$ | 0 | 0 |
| $T_6$ | 0 | 0 | $c_{66}$ | 0 | $-e_{26}$ | $-e_{36}$ |
| $D_1$ | $+e_{12}$ | $+e_{14}$ | 0 | $\varepsilon_{11}$ | 0 | 0 |
| $D_2$ | 0 | 0 | $+e_{26}$ | 0 | $\varepsilon_{22}$ | $\varepsilon_{23}$ |
| $D_3$ | 0 | 0 | $+e_{36}$ | 0 | $\varepsilon_{23}$ | $\varepsilon_{33}$ |

Figure 2. Constitutive matrix for piezoelectric case.

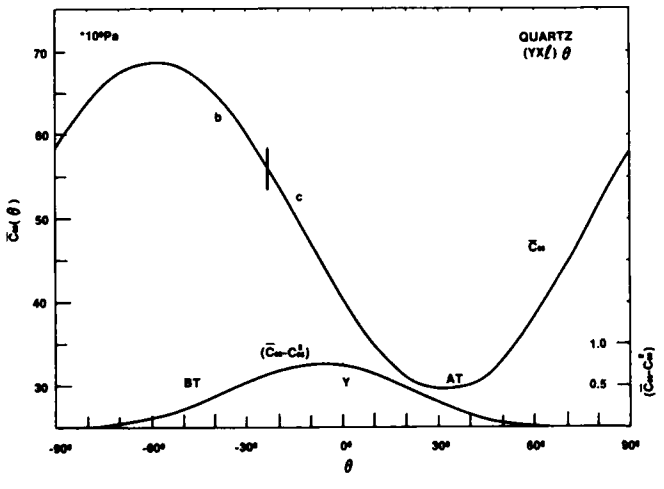



Figure 3. Piezoelectrically stiffened pure shear mode elastic constant for rotated-Y-cuts of quartz.

The direction of phase progression is along the thickness, but the direction of energy propagation is generally oblique to $X_2'$, although it remains normal to the digonal axis $X_1$. The difference in these directions can be quite appreciable, as seen in Fig. 5. Nearly all of the deviation from the thickness direction arises from elastic terms; the electromagnetic contribution in the QS approximation

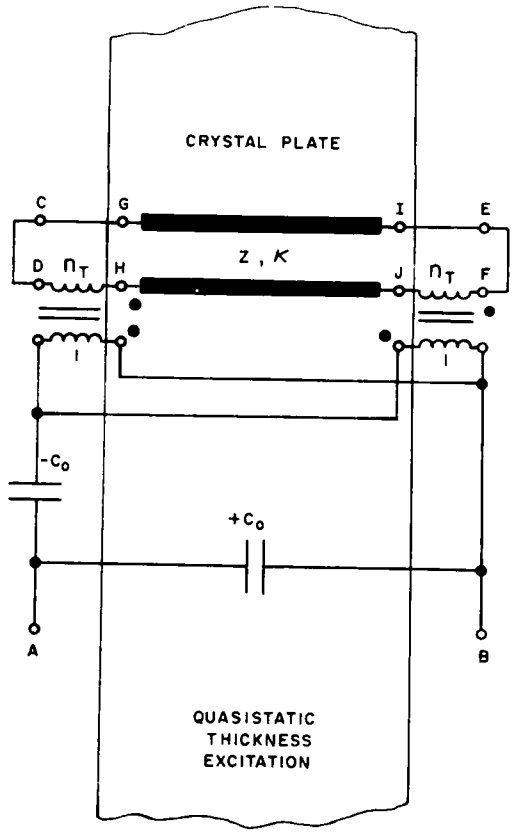



Figure 4. Exact equivalent network for thickness excitation of a single, simple thickness mode, within the quasistatic approximation.

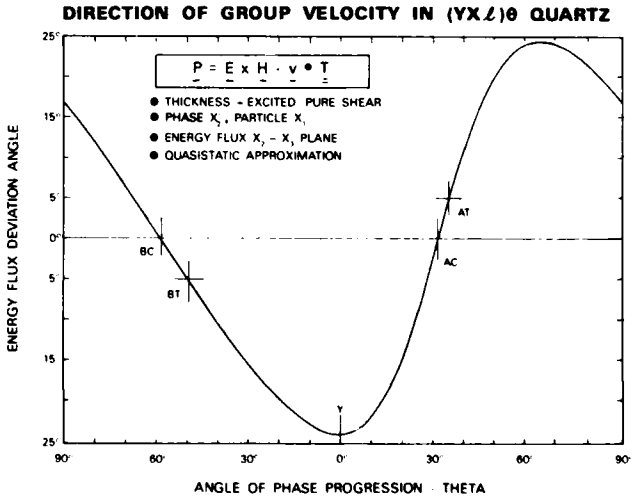



Figure 5. Energy flux deviation angle for rotated-Y-cuts of quartz; quasistatic approximation.

is seen in Fig. 6 to be well under one half of a degree at most.

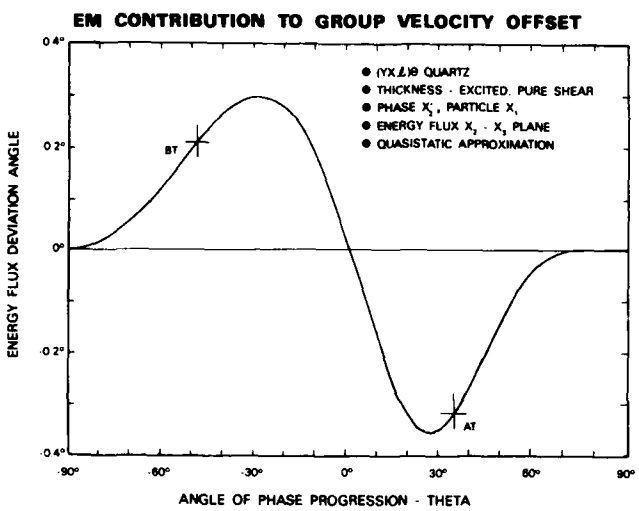



Figure 6. Electromagnetic contribution to the energy flux deviation angle for rotated-Y-cuts of quartz; quasistatic approximation.

Figure 7 gives the lateral field piezocoupling factors $\underline{k}(\theta)$ for each of the three simple thickness modes of rotated-Y-cut quartz when the exciting electric field is directed along the digonal axis. It is seen that for the AT cut, with $\theta = 35.25°$, the pure shear mode, (designated "c mode"), has a value of zero; it cannot be driven. The corresponding graphs for LE with field normal to the digonal axis are seen in Fig. 8. It is seen that the c mode can be driven with this field direction in the AT cut. The situations are identical for the BT cut at $\theta = -49.20°$. The crossover in the mode designations at $\theta \cong -24°$ leads to calling the pure-shear mode in the BT cut the "b mode"; it is still LE-driven by an $X_3'$- directed electric field.

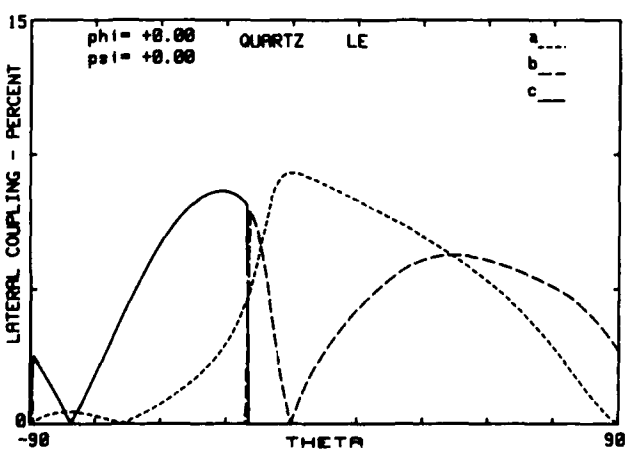



Figure 7. Lateral-field piezoelectric coupling coefficients for the three simple thickness modes of rotated-Y-cut quartz plates with driving electric field along the digonal axis.

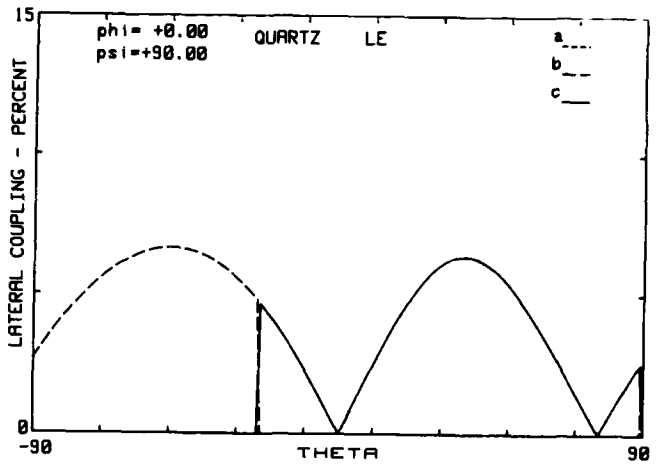



Figure 8. Lateral-field piezoelectric coupling coefficients for the three simple thickness modes of rotated-Y-cut quartz plates with driving electric field normal to the digonal axis.

Lateral-field coupling factors for AT-cut quartz are plotted as function of azimuth about the thickness direction in Fig. 9. Psi equals zero corresponds to the driving field along the digonal ($X_1$) axis. Additional curves for the BT cut, etc., are given in [20].

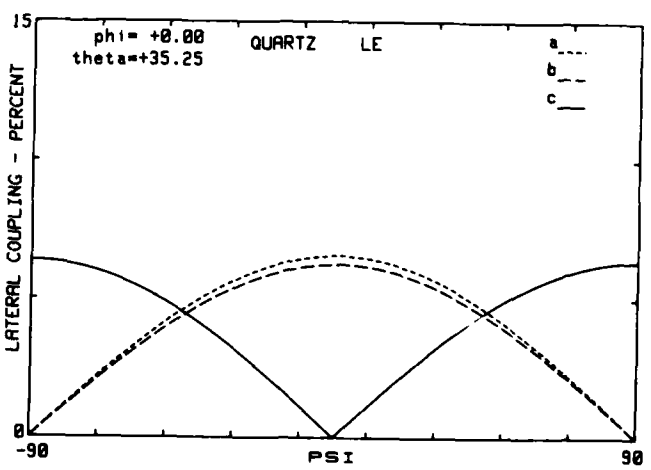



Figure 9. Lateral-field piezoelectric coupling coefficients for the three simple thickness modes of an AT-cut quartz plate as function of azimuth.

Figure 10 corresponds to Fig. 4 when the field is lateral. Notice the reorientation of the static capacitor and the terminals. The piezotransformer turns ratio $n_L$ is directly proportional to $\underline{k}$. In both Figs. 4 and 10, a single mode is represented; extension to two or three modes is discussed in [20] and references therein. The inductors $L_1$ and $L_2$ in Fig. 10 represent the effects of inertial mass-loading on the plate major surfaces, as might occur from depositions placed thereon for the purpose of frequency trimming, control of unwanted modes,

or alteration of other of the resonator's environmental coefficients.

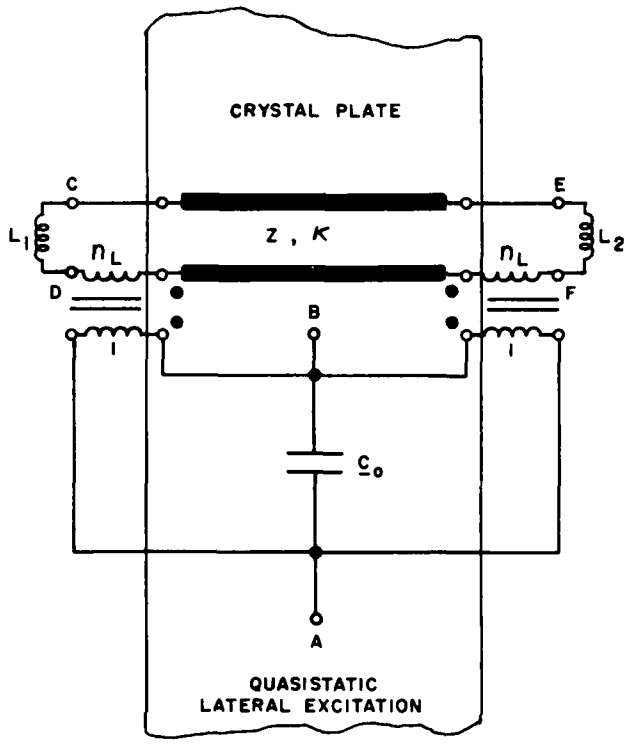



Figure 10. Exact equivalent network for lateral excitation of a single, simple thickness mode, within the quasistatic approximation.

Maxwell's Equations for Plane Harmonic Waves

Maxwell's equations are specialized in Fig. 11 for the case of plane, harmonic waves. For nonmagnetic media such as quartz, the permeability tensor takes the simple form $\mu_{ij} = \mu_o \delta_{ij}$. Thus, irrespective of the medium, the field vectors $\underline{D}$ and $\underline{B}$ will be perpendicular to the propagation wavevector $\underline{\kappa}$, and, furthermore, $\underline{D}$, will be perpendicular to $\underline{H}$ and $\underline{B}$ will be perpendicular to $\underline{E}$. Since we have taken $\underline{B} = \mu_o \underline{H}$, this additionally makes $\underline{H}$ parallel to $\underline{B}$. The vectoral relations are shown in Fig. 11.

Plane electromagnetic waves incident at an angle to a plane interface, separating two isotropic dielectrics, are partially reflected and partially refracted. The two cases that arise are depicted in Fig. 12. A distinction is made according to whether the $\underline{E}$ field incident is perpendicular to the plane of incidence (s = senkrecht case), or in the plane of incidence (p = parallel case). When the angle of incidence is zero, one has normal incidence and the distinction between the cases vanishes [21]-[26].

Figure 13 plots the power reflection coefficients for the p and s cases, as function of angle of incidence, for two dielectric materials, BK7 glass, and high-purity, semi-insulating gallium arsenide. For normal incidence, the normalized power reflection coefficient P is

$$P = (n-1)^2/(n+1)^2.$$

# MAXWELL'S EQUATIONS

$$\nabla \times \underline{H} = + \dot{\underline{D}} \qquad \nabla \times \underline{E} = -\underline{B}$$

PLANE WAVES : $\exp[i(\omega t - \underline{\kappa} \cdot \underline{x})]$

$$\underline{\kappa} \times \underline{H} = -\omega \underline{D} \qquad \underline{\kappa} \times \underline{E} = \omega \underline{B}$$

$\underline{D}, \underline{B}, \perp \underline{\kappa}$ ; $\underline{D} \perp \underline{H}$ and $\underline{B} \perp \underline{E}$

If $\underline{B} = \mu_0 \underline{H}$, $\underline{H} // \underline{B}$ :

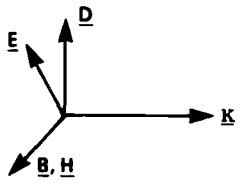
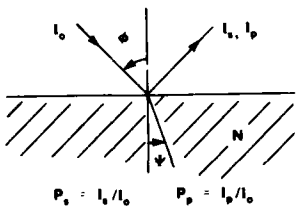



Figure 11. Maxwell's equations, specialized for plane harmonic wave propagation.

## POWER REFLECTION COEFFICIENTS

SNELL $\quad N = \dfrac{\sin\phi}{\sin\psi}$




FRESNEL

$P_s = I_s/I_o \qquad P_p = I_p/I_o$

$$P_s = \frac{\sin^2(\phi-\psi)}{\sin^2(\phi+\psi)} \qquad P_p = \frac{\tan^2(\phi-\psi)}{\tan^2(\phi+\psi)}$$

$$R = I_s/I_p = P_s/P_p = \frac{\cos^2(\phi-\psi)}{\cos^2(\phi+\psi)}$$

Figure 12. Power reflection formulas for optical plane waves incident obliquely upon a plane, isotropic interface.

The index of refraction n is just the square root of the dielectric permittivity ratio of the two media.

Plane electromagnetic wave propagation within uniaxial, dielectric crystals presents another complexity [27]-[31]. The relations from Fig. 11 are still valid, but it is found that the orientation of the electric vectors with respect to the optic axis of the crystal leads to two waves with differing speeds. Figure 14 shows the two cases. The first, where $\underline{D}$ and $\underline{E}$ are normal to the optic axis, is referred to as the "ordinary" wave or ray. The wave speed is isotropic for all directions of the propagation wavevector. The second case, where

B and H are normal to the optic axis, is referred to as the "extraordinary" wave or ray. Its wave speed is a function of the angle between the D vector and the optic axis, equaling the speed of the ordinary wave when D is perpendicular to the optic axis.

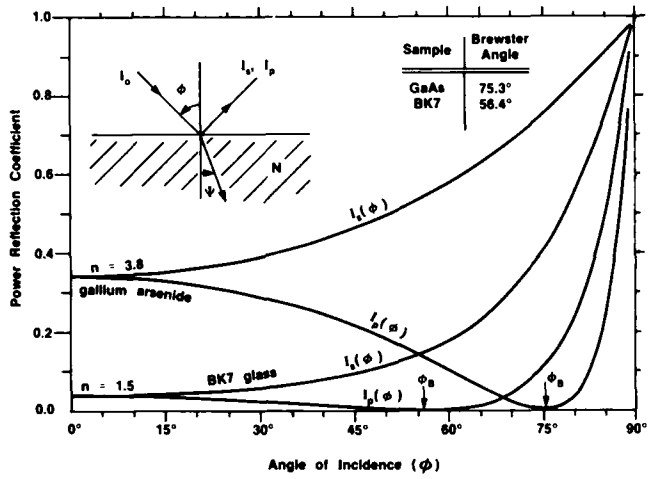



Figure 13. Power reflection coefficients versus incidence angle for two dielectrics.

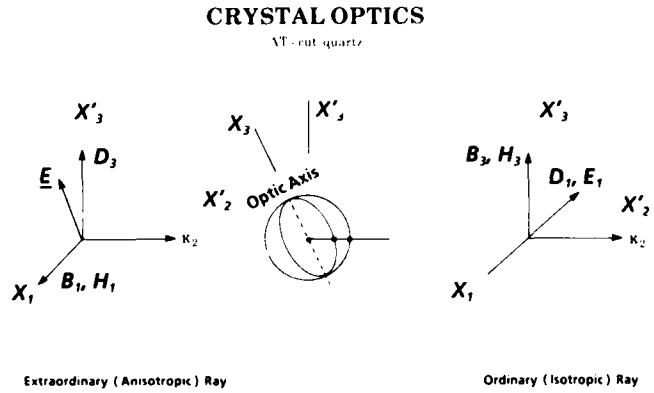



Figure 14. Plane electromagnetic wave propagation in a positive uniaxial crystal.

## Mindlin's Case

Mindlin's solution can now be sketched by putting together the parts discussed briefly above. Mindlin [16], [17] considered an AT-cut quartz plate, driven by mechanical tractions on the major surfaces, and showed that the piezoelectric coupling led to electromagnetic radiation from the plate. [He later also showed that an isotropic, elastic body with a polarization gradient could also radiate [32]. We do not consider this situation].

We start with a rotated-Y-cut plate of quartz, and consider the equations that describe the full coupling of the elastic and electromagnetic fields. The full matrix is seen in Fig. 15. This is to be compared with Fig. 2 for the elastic case.

In Fig. 15 the variables have been rearranged to sort them into two portions. The 4-by-4 in the lower right corresponds to the problem which, with boundary conditions added, is the problem Mindlin solved. It consists of one quasi-acoustic (or quasi-elastic) mode coupled to one quasi-electromagnetic mode. Furthermore, the electromagnetic (EM) mode is the one that, when uncoupled from the elastic mode, constitutes the extraordinary wave seen in Fig. 14. In the absence of coupling with the EM mode, the acoustic mode is the pure shear mode in rotated-Y-cut quartz.

In the center of the Fig. 15 matrix is the uncoupled relation $B_2 = \mu_0 H_2$ which is not further considered.

| | $S_2$ | $S_4$ | $E_1$ | $H_3$ | $H_2$ | $S_6$ | $E_2$ | $E_3$ | $H_1$ |
|---|---|---|---|---|---|---|---|---|---|
| $T_2$ | $c_{22}$ | $c_{24}$ | $-e_{14}$ | 0 | 0 | 0 | 0 | 0 | 0 |
| $T_4$ | $c_{24}$ | $c_{44}$ | $-e_{14}$ | 0 | 0 | 0 | 0 | 0 | 0 |
| $D_1$ | $+e_{12}$ | $+e_{14}$ | $\varepsilon_{11}$ | $\frac{-K_2}{\omega}$ | 0 | 0 | 0 | 0 | 0 |
| $B_3$ | 0 | 0 | $\frac{-K_2}{\omega}$ | $\mu_0$ | 0 | 0 | 0 | 0 | 0 |
| $B_2$ | 0 | 0 | 0 | 0 | $\mu_0$ | 0 | 0 | 0 | 0 |
| $T_6$ | 0 | 0 | 0 | 0 | 0 | $c_{66}$ | $-e_{26}$ | $-e_{36}$ | 0 |
| $D_2$ | 0 | 0 | 0 | 0 | 0 | $+e_{26}$ | $\varepsilon_{22}$ | $\varepsilon_{23}$ | 0 |
| $D_3$ | 0 | 0 | 0 | 0 | 0 | $+e_{36}$ | $\varepsilon_{23}$ | $\varepsilon_{33}$ | $\frac{+K_2}{\omega}$ |
| $B_1$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | $\frac{+K_2}{\omega}$ | $\mu_0$ |

Figure 15. Matrix for elastoelectrodynamics problem.

The 4-by-4 matrix in the upper left of Fig. 15 represents the two remaining simple thickness acoustic (the quasi-shear and quasi-extensional) modes coupled to the EM wave that, in the absence of such coupling, becomes the ordinary EM wave. This second problem is similar to Mindlin's case [33].

Mindlin [17] reduced the elastic and electromagnetic equations for plane wave propagation in rotated-Y-cut quartz, corresponding to the pure shear wave fully coupled to the extraordinary optical wave, to the three equations at the top of Fig. 16. By eliminating the variable $E_2$, one arrives at the two modified equations that are given below those equations.

Assuming plane wave propagation, these equations give the dispersion relation shown in Fig. 17. The single factor that expresses the coupling between the acoustic and optic waves is $\bar{k}_{36}^2$. In Fig. 18 are given the approximate roots, and application to AT-cut quartz. From the approximation used it is seen that the factor $\frac{1}{2}\bar{k}_{36}$ splits the two normalized roots ($v_{ac}/\bar{v}$) and ($v_{op}/\bar{c}$) equally. The piezocoupling is exceedingly small, as the example shows. The acoustic wave is slowed by 38 angstroms per second, while the optic wave is speeded up by 0.16 mm per second.

$$c_{66}\, u_{1,22} - e_{26}\, E_{2,2} - e_{36}\, E_{3,2} = \rho\, \ddot{u}_1$$

$$e_{26}\, u_{1,2} + \varepsilon_{22} E_2 + \varepsilon_{23} E_3 = 0$$

$$e_{36}\, \ddot{u}_{1,2} + \varepsilon_{23} \ddot{E}_2 + \varepsilon_{33} \ddot{E}_3 = \mu^{-1}_0 E_{3,22}$$

Eliminate $E_2$:

$$\bar{c}_{66}\, u_{1,22} - \bar{e}_{36} E_{3,2} = \rho\, \ddot{u}_1$$

$$\bar{e}_{36}\, \ddot{u}_{1,2} + \bar{\varepsilon}_{33} \ddot{E}_3 = \mu^{-1}_0 E_{3,22}$$

where $\bar{c}_{66} = c_{66} + e^2_{26} / \varepsilon_{22}$

$$\bar{e}_{36} = e_{36} - e_{26} \varepsilon_{23} / \varepsilon_{22}$$

$$\bar{\varepsilon}_{33} = \varepsilon_{33} - \varepsilon^2_{23} / \varepsilon_{22} = \det \varepsilon / \varepsilon_{22}$$

**Figure 16. Reduction of equations.**

PLANE WAVE $\exp[i(\omega t - \kappa_2 x_2)]$ ; LEADS TO

DISPERSION RELATION $\kappa^4 - \zeta_2 \kappa^2 + \zeta_3 = 0$

where $\zeta_2 = (\omega / \bar{v})^2 [1 + \bar{\beta}^2 (1 + \bar{k}^2_{36})]$

$\zeta_3 = (\omega / \bar{v})^2 (\omega / \bar{c})^2$

and $\bar{v}^2 = \bar{c}_{66} / \rho$ ; $\bar{c}^2 = (\mu_0 \bar{\varepsilon}_{33})^{-1}$ ; $\bar{\beta} = \bar{v} / \bar{c}$

$\bar{k}^2_{36} = \bar{e}^2_{36} / (\bar{\varepsilon}_{33} \bar{c}_{66})$

ALTERNATIVE FORM

$$(v^2 / \bar{v}^2 - 1)(1 - \bar{c}^2 / v^2) = \bar{k}^2_{36}$$

$\bar{c}$: velocity of extraordinary optical wave

$\bar{v}$: velocity of slow shear acoustic wave (piezo-stiffened)

**Figure 17. Development of dispersion relation.**

Approximations : $(\bar{\beta} \bar{k}_{36} \ll 1)$

Acoustic branch $v^2_{ac} \approx \bar{v}^2 (1 + \bar{\beta}^2 \bar{k}^2_{36}) \quad < \bar{v}^2$

Optical branch $v^2_{op} \approx \bar{c}^2 (1 - \bar{\beta}^2 \bar{k}^2_{36}) \quad > \bar{c}^2$

equal splitting

$v^2_{ac}$ correct to $O\ \bar{\beta}^2$ ; $v^2_{op}$ correct to $O\ \bar{\beta}^4$

AT-CUT QUARTZ $\bar{c} = 1.40 \times 10^3$ m s $(n_e = c / \bar{c} = 2.14)$

$\bar{v} = 3.32 \times 10^3$ m s; $\bar{\beta} = 2.37 \times 10^{-1}$; $\bar{k}_{36} = 6.36\%$

Elastoelectromagnetic coupling factor $\bar{\beta} \bar{k}_{36} = 1.5$ ppm

$v_{ac} = \bar{v} (1 - (\bar{\beta} \bar{k}_{36})^2 / 2) = \bar{v} - 38$ Å second

$v_{op} = \bar{c} (1 + (\bar{\beta} \bar{k}_{36})^2 / 2) = \bar{c} + 0.16$ mm second

**Figure 18. Approximations for the wave velocities.**

The situation that occurs may be modeled by transmission lines coupled at the plate boundaries [18]-[19]. When the surfaces are coated with conductors, so that radiation is prevented, the situation is that shown in Fig. 19. The more general situation is shown in Fig. 20, where the interface networks transform from the normal coordinates for each mode to the purely electromagnetic and purely mechanical variables at the plate surfaces. In Fig. 21 is given the network representation of Professor Mindlin's solution to the problem. Within the crystal, two modes propagate. One is essentially acoustic, and the other is essentially electromagnetic. At the boundaries of the plate, these are coupled to the single optical mode that propagates in free space, representing the radiating electromagnetic energy. In Mindlin's solution the plate is driven by mechanical surface tractions, which would be represented by voltage sources attached at the mechanical ports of Fig. 20. The electromagnetic ports are attached to transmission lines leading away from the boundaries, as in Fig. 21.

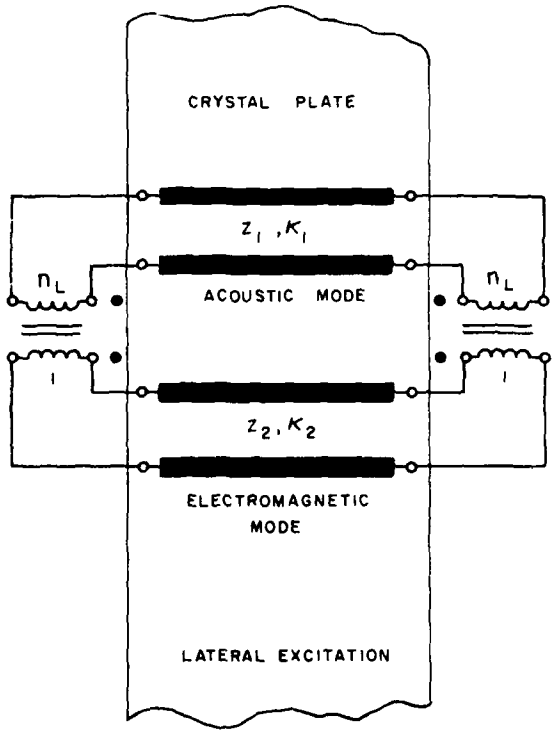



**Figure 19. Exact equivalent network for the coupling of modes when radiation is prevented.**

## Conclusion

This brief sketch gives only an indication of one of the many beautiful problems that Professor Mindlin solved. It moreover introduces electrical network elements, for which I hope he will forgive me. His solution provides the connection between the approximate QS situation of Fig. 10 and the exact solution of Fig. 19, and resolves the contradiction that appears when lateral-field excitation is modeled. Because a propagating wave in a piezoelectric medium always produces an electric field along the direction of travel ($E_2$), the capacitor plates of $C_o$ in Fig. 10 (which should

extend across the whole thickness) short out this field. In actuality, the capacitor $C_o$ in Fig. 10 is only the QS approximation to the optic (electromagnetic) TL in Fig. 19 or Fig. 20; this TL supports an $E_2$ component. Further considerations relating to the QS approximation are found in [34].

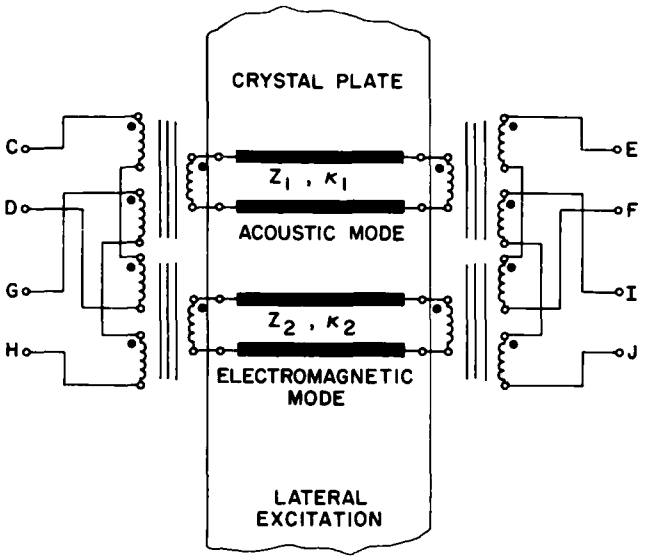
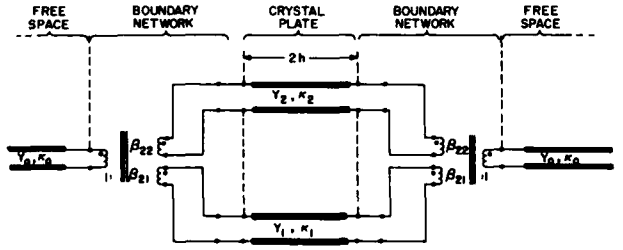



Figure 20. Exact equivalent network representing the acoustic and electromagnetic modes within the crystal plate and the transformation to boundary variables.




ELECTROMAGNETIC-MECHANICAL COUPLING IN A ROTATED Y-CUT QUARTZ PLATE WITH TRACTION-FREE BOUNDARIES.

Figure 21. Exact equivalent network representing Mindlin's problem. The optical mode propagating outside the crystal couples at the boundaries to the two plate modes.

## References


[1] J. V. Atanasoff and P. J. Hart, "Dynamical Determination of the Elastic Constants and Their Temperature Coefficients for Quartz," Phys. Rev., Vol. 59, Jan 1941, pp. 85-96.

[2] A. W. Lawson, "Comments on the Elastic Constants of Alpha-Quartz," Phys. Rev., Vol. 59, 1941, pp. 838-839.

[3] A. W. Lawson, "The Vibration of Piezoelectric Plates," Phys. Rev., Vol. 62, July 1942, pp. 71-76.

[4] R. Bechmann, "Über Dickenschwingungen piezoelektrischer Kristallplatten," Archiv der Elektrischen Übertragung, Vol. 6, Sep 1952, pp. 361-368.

[5] R. Bechmann, "Über Dickenschwingungen piezoelektrischer Kristallplatten," Nachtrag, Archiv der Elektrischen Übertragung, Vol. 7, July 1953, pp. 354-356.

[6] H. F. Tiersten, "Thickness vibration of piezoelectric plates," J. Acoust. Soc. Amer., Vol. 35, No. 1, January 1963, pp. 53-58.

[7] J. F. Havlice, W. L. Bond, and L. B. Wigton, "'Elastic' Poynting vector in a piezoelectric medium," IEEE Trans. Sonics Ultrason., Vol. SU-17, No. 4, October 1970, pp. 246-249.

[8] G.S. Kino, Acoustic Waves: Devices, Imaging, and Signal Processing, Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[9] J. J. Kyame, "Wave Propagation in Piezoelectric Crystals," J. Acoust. Soc. Amer., Vol. 21, No. 3, May 1949, pp. 159-167.

[10] H. Pailloux, "Piézoelectricité Calcul des Vitesses de Propagation," J. Phys. Rad., Vol. 19, May 1958, pp. 523-525.

[11] V. Alda, K. Hruška, and J. Tichý, "Propagation of Waves Through Infinite Piezoelectric Medium," Czech. J. Phys., Vol. 13, 1963, pp. 345-366. (In Czech).

[12] K. Hruška, "The Rate of Propagation of Ultrasonic Waves in ADP in Voigt's Theory," Czech. J. Phys., Vol. B16, 1966, pp. 446-453.

[13] K. Hruška, "Relation Between the General and the Simplified Condition for the Velocity of Propagation of Ultrasonic Waves in a Piezoelectric Medium," Czech. J. Phys., Vol. B18, 1968, pp. 214-221.

[14] J. J. Kyame, "Conductivity and Viscosity Effects on Wave Propagation in Piezoelectric Crystals," J. Acoust. Soc. Amer., Vol. 26, No. 6, Nov 1954, pp. 990-993.

[15] A. R. Hutson and D. L. White, "Elastic Wave Propagation in Piezoelectric Semi-conductors," J. Appl. Phys., Vol. 33, No. 1, 1962, pp. 40-47.

[16] R. D. Mindlin, "Theory of vibrations of plates," Proc. 26th Annual Frequency Control Symposium, June 1972, p. 84.

[17] R. D. Mindlin, "Electromagnetic radiation from a vibrating quartz plate," Int. J. Solids Structures, Vol. 9, 1973, pp. 697-702.

[18] A. Ballato, "Networks for crossed-field and in-line excitation of bulk and surface acoustic waves," Proc. Symp. Optical and Micro-Electronics, Polytechnic Institute of New York MRI Symp. Proc. Vol. 23, April 1974, pp. 599-615.

[19] A. Ballato, "Bulk and surface acoustic wave excitation and network representation," Proc. 28th Annual Frequency Control Symposium, May 1974, pp. 270-279.

[20] A. Ballato, E. R. Hatch, M. Mizan, and T. J. Lukaszek, "Lateral field equivalent networks and piezocoupling factors of quartz plates driven in simple thickness modes," IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control, Vol. UFFC-33, No. 4, July 1986, pp. 385-392.

[21] A. R. von Hippel, Dielectrics and Waves, M.I.T. Press, Cambridge, MA, 1954.

[22] A. Sommerfeld, Optics, Lectures on Theoretical Physics, Vol. IV, Academic Press, New York, 1964.

[23] O. S. Heavens, Optical Properties of Thin Solid Films, Dover Publications, New York, 1965.

[24] Handbook of Optics, (W. G. Driscoll, ed.), McGraw-Hill, New York, 1978.

[25] E. Collett, "Digital Refractometry," Optics Commun., Vol. 63, No. 4, 15 August 1987, pp. 217-224.

[26] W. A. Shurcliff, Polarized Light, Harvard Univ. Press, 1962.

[27] R. E. Collin, Field Theory of Guided Waves, McGraw-Hill, New York, 1960.

[28] M. Born and E. Wolf, Principles of Optics, 3rd Ed., Pergamon Press, New York, 1964.

[29] E. A. Wood, Crystals and Light, Van Nostrand Co., Princeton, New Jersey, 1964.

[30] H. J. Juretschke, Crystal Physics, W. A. Benjamin, Reading, Massachusetts, 1974.

[31] F. Wooten, "Reflectivity of Uniaxial Absorbing Crystals," Applied Optics, Vol. 23, No. 23, 1 Dec 1984, pp. 4226-4227.

[32] R. D. Mindlin, "Electromagnetic radiation from a vibrating, elastic sphere," Technical Report No. 28, Office of Naval Research, Contract No. N00014-67A-0108-0027, January 1974, 25 pp. AD-780 244.

[33] A. Ballato, forthcoming.

[34] P. C. Y. Lee, "Electromagnetic radiation from an AT-cut quartz plate under lateral-field excitation," IEEE Ultrasonics Symposium Proc., October 1988, in press.

42nd Annual Frequency Control Symposium - 1988

# ACCELERATION INSENSITIVITY OF THICKNESS FREQUENCIES OF DOUBLY ROTATED QUARTZ CRYSTAL DISKS

P. C. Y. Lee and M. S. H. Tang
Department of Civil Engineering and Operations Research
Princeton University
Princeton, NJ 08544

## Summary

Initial fields of displacement gradients and strains caused by steady accelerations of arbitrary direction are calculated from finite element solutions of Mindlin's two-dimensional, first-order equations of equilibrium for crystal plates.[1] The frequency equation of incremental thickness vibrations of circular crystal disks under initial stresses is obtained by the extension of a recently obtained solution for energy-trapped thickness vibrations[2] and solved by a perturbation method.

Acceleration sensitivity of SC- and AT-cuts of quartz disks are studied systematically. It is found: (1) Among various support configurations, a six-point mount gives the lowest acceleration sensitivity ($\Gamma$) for SC-cut, and three-point and four-point mounts give the lowest values of $\Gamma$ for an AT-cut. (2) For four-point mounted SC-cut, $\alpha \cong 8°$ gives the lowest value of $\Gamma$. For three-point and four-point mounted AT-cut, $\alpha = 0$ gives the lowest values of $\Gamma$. (3) For disks with diameter d = 14 mm and thickness 2b = 0.2 to 1.1 mm. acceleration sensitivity varies linearly with respect to the plate thickness for SC-cuts, but $\Gamma$ is practically independent of thickness 2b for the AT-cut. For 2b < 0.2 mm, $\Gamma$ is approximately inversely proportional to the plate thickness for both cuts. (4) Values of $\Gamma$ are calculated for doubly rotated cuts with four-point mount ($\alpha = 0$) along the upper locus of zeros of first-order temperature coefficient of frequency of quasi-shear mode c, i.e., along $\theta = 35.25° - (11/180)\phi$ with $\theta > 0$.[3] In the neighborhood of AT ($\phi = 0$, $\theta = 35.25°$) and SC ($\phi = 21.43°$, $\theta = 34°$) cuts, loci of minimum acceleration sensitivity as function of $\phi$ and $\theta$ are obtained. It is found that for four-point mounted circular disks the acceleration-insensitive cut is at $\phi = 0$, $\theta = 34.20°$. This cut in brief may be called AI(4)-cut. (5) Support vibrations of four-point mounted crystal disks are studied for various types of support structures by the finite element method. Resonant frequencies and mode shapes for the first five modes are obtained for SC- and AT-cuts.

## 1. Introduction

In a previous paper,[4] results of parametric study on the acceleration sensitivity of SC-cut quartz disk were reported and the background of related previous investigations were reviewed.

In the present paper, we adopt the same notations and definitions for acceleration sensitivities by letting

$$\Gamma = \sqrt{\Gamma_\perp^2 + \Gamma_n^2} = \sqrt{\Gamma_1^2 + \Gamma_2^2 + \Gamma_3^2} \qquad (1)$$

where $\Gamma$ is the total acceleration sensitivity (per g), $\Gamma_\perp$ and $\Gamma_n$ are, respectively, the thickness and lateral acceleration sensitivities, and $\Gamma_i$, i = 1,2,3, represents changes of frequencies due to 1g acceleration in the $x_i$-axis direction. Although $\Gamma$ can be computed for thickness modes a, b, and c, it shall mean the acceleration sensitivity of mode c only, unless it is specifically stated otherwise.

We consider doubly rotated, circular quartz disks with diameter d = 14.0 mm and thickness 2b = 1.1469 mm vibrating in the fundamental thickness modes. The disk is supported by a number of metal ribbons at the edge of the disk. The ribbon is made of molybdenum ($E = 27.579 \times 10^9$ dynes/mm$^2$, $\nu = 0.32$) with length $\ell = 1.143$ mm and rectangular cross-section (0.0177 mm × 1.524 mm). The contact condition between the plate and the ribbon is assumed to be "rigid", i.e., it can transmit both moments and forces.

## 2. Acceleration Sensitivity versus Support Configuration

To examine the effect of different support configurations on the acceleration sensitivity, six different mounting arrangements as shown in Table I are considered. For a three-point mounted AT-cut disk, $\Delta f/f_0$ vs. $\psi_1$ and $\Delta f/f_0$ vs. $\psi_2$ are calculated and plotted in Figs. 1 and 2, respectively, from which we have

$$\Gamma = \sqrt{\Gamma_\perp^2 + \Gamma_n^2}$$

$$= \sqrt{(0.65 \times 10^{-11})^2 + (2.85 \times 10^{-11})^2}$$

$$= 0.29 \times 10^{-10}/g$$

Similar calculations are made for AT- as well as SC-cut disks with four, six, eight, twelve and forty-four point mounts. Results are listed in Table I. It can be seen that for an SC-cut, a six-point mounted disk gives the lowest value of $\Gamma = 0.45 \times 10^{-10}/g$. For AT-cut, both the three-point and four-point mounted disks have the lowest value of $\Gamma \cong 0.28 \times 10^{-10}/g$

## 3. Acceleration Sensitivity versus Mount Orientation

Acceleration sensitivity $\Gamma$ as a function of mount orientation angle $\alpha$ for a four-point mounted SC-cut disk is shown in Fig. 3 which is slightly different from the curve in Fig. 7 of Ref. 4 due to presently modified frequency equation for taking into account the energy-trapped thickness modes. It may be seen from Fig. 3 that at $\alpha \cong 8°$, $\Gamma$ attains a minimum value of $0.911 \times 10^{-10}/g$. The predicted values of $\Gamma$ for $\alpha = 0°$ to 45° are replotted in Fig. 4 in order to compare with measured values by Filler, Kosinski, and Vig[5] (represented by solid dots) and measured values by Tuladhar[6] (represented by heavy vertical lines at $\alpha = 0°$ and 45°). It may be seen that the predicted curve appears to be near the lower bound of experimental dots. In comparing with Tuladhar's data one should keep in mind that his measurements were for 3rd and 5th overtone crystals in the frequency ranges 75 to 80 MHz and 115 to 120 MHz and with different types of crystal holders.

The $\Gamma$ vs. $\alpha$ curve for a four-point mounted AT-cut disk is shown in Fig. 5 from which we see that $\Gamma$ attains a minimum value at $\alpha = 0$.

In Fig. 6, the solid line represents the $\Gamma$ vs. $\alpha$ curve for a three-point mounted AT-cut disk while the dotted line represents $\Gamma$ vs. $\alpha$ curve for the same disk with four-point mount. Tuladhar's measured values for crystal types A, C, D of three-point mount at $\alpha = 180°$ and for crystal type F for four-point mount at $\alpha = 0°$ are also given in Fig. 6 for comparison. It can be seen that both three-point and four-point mounts attain minimum value of $\Gamma$ at $\alpha = 0$ (as listed in Table I), but the four-point mounting configuration is much less sensitive to the changes of orientation angle $\alpha$.

## 4. Acceleration Sensitivity versus Plate Thickness

Acceleration sensitivity of an AT-cut disk is calculated as a function of plate thickness 2b and is plotted in Figs. 7 and 8 for four-point and three-point mounts, respectively.

CH2588-2/88/0000-014 $1.00 © 1988 IEEE

It may be seen that the lateral acceleration sensitivity $\Gamma_{\parallel}$ is practically independent of plate thickness. In the case of a four-point mounted SC-cut, $\Gamma_{\parallel}$ is found to be linearly proportional to the plate thickness (see Fig. 8, Ref. 4). However, the thickness acceleration sensitivity $\Gamma_{\perp}$ is found to be approximately inversely proportional to the plate thickness and becomes predominant when $2b \leq 0.2$ mm for both AT- and SC-cuts.

## 5. Acceleration Sensitivity versus Plate Orientation

Acceleration sensitivity $\Gamma$ for doubly rotated cuts with four-point mount and $\alpha = 0$ are computed for plate orientation angles $\phi$ and $\theta$ following the upper locus of zeros of first-order temperature coefficient of frequency of quasi-shear mode c, i.e., following the straight line $\theta = 35.25° - (11/180)\phi$ and $\theta > 0$.[3] The resulting $\Gamma$ vs. $\phi$ curve is plotted in Fig. 9 which shows that along the upper locus line AT-cut has the lowest value of $\Gamma$, although SC-cut has the minimum frequency changes under isotropic in-plane (thin film) stress.[5]

Further calculations of $\Gamma$ for four-point mounted disks are made for $\phi = 0, \pm1°, \pm2°, \pm4°$ and $\theta = 30°$ to $40°$ and results are given in Fig. 10. By connecting the minimum points of these curves, the locus of $\Gamma$ minimum is obtained as shown by the dotted line in Fig. 10. Thus we locate a local minimum of $\Gamma = 0.2 \times 10^{-11}/g$ at $\phi = 0$ and $\theta = 34.20°$. This acceleration-insensitive cut with a four-point mount may be designated in brief as AI(4)-cut.

Similar calculations of $\Gamma$ are made for $\phi = 18°$ to $28°$ and $\theta = 30°$ to $40°$ in the neighborhood of SC-cut and results are plotted in Fig. 11. It can be seen that values of $\Gamma$ minimum of each curve decreases as $\phi$ increases, but there is no local minimum in the region considered.

## 6. Support Vibrations Affected by Support Structure

Support vibrations of four-point mounted crystal disks with various types of support structures are studied by the finite element method.

Details of a general support structure and dimensions for support types A, B, C, D, and E are given in Fig. 12. We note that for support type A, $\ell_1 = \ell_2 = 0$, i.e., there is no horizontal structural part, while for the other four types of support there is always a non-zero horizontal structural part. It is mainly due to this difference that vibrational modes of the plate with support type A are quite different from those of the plate with support types B, C, D, and E.

The first five vibrational modes, in the order of increasing frequencies, for plate with support type A are depicted in Fig. 13. We see that the first mode is predominantly the first flexural mode of plate with little column deflections, the second and third modes are predominantly flexural modes of columns in the $x_1$ and $x_3$ directions respectively, and the fourth and fifth modes are the second flexural modes of plate in the $x_1$ and $x_3$ directions, respectively.

The first five modes for plate with support types B, C, D, and E are given in Fig. 14. In this case, the first mode is predominantly plate translational mode, the second and third modes are predominantly plate rotational modes about the $x_3$ and $x_1$ axes respectively, and the fourth and fifth modes are flexural modes of columns in the $x_1$ and $x_3$ directions, respectively.

Resonant frequencies corresponding to these five modes for support types A, B, C, D, and E are computed for SC-cut disk and listed in Table II. It can be seen that more flexible are the supports the lower values are the resonant frequencies.

Similar computations are also made for AT-cut plate. Values of resonant frequencies of corresponding modes are very close to those for SC-cut. This is owing to the fact that both cuts have same value of $c_{11}$ and slightly different values for $c_{33}$ and frequencies of flexural vibrations depend mainly on the values of these two elastic stiffness coefficients. Low frequency structural modes had also been studied by Weglein, Hu, and Margagliano by the use of the finite element method.[8]


## Acknowledgment

This work was supported by the U.S. Army Research Office, Contract No. DAAL03-87-K-0125.


## References


1. P. C. Y. Lee and M. S. H. Tang, Proc. 40th Ann. Freq. Cont. Symp., pp. 152-160, 1986.
2. P. C. Y. Lee, X. Guo, and M. S. H. Tang, J. Appl. Phys. 63(6) pp. 1850-1856, 1988 and in IEEE Ultrasonics Symp. Proc., 1987.
3. A. Ballato, "Doubly rotated thickness mode plate vibrators," in Physical Acoustics: Principles and Methods, (W. P. Mason and R. N. Thurston, eds.), Vol. 13, Chap. 5, Academic Press, New York, pp. 115-181, 1977.
4. P. C. Y. Lee and M. S. H. Tang, Proc. 41st Ann. Freq. Cont. Symp. pp. 277-281, 1987.
5. R. L. Filler, J. A. Kosinski and J. R. Vig, Proc. 37th Ann. Freq. Cont. Symp., pp. 265-271, 1983.
6. K. K. Tuladhar, IERE (Brit.) Proc. Int'l. Conf. on Freq. Control and Synthesis, U. Surrey, pp. 67-73, 1987.
7. A. Ballato and M. Mizan, IEEE Trans. Sonics Ultrason., Vol. SU-31, pp. 11-18, 1984.
8. R. D. Weglein, W. C. Hu, and A. F. Margagliano, IEEE Ultrasonics Symp. Proc., pp. 378-382, 1984.


Table I. Acceleration Sensitivity of Quartz Disks for Various Support Configurations

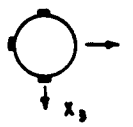
| Configuration | No. of Supports | Acceleration Sensitivity $\Gamma(10^{-10}/g)$ SC-cut | AT-cut |
|---|---|---|---|
|  | 3 | 8.91 | 0.29 |
| | 4 | 2.54 | 0.27 |
| | 6 | 0.45 | 0.85 |
| | 8 | 1.72 | 1.03 |
| | 12 | 1.91 | 0.89 |
| | 44 | 2.29 | 0.81 |

Table II. The first five frequencies of SC–cut quartz disk with four–point support of types A, B, C, D, and E.

| Mode No. | Frequency (kHz) | | | | |
|---|---|---|---|---|---|
| | A | B | C | D | E |
| 1 | 21.77 | 4.69 | 1.39 | 1.30 | 1.16 |
| 2 | 27.04 | 7.27 | 2.21 | 2.00 | 1.80 |
| 3 | 27.47 | 7.27 | 2.21 | 2.00 | 1.80 |
| 4 | 35.45 | 23.11 | 14.98 | 14.42 | 6.49 |
| 5 | 36.23 | 23.25 | 15.00 | 14.43 | 6.49 |

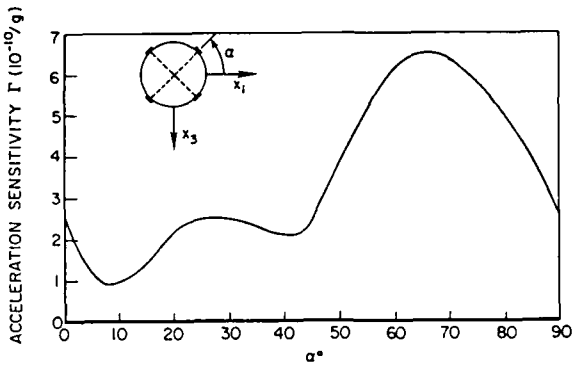



Fig. 3 Acceleration sensitivity $\Gamma$ vs. $\alpha$ for four–point mounted SC–cut resonators.

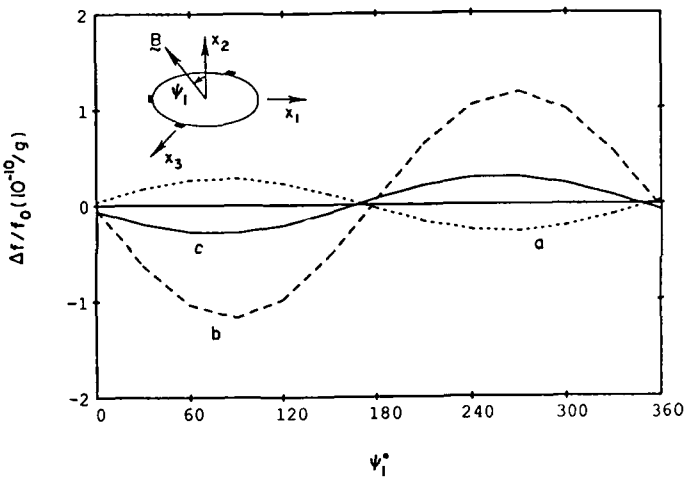



Fig. 1 $\Delta f/f_0$ vs. $\psi_1$ for a three–point mounted AT–cut resonator.

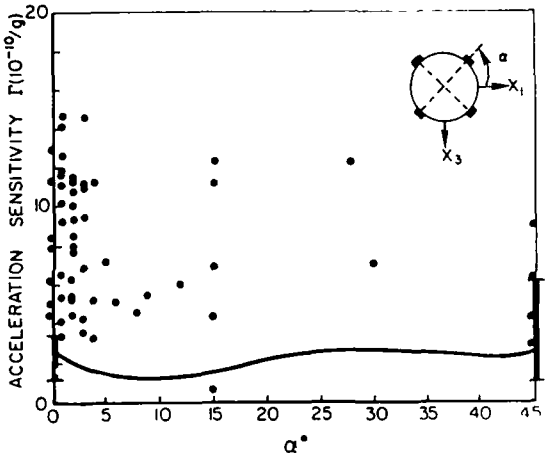



Fig. 4 Comparison of predicted $\Gamma$ vs. $\alpha$ curve with measured values of Filler, Kosinski and Vig (by dots) and of Tuladhar (by vertical lines) for four–point mounted SC–cut resonators.

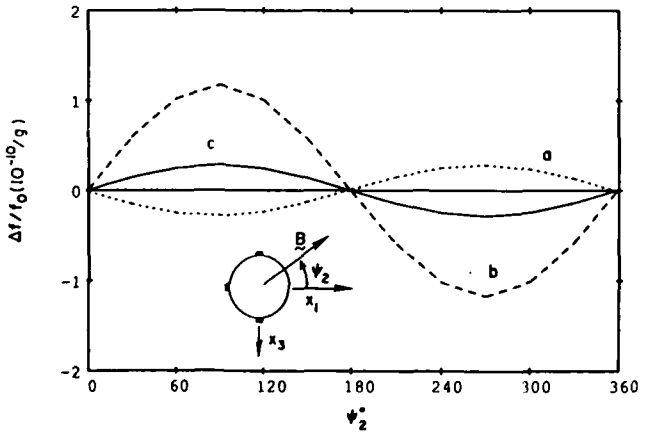



Fig. 2 $\Delta f/f_0$ vs. $\psi_2$ for a three–point mounted AT–cut resonator.

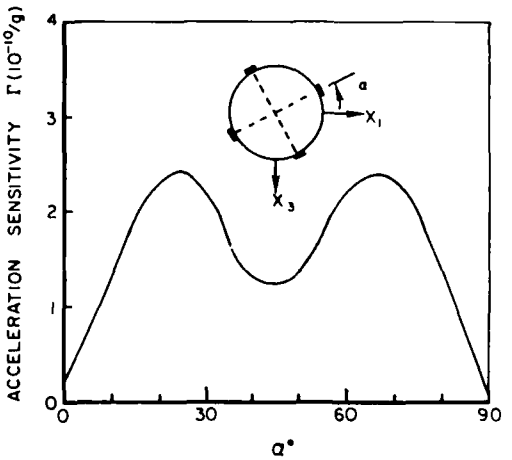



Fig. 5 $\Gamma$ vs. $\alpha$ for four–point mounted AT–cut resonators.

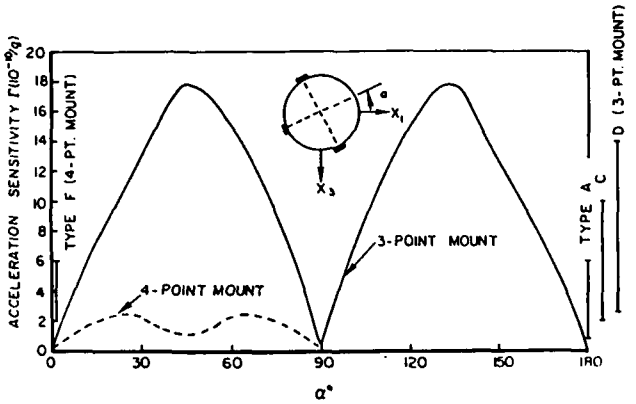



Fig. 6 Γ vs. $\alpha$ for three–point and four–point mounted AT–cut resonators and and measured values by Tuladhar.

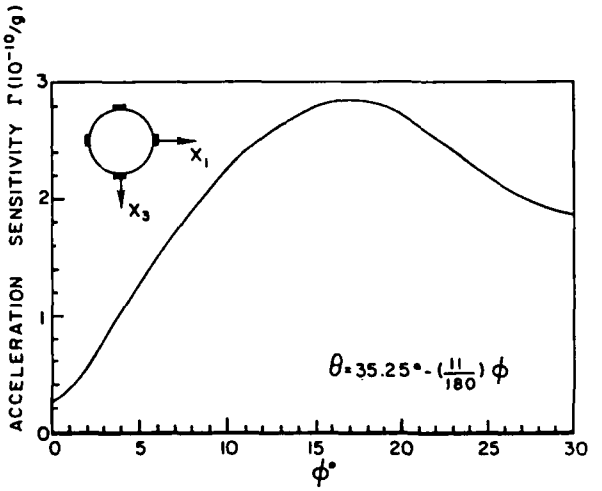



Fig. 9 Γ vs. $\phi$ for four–point mounted disks along the upper locus of zeros of first–order temperature coefficient of frequency of quasi–shear mode c.

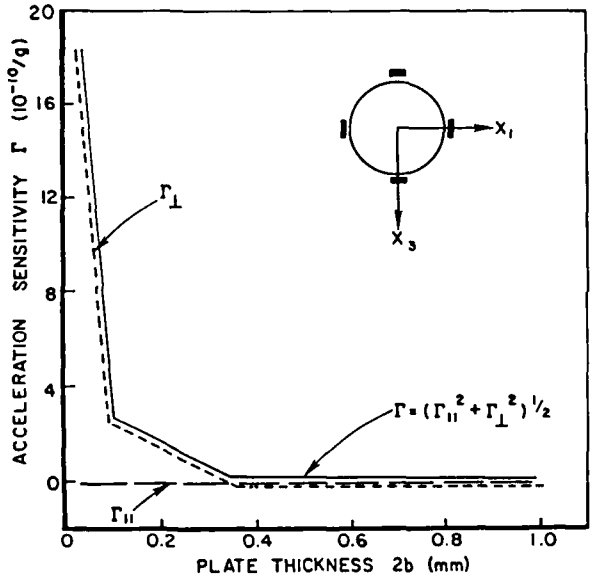



Fig. 7 Γ vs. plate thickness 2b for four–point mounted AT–cut resonators.

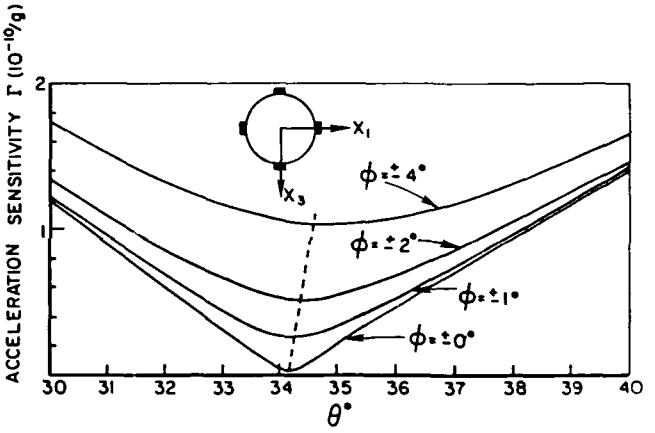



Fig.10 Γ vs. $\theta$ for four–point mounted disks and for $\phi = 0, \pm 1°, \pm 2°$ and $\pm 4°$.

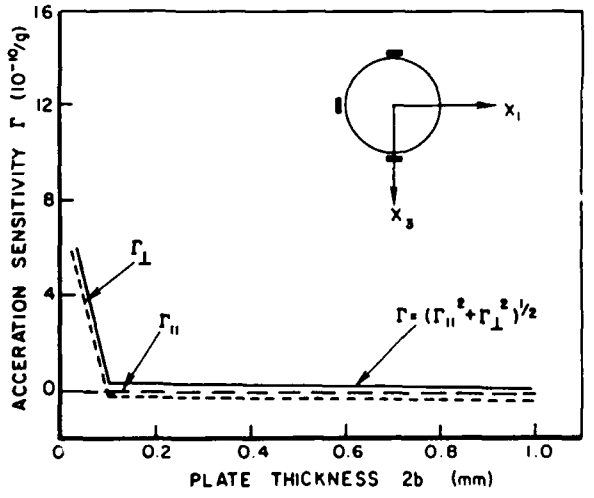



Fig. 8 Γ vs. plate thickness 2b for three–point mounted AT–cut resonators.

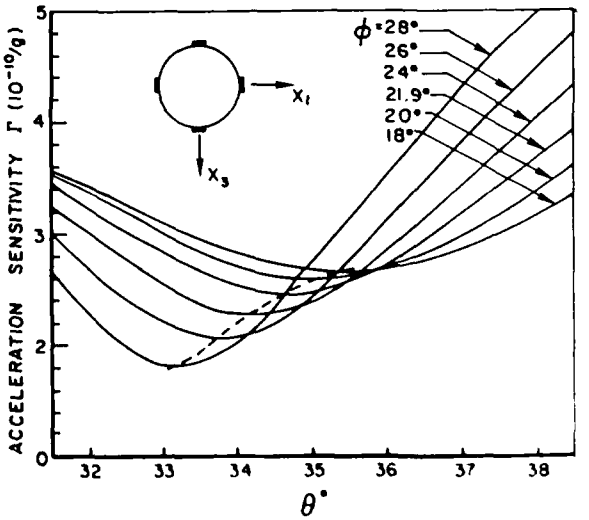
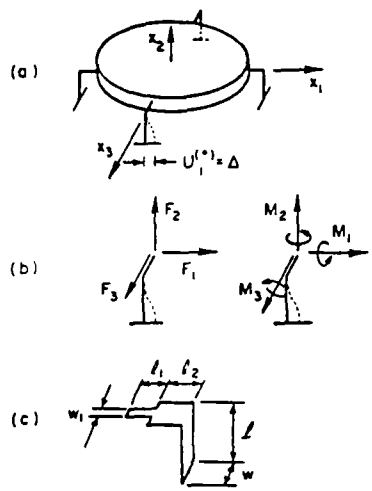



Fig.11 Γ vs. $\theta$ for four–point mounted disks and for $\varphi$ = 18°~28°.




| Support Type | $\ell$ | $\ell_1$ | $\ell_2$ | w | $w_1$ (in mm) |
|---|---|---|---|---|---|
| A | 1.143 | 0 | 0 | 1.524 | 0 |
| B | 1.143 | 0 | 0.2 | 1.524 | 0 |
| C | 1.143 | 0.2 | 0.2 | 1.524 | 0.505 |
| D | 1.143 | 0.4 | 0 | 1.524 | 0.505 |
| E | 1.143 | 0.4 | 0 | 0.761 | 0.505 |

Fig.12 (a) Four-point mounted resonators. (b)Induced forces and moments due to initial displacement. (c) Letter symbols for a general support structure, and dimensions for support types A, B. C, D, and E.

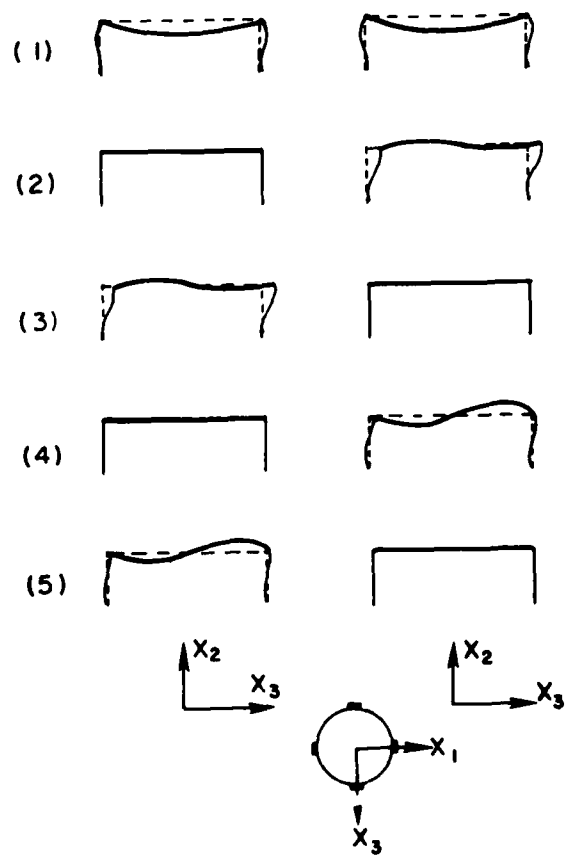



Fig.13 First five vibrational modes of an SC-cut resonator with four-point support of Type A.

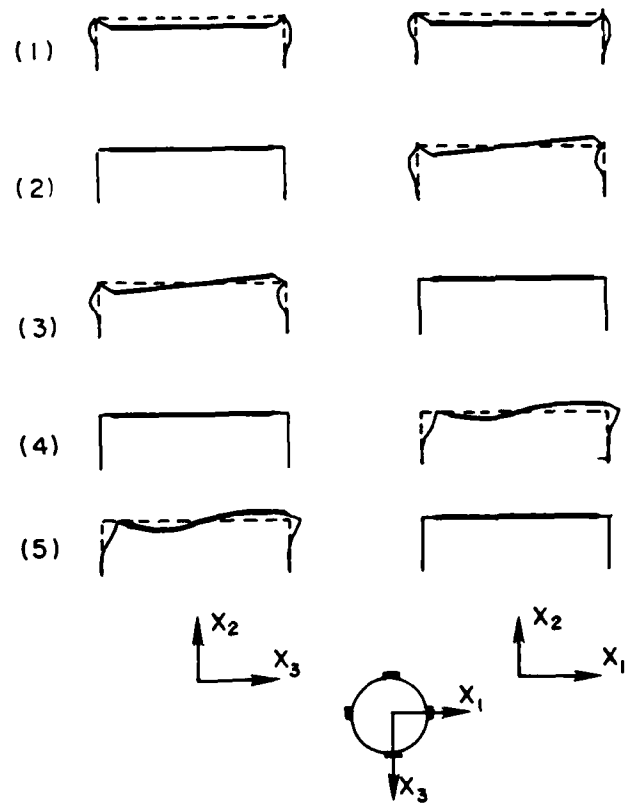



Fig.14 First five vibrational modes of an SC-cut resonator with four-point support of types B, C, D, and E.

42nd Annual Frequency Control Symposium - 1988

# LINEAR MODEL OF THE CONTOURED RESONATORS

J. Détaint[1], H. Carru[1] J. Schwartzel[1]
B. Capelle[2], A. Zarka[2]

[1] Centre National d'Etudes des Télécommunications
92220 Bagneux France
[2] Laboratoire de Minéralogie et de Cristallographie
Université P et M. Curie 75005 PARIS

## ABSTRACT

This analysis is based upon the approximate equation to the thickness vibration modes established by H.F. Tiersten and coworkers (1).

In a first part, a semi-algebraical method of resolution is investigated for partially or totally contoured resonators. In this method, the eigen solutions at V = 0 are obtained by separation of the approximate equation in transformed rectangular or polar coordinates for the different regions of the resonator. These solutions are obtained in the form of series of products of trigonometric and Bessel functions for the plane regions, or of products of parabolic cylinder functions for the regions with a spherical contour. A particular type of contour presenting the same lateral anisotropy as the crystal plate is also considered (an elliptical paraboloid). In this case, *the eigen solutions* are obtained as series of products of trigonometric and Whittaker functions and the eigen frequencies and modes have interesting properties. A discretization of the boundary and continuity conditions is used to obtain an equation to the eigen frequencies. The forced modes can be obtained as a linear combination of the eigen modes. The interest of this method is its precision and the moderate amount of needed computation.

In a second part, a totally numerical solution of the approximate equation, by the finite element method is proposed. This method permits to consider practically any type of thickness mode resonator. it is applied here to some typical examples of contoured resonators. The results are compared with those obtained with the previous method and with experimental results.

## INTRODUCTION

The totally or partially contoured thickness mode resonators are of very common use for frequency generation or filtering. They present the advantage to have, in the frequency range 0.5 - 10 MHz, high Q factors together with reasonable lateral dimensions. They are in some manner a kind of miniaturized resonators). The plano-convexe and the bi-convexe resonators are characterized by nearly Gaussian modes, (decreasing as $e^{-x^2}$), that confer high Q factors on these resonators, especially when overtones modes are used.

The partially contoured ones (bevelled resonators) are used when it is necessary to have plates with a reduced diameter to thickness ratio together with a high coupling coefficient ($C_1/C_0$). They can also be made to present very few unwanted responses, this fact is of particular interest for filter applications. All these favourable properties are the consequence of the occurence in these resonators of a particular type of energy trapping that results mostly from the lateral variations of the thickness.

Mathematical models of these resonators having more and more accuracy were presented in the recent years. Most of them were derived from the recent developments of the theories of thickness vibrations of the piezoelectric plates established by Professor Tiersten for the resonators vibrating near one overtone of a thickness mode having a thickness dependance essentially similar to that of the corresponding one dimensional mode of an infinite piezoelectric plate was used (1) (2) (3).

Models of plano-convex resonators of successively increased generality and accuracy were established by Tiersten and coworkers (2) (3) (4). Resonators with cylindrical bevels were considered by Vangheluwe (5) and Sinha (6). Plane and corrugated resonators with arbitrary electrode geometry were then considered (7). Recently Dulmet and Tiersten (8) have treated the case of plano-convex fully electroded rectangular plates (finite dimension). In these proceeding, Peach (9) considers a method to establish other approximate equations accounting for the lateral variations of the thickness of the plate.

In this paper, we present semi-algebrical solutions of the approximate equation for several types of partially electroded contoured resonators with finite dimensions. A finite element solution of this equation is then proposed for several case of contoured resonators.

## I - THE APPROXIMATE EQUATION

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 (noted (1)). 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 $x_1$ and $x_3$ choosen to eliminate the mixed derivatives (3) the following form :

$$M_n' \frac{\partial^2 \tilde{u}_1^n}{\partial x_1^2} + P_n' \frac{\partial^2 \tilde{u}_1^n}{\partial x_3^2} - \frac{n^2\pi^2 C^*}{4h^2} \tilde{u}_1^n + \rho\omega^2 \tilde{u}_1^n$$

$$= \rho\omega^2 (-1)^{(n-1)/2} \frac{e_{26}\, 4V_0\, e^{j\omega t}}{c^{(1)} n^2 \pi^2}$$

In this equation, $\tilde{u}_1^n(x_1,x_3,t)$ is the lateral dependance of $\tilde{u}_1^n(x_1,x_2,x_3,t) = \tilde{u}_1^n \sin(n\pi x_2/2h)$. $\tilde{u}_1$ arises from a transformation of the usual component of the displacement $u_1$ and of the potential $\Phi$, made to replace the inhomogeneous boundary condition $\Phi = \pm V_0 e^{j\omega t}/2 at x_2 = \pm h$ by the homogeneous conditions $\tilde{\Phi} = 0$. The inhomogeneous term of the equation results from this transformation.

$$u_1^n = \tilde{u}_1^n - \frac{e_{26}\, V_0\, e^{j\omega t}}{c^{(1)} 2h} \cdot x_2 \; ;$$

$$\Phi = \tilde{\Phi} + \frac{x_2 V_0 e^{j\omega t}}{2h} \; ; \; u_1 = \sum_n u_1^n$$

$C^*$ is either $\bar{c}^{(n)}$ for the unelectroded part of the resonator or $\hat{c}^{(n)}$ for the electroded part. $\bar{c}^{(n)}$ is the stiffened elastic constant relative to the corresponding one dimensional mode (Eigen-value of the Christoffel Matrix).

CH2588-2/88/0000-019 $1.00 C 1988 IEEE

$$\tilde{c}^{(1)} = \bar{c}^{(1)}\left(1 - 8\,k^{(1)2}/n^2\pi^2 - 2\hat{R}\right)$$

$\tilde{c}^{(1)}$ is a constant that includes the electrical and mechanical effects of the metallization. $4k^{(1)2}/n^2\pi^2$ is the relative frequency lowering due to the electrical effect of the metallization ; $\hat{R} = 2\rho'h'/\rho h$ is the relative frequency lowering due to the inertia of the electrodes (figure 1). $\bar{c}^{(1)} = \bar{c}^{(1)}\,(1-k^{(1)2})$ is a pseudo ordinary elastic constant. $k^{(1)}$ is the coupling coefficient of the corresponding one dimensional mode. $M'_n$, $P'_n$ are intricate functions of the material constants and of the plate orientation (3).

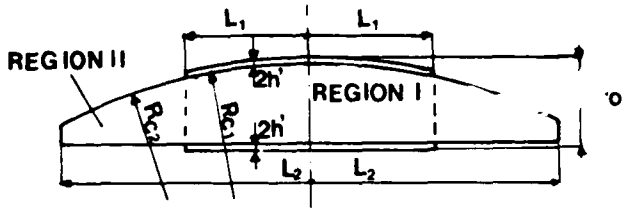



Figure 1 : Contoured resonators .

**Boundary and continuity conditions** : The traction free conditions on the major 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 (figure 1). On the external edges of the resonators, the approximate boundary conditions to be applied are not uniquely determined (7) and we have to choose in each case what type of mechanical condition, compatible with the scalar equation is the best approximation of the traction free conditions. However, for a well designed resonator this choice is of small importance since $u_1$, and the corresponding energy are extremely small near the edge.

## II - SEMI ALGEBRAICAL MODEL

This model is based upon algebraical solutions of the approximate equation. As for the case of the plane resonator (7) we are looking for solutions, that are, for each region of the resonator, in the form of an infinite series of eigen functions which are exact solutions of the equation. As these solutions are obtained by separation in coordinates systems for which the electrode boundary and/or the plate boundary are not, in general, lines with one constant coordinate, the continuity and boundary conditions are expressed at a discrete number of points. This permits to obtain, by truncation of the series, an approximate solution that converges toward the exact solution when the number of points is increased. Most often, the convergence is very fast, so that a limited amount of points (of terms) is required to have a great precision.

Separable forms for the thickness variations : The more commoun contoured resonators are the plano-convex, the biconvex and the bevelled plates. In some instances, plano-convex or biconvex resonators that present a second shorter radius of curvature (bevel) near their edges are used (Figure 2). All these resonators have spherical contouring. The most commonly used separable approximation of the term $(1/2h)^2$ that appears in the partial derivative equation is a first order (2) approximation in $r^2 = (x_1^2 + x_3^2)$ :

$$\left(\frac{1}{2h}\right)^2 = \left(\frac{1}{2h_0}\right)^2\left(1 + \frac{x_1^2+x_3^2}{2\,R_c h_0}\right) \qquad [1]$$

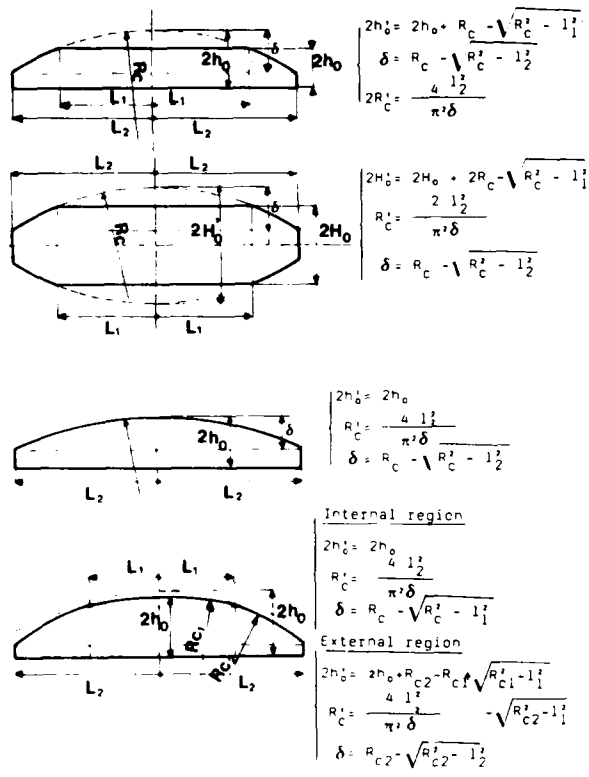



Figure 2: Considered geometries

The same approximation applied to many surfaces tangent to or oscullating the sphere on the $x_2$ axis (figure 2) will lead to the same expression. This is the case for the paraboloid of revolution $-x_2 = (x_1^2 + x_3^2)/2p$ with p = Rc (radius of curvature at the summit ($x_3$=$x_1$=0)). Another approximation may be usefull when a precise representation of the thickness variations near the edges of the resonator is needed ; one was proposed in reference (10) :

$$\left(\frac{1}{2h}\right)^2 = \left(\frac{1}{2h_0}\right)^2\left(1 + \frac{\pi^2\delta\,(x_1^2+x_3^2)}{8\,h_0\,l_2^2}\right) \qquad [2]$$

where $\delta$, $l_2$ are defined on figure 2.

If we consider $R'_c = 4l_2^2/\Pi^2\delta$ and, to include the case of all contoured resonators considered in figure 2, if we introduce the quantity $h'_0$ (defined on figure 2), the form of relation [1] can be conserved for all cases.
More over, using this formula, it is possible with a numerical approximation technique to find $R'_c$ and $h'_0$ in order to optimize the representation of a given region of the resonator.

$$\left(\frac{1}{2h}\right)^2 = \left(\frac{1}{2h'_0}\right)^2\left(1 + \frac{x_1^2+x_3^2}{2R'_c h'_0}\right) \qquad [3]$$

As can be seen on figure 3 the approximation given by expression [1] is generally better near the center of the resonator whereas the approximation [2] is better near the edges of the plates.

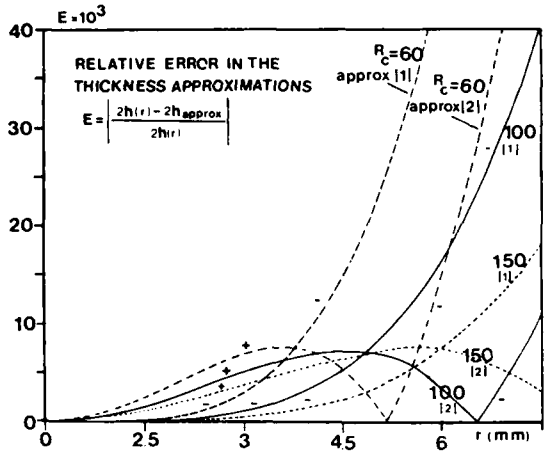



Figure 3: Thickness approximations [1] and [2]

Although they seem to constitue less precise approximations to the spherical contour, it is interesting from a mathematical point of view and most probably for practical applications to consider the case of non spherical (anitropic) contours that lead to separable expression of $(1/2h)^2$ like :

$$\left(\frac{1}{2h}\right)^2 = \left(\frac{1}{2h'_0}\right)^2 \left(1 + \frac{x_1^2}{2R'_1 h'_0} + \frac{x_3^2}{2R'_3 h'_0}\right) \qquad [4]$$

Simple surfaces (elliptical paraboloids, ellipsoids) corresponds to this case. We will observe, that a particular case of such non spherical contours, leads to interesting properties for the eigen frequencies and eigen modes ; this arises when the anisotropy of the thickness variations corresponds exactly to the lateral anisotropy of the plate.

Eigen solution at V = 0 for the spherical regions : In the following $C^*$ is equal to $\bar{c}^{(1)}$ if the spherical portion consider is electroded or to $\bar{c}^{(2)}$ in the other case. The homogeneous form of the equation is

$$M'_n u^n_{1,11} + P'_n u^n_{1,33} + 4\pi^2 \rho \left(f^2 - f^2_{C1}\left(1 + \frac{x_1^2 + x_3^2}{2h'_0 R'_c}\right)\right) \hat{u}_1 = 0 \qquad [5]$$

where :

$$f_{C1} = \frac{n}{4h'_0} \sqrt{\frac{C^*}{\rho}}$$

The transformation of coordinates (5)

$$\begin{cases} X_1 = (C^*/M'_n)^{1/4} . \gamma . x_1 \\ \text{with } \gamma = \left(\dfrac{n^2\pi^2}{2R'_c h'^3_0}\right)^{1/4} \\ X_3 = (C^*/P'_n)^{1/4} . \gamma . x_3 \end{cases} \qquad [6]$$

permits to obtain a solution in the separate form $\hat{u}_1 = F(X_1).G(X_3)$ where F and G are solutions of the parabolic cylinder equations

$$\begin{cases} F'' - \left(\dfrac{X_1^2}{4} + a_1\right) F = 0 \\ G'' - \left(\dfrac{X_3^2}{4} + a_3\right) F = 0 \end{cases} \qquad [7]$$

The two, non independant, separation constants have to verify the expression :

$$4\pi^2 \rho (f^2 - f^2_{C_1}) = -a_1 M'^{1/2}_n C^{*1/2} \gamma^2 - a_3 P'^{1/2}_n C^{*1/2} \gamma^2 \qquad [8]$$

or

$$-a_1 \left(\frac{M_n}{C^*}\right)^{1/2} - a_3 \left(\frac{P_n}{C^*}\right)^{1/2} = A = \frac{n\pi}{2} \left(\frac{B'_c}{2h'_0}\right)^{1/2} \left(\frac{f^2 - f^2_{C_1}}{f^2_{C_1}}\right)$$

The most usual standart pair of solutions of the parabolic cylinder equation are the Withittaker's functions V(a,X), U(a,X)[11]. The U function is often refered as $D_\upsilon(X)$ with a $a = -\upsilon - 1/2$. Since we are considering the symetrical eigen modes of a symetrical finite plate with symetrical electrodes, we can choose to use, as Dulmet and Tiersten (8) the even series solution of the parabolic cylinder equation $y_1$ (a,X) (11). The even part $D_\upsilon$ of $D_\upsilon$ can be expressed in the form of a series, it is also a solution.

$$D^1_\upsilon = 2^{-\upsilon/2} e^{-X^2/4} H^1_\upsilon \left(\frac{X}{\sqrt{2}}\right) \qquad [9]$$

$$H^1_\upsilon = \frac{1}{2\Gamma(-\upsilon)} \sum_{m=0}^{\infty} \left(\frac{\Gamma\left(\frac{2m-\upsilon}{2}\right) \left(\frac{2X}{\sqrt{2}}\right)^{2m}}{(2m)!}\right)$$

where : $H^1_\upsilon$ is even part of the Hermite function (12).

Gauss Hermite approximation : The solution given by Tiersten (1) consists for an entirely electroded plano-convexe resonator to take a solution of the parabolic cylinder equations , the $D_\upsilon$ function with even integer values of the $\upsilon$ index. These functions are the only parabolic cylinder functions that are bounded in the interval $(-\infty + \infty)$. In this case :

$$D_m = 2^{-m/2} e^{-X^2/4} H_m (X/\sqrt{2})$$

where $H_m$ is an Hermite polynomial.

General case, It is possible to obtain a series solution in the following manner :
if we set

$$\begin{cases} a_1^m \left(\dfrac{M_n}{C^*}\right)^{1/2} = \dfrac{A}{2} + f(m) \\ a_3^m \left(\dfrac{P_n}{C^*}\right)^{1/2} = \dfrac{A}{2} - f(m) \end{cases} \qquad \text{with } m = 1, 2, 3, .$$

$a_1^m$ and $a_3^m$ verify the relation [7] and a solution of [5] is given by :

$$\hat{u}_1 = F(a_1^m, X_1).G(a_3^m, X_3)$$

where F and G are parabolic cylinder functions. A series solution can be obtained, (if f(m) is choosen to obtain a convergence of the series) as :

$$\hat{u}^n_1 = \sum_{m=0}^{\infty} C^m y_1 (a_1^m, X_1).y_1(a_3^m, X_3) \qquad [10]$$

in our applications we have taken f(m) = 2m (m even).

Eigen solutions for plane regions (bevelled resonators) : The homogeneous equation is now :

$$M'_n \hat{u}_{1,11} + P'_n \hat{u}_{1,33} + 4\pi^2\rho(f^2 - f^2_{C_1})\hat{u}_1 = 0 \quad [11]$$

To have common notations with the preceedings paragraph we use the coordinate transformations :

$$\begin{cases} X_1 = \left(\frac{C^*}{M'_n}\right)^{1/2} \gamma_2 x_1 \\ X_3 = \left(\frac{C^*}{P'_n}\right)^{1/2} \gamma_2 x_3 \end{cases} \qquad \begin{cases} X_1 = r\cos t \\ \text{with } \gamma_2^4 = \frac{n^2\pi^2}{2ho^4} \\ X_3 = r\sin t \end{cases}$$

This leads to the equation :

$$\hat{u}_{1,rr} + \frac{1}{r}\hat{u}_{1,r} + \frac{1}{r^2}\hat{u}_{1,tt} + A'\hat{u}_1 = 0 \quad [12]$$

$$\text{with } A' = \frac{n\pi}{4}(f^2 - f^2_{C_1})/f^2_{C_1}$$

that can be separated as $u_1 = R(r).T(t)$ in :

$$\begin{cases} r^2R'' + rR' + R(A'^2r^2 - \nu^2) = 0 \\ T'' + \nu^2 T = 0 \end{cases} \quad [13]$$

the first 2nd order differential equation is a Bessel equation that has for solution the Bessel functions of first kind $J_\nu(r\sqrt{A'})$ and of 2nd kind $Y_\nu(r\sqrt{A'})$. The second equation has for solutions cos $\upsilon$ t and sin $\upsilon$ t ; since the solution $U_1(r,t)$ must be such that $U_1(r,t) = U_1(r,t+2m\pi)$, $\upsilon$ has to be an integer $\upsilon = m$.. Since the $Y_m$ functions are not finite at the origin, for the electroded central region of a bevelled resonator, the symetrical solutions can be expressed as (7) :

$$\hat{u}_1 = \sum_{m=0}^{\infty} B_m J_m(r\sqrt{A'}) \cos mt \quad [14]$$

with m even integer

Eigen solutions for non spherical contours with the expression of $1/h^2(x_1, x_3)$ given previously and a coordinate transformation defined by :

$$\begin{cases} X_1 = \left(\frac{C^*}{M'_n}\right)^{1/4} \gamma_1 x_1 ; \gamma_1^4 = \frac{n^2\pi^2}{2R_1h_0^3} \\ X_3 = \left(\frac{C^*}{P'_n}\right)^{1/4} \gamma_3 x_3 ; \gamma_3^4 = \frac{n^2\pi^2}{2R_3h_0^3} \end{cases}$$

The homogeneous form of the approximate equation can also be separated in two parabolic cylinder equation.

$$\hat{u}_1 = F(X_1) \cdot G(X_3)$$

$$\begin{cases} F'' - \left(\frac{X_1^2}{4} + a_1\right)F = 0 \\ G'' - \left(\frac{X_3^2}{4} + a_3\right)G = 0 \end{cases} \quad [15]$$

The two non independant separation constants $a_1$ and $a_3$ have to verify the relation :

$$-a_1 M'^{1/2}_n C^{*1/2} \gamma_1^2 - a_3 P'^{1/2}_n C^{*1/2} \gamma_3^2 = 4\pi^2\rho(f^2 - f^2_{C_1}) \quad [16]$$

Gauss-Hermite approximation : for this case again this approximation consist to take a solution in the form :

$$\hat{u}_1 = D_m(X_1) D_{m'}(X_3)$$

$$\text{with : } m = -a_1 - 1/2 \; ; \; m' = -a_3 - 1/2$$
$$m, m' \text{ even integers}$$

with m and m' even integers (symetric modes) and : $m = -a_1 - 1/2$ ; $m' - a_3 - 1/2$.
The eigen frequency equation is then obtained from [16] :

$$f^2_{nmm'} = f^2_{C_1}\left[1 + \frac{2}{n\pi} \cdot B\right] \quad [17]$$

$$B = (m + \tfrac{1}{2})\left(\frac{M'_n 2ho}{C^* R_1}\right)^{1/2} + (m' + \tfrac{1}{2})\left(\frac{P'_n 2h_0}{C^* R_3}\right)^{1/2}$$

In this relation we can observe that, when $R_1/R_3 = M'_n/P'_n$ the resonance frequency of the nmm'th mode is only dependant of the sum m + m'.

$$f^2_{nmm'} = f^2_{C_1}\left[1 + \frac{2}{n\pi}\left(\frac{2ho}{R'}\right)^{1/2}(m + m' + 1)\right]$$

[18]

So that all anharmonics with constant m + m' have the same eigen frequencies. We can observe that this occurs when the lateral thickness variations have the same lateral anisotropy as the plate. A similar eigen-frequency equation would have been found for the case of an ordinary contoured resonator made in a plate with no lateral anisotropy ($M'_n = P'_n$ ; $R_1 = R_3$). A resonator of the latter type can have the same eigen frequency $f_{nmm'}$ as a conventional plano convex resonator of the same thickness $2h_0$ if :

$$R' = R_c\left(\frac{m + m' + 1}{(m + \frac{1}{2})\left(\frac{M'_n}{C^*}\right)^{1/2} + (m' + \frac{1}{2})\left(\frac{P'_n}{C^*}\right)^{1/2}}\right)^2 \quad [19]$$

When this relation is verified, the nmm' th eigen modes of both types of resonators (normalized to 1 at the origin) are related by a simple geometric transformation. This anisotropic type of contour, is also optimal from the point of view of anharmonic mode spacing. This is illustrated in figure 4. for a conventional AT planoconvex resonator (2) with $2h_0 = 0{,}68\,785$ mm $R_c = 106$ mm, $\hat{R} = 1.86\,10^{-3}$ and an unconventional AT resonator with $R'_c = 34{,}80$ mm ; $R_1 = R'M'_n/C^* = 132.1mm$; $R_3 = R'P'_n/C^* = 82.7mm$ ; the same $h_0$ and the same $\hat{R}$.

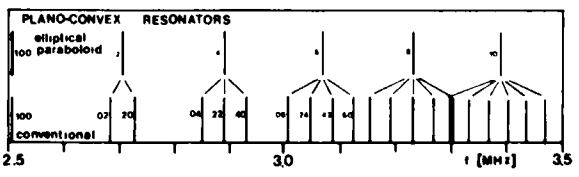



Figure 4: Anharmonic spectrum of a non-spherical resonator.

From formula [17] we can observe that it is possible to manage the anharmonic spectrum of a non spherical resonator using the ratio $R_1/R_3$. This may be usefull for double rotated cuts.

Resonators with finite plate and electrode dimension : A series solution can be obtained using the method previously described for the spherical contours

Other solution for the case when $R_1/R_3 = M_n/P_n$ :

If we substitute R' in the coordinate transformation we obtain :

$$\begin{cases} X_1 = \left(\frac{C^*}{M'_n}\right)^{1/2} \cdot \left(\frac{n^2\pi^2}{2\,R'\,h_0^3}\right)^{1/4} \cdot x_1 \\ X_3 = \left(\frac{C^*}{P'_n}\right)^{1/2} \cdot \left(\frac{n^2\pi^2}{2\,R'\,h_0^3}\right)^{1/4} \cdot x_3 \end{cases} \quad [20]\ [21]$$

with the coordinate transformation :

$X_1 = r\cos t$ ; $X_3 = r \sin t$

The homogeneous approximate equation is transformed into:

$$\hat{u}_{1,rr} + \frac{1}{r}\hat{u}_{1,r} + \frac{1}{r^2}\hat{u}_{1,tt} + \left(A - \frac{r^2}{4}\right)\hat{u}_1 = 0 \quad [22]$$

$$A = \frac{nm}{2}\left(\frac{R'}{2ho}\right)\left(\frac{f^2 - f_{C_1}^2}{f_{C_1}^2}\right) \quad [23]$$

This equation can be separated in two ordinary differential equations :

$$\hat{u}_1 = R(r) \cdot T(t)$$

$$\begin{cases} T'' + \mu^2 T = 0 \\ R'' + \frac{1}{r}R' + \left(A - \frac{r^2}{4} - \frac{\mu^2}{r^2}\right)R = 0 \end{cases} \quad [24]$$

The solution of the first equation in [24] are the trigonometric functions cos(μt) and sin(μt) ; since $\hat{u}_1$ must be such that $\hat{u}_1 = \hat{u}_1(t + 2\pi)$, the separation constant has to be an integer ; μ = m. With the transformation $R(r) = W(r^2/2)/r$ the second differential equation becomes :

$$W'' + \left(-1/4 + \frac{A/2}{r^2/2} + \frac{1/4 - m^2/4}{(r^2/2)^2}\right)W = 0$$

The equation [25] is a confluent hypergeometric equation in the form given by Whittaker (13) which has for solutions the Whittaker's functions M and W.

$$R_m(r) = \frac{1}{r}B^m M_{A/2,\, m/2}(r^2/2) + C^m W_{A/2,\, m/2}(r^2/2)$$

The M functions are finite for r = 0, and the W functions finite at +∞. Their numerical evaluations are respectively as easy as for the y1 and the D functions (11) (13). The Bessel functions, the parabolic cylinder functions and the trigonometric functions can be expressed in terms of the Whittaker's functions M and W (11).

A general expression for the symetrical eigen modes of such non spherical resonators is thus :

$$\hat{u}_1^n = \sum_{m=0}^{\infty} R_m(r) \cdot \cos mt \quad [25]$$

with m even

For the electroded region of the resonator we can retain only the M functions whereas for the external region it is possible to consider only the W functions, especially if the amplitude of $\hat{u}_1$ is very small at the edge of the resonator.

Expression of the continuity and boundary conditions

At all points of the electrode edge we have to specify the continuity of $\hat{u}_1$ and of its normal derivative. As previously discussed (7) (14), this can be approximated by the expression of these conditions at only a discrete number $q_1$ of points (figure 4). If the eigen solutions for the two regions are :

$$\begin{cases} u_1^I = \sum_{m=0} D_m H_m(x_1, x_3) \\ u_1^{II} = \sum_{m=0} E_m K_m(x_1, x_3) \end{cases} \quad [26]$$

at the point P $(x_1^p, x_3^p)$ we have :

$$\begin{cases} u_1^I(x_1^P, x_3^P) = u_1^{II}(x_1^P, x_3^P) \\ n_1 \frac{\partial u_1^I(P)}{\partial x_1} + n_3 \frac{\partial u_1^I(P)}{\partial x_3} = n_1 \frac{\partial u_1^{II}(P)}{\partial x_1} + n_3 \frac{\partial u_1^{II}(P)}{\partial x_3} \end{cases} \quad [27]$$

where ; $\vec{N}(n_1, n_3)$ is the normale to the boundary between region I and region II at P. (figure5 )

At the resonator edge we have to express a boundary condition compatible with the form of the approximate equation. Since only Dirichlet, Neumann or mixed conditions relative to the lateral part of $\hat{u}_1$ ie $\hat{u}_1(x_1, x_3)$ are compatibles with the equation, we have to choose among these ones the best approximation to the zero traction conditions. For well designed resonators this choice is of little importance.

For example, if we choose the conditions $\hat{u}_1 = 0$, at the $q_2$ points we obtain $q_2$ conditions of the form.

$$0 = \sum_m E_m K_m(x_1^P, x_3^P) \quad [28]$$

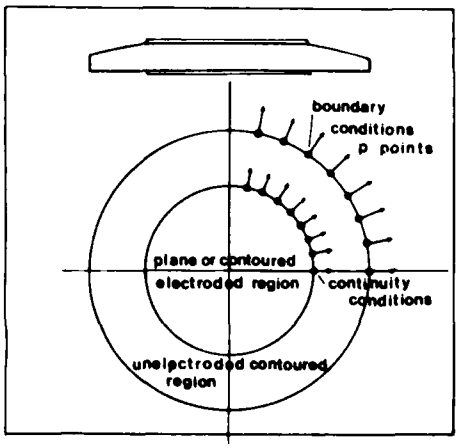



Figure 5: Discretisation at the electrode and plate edges.

If the number of points at the electrode edge and at the plate edge are chosen equal ($q_1=q_2$), we have $2q_1+q_2=3q$ conditions so that the series solution for $u_1^I$ can be truncated to 2q terms and the series for $u_1^{II}$ has then to be truncated to q terms.
The ensemble of the relations [27] and [28] written for the q points at the electrode edge and the q points at the plate edge constitutes thus an homogeneous linear system in the coefficients $D_0, D_2, \ldots D_{4q-2}$, $E_0, E_2, \ldots, E_{2q-2}$. To have non trivial solutions, the determinant of this system must vanish ; this condition constitutes the eigen frequency equation. This equation is solved numerically so that we obtain the eigen frequencies $f_n^1, f_n^2, f_n^3, \ldots, f_n^j, \ldots$
For each of the eigen frequencies the homogeneous system is solved and the coefficients $D_0^j, D_2^j, D_4^j, \ldots, D_{4q-2}^j, E_0^j, \ldots E_{2q-2}^j$ are obtained as a function of one of them. Thus the eigen modes are obtained in the form :

$$\begin{cases} u_1^{nsI} = e^{j\omega t}\left[\sum_{m=0}^{4q-2} D_m^s \; H_m(x_1,x_2,f_n^s)\right]\sin(m\pi x_2/2h) \\ u_1^{nsII} = e^{j\omega t}\left[\sum_{m=0}^{2q-2} E_m^s \; K_m(x_1,x_2,f_n^s)\right]\sin(m\pi x_2/2h) \end{cases} \quad [29]$$

where : m is even
$H_m(x_1, x_2)$ ; $K_m(x_1, x_2)$ are given by relations [10] or [14] or [25] according to the geometry of the coutoured resonator.

Forced solutions (principle)

The forced modes are expressed as a sum of all the eigen modes (7) :

$$\hat{u}_1^n = \sum_s H^{ns}\,\hat{u}_1^{ns}\sin(m\pi x_2/2h) \quad , \quad u_1 = \sum u_1^n \quad [30]$$

$$\iint_{ST} \hat{u}_1^{ns}\,\hat{u}_1^{nv}\,ds = \delta_{sv}\,N^{(n)}(s) \quad [31]$$

Using the orthogonality of the eigen modes ; the fact that $u_1^{nv}$ is a solution of the homogeneous approximate equation, and the fact that $u_1^n$ is a solution of the inhomogeneous equation we can find the values of H= (3), (7). Then using the constitute equation giving $D_2$, the electrical admittance of the n*overtone is found as :

$$Y = j\omega\left[\left(C_z - \sum_n \sum_s C_{ns}\right) - \sum\sum C_{ns}\left(\frac{f^2_{ns}}{f^2 - f^2_{ns}}\right)\right] \quad [32]$$

where :

$$C_z = \frac{S_e'}{2h}\,\varepsilon_{22}\,(k^2+1) \quad \text{(capacitance at f = 0)}$$

$$C_{ns} = \left(\frac{4k^2\,\varepsilon_{22}}{n^2\pi^2 h_0}\right)\frac{\left(\iint_{se}\hat{u}_1^{ns}\,ds\right)^2}{N^{(n)}(s)}$$

$C_{ns}$ is the dynamic capacitance.

$S_e'$ surface corrected for the thickness variations.

Properties of this method of solutions : A computer program was implemented for the case of plano convexe, biconvexe and bevelled resonators. As previously discussed, the most important factor for the precision of the model is the rate of convergence of the series solutions. As in the case of the plane resonator (7), the rate of convergence of the numerical solution toward the exact solution can be appreciated considering the eigen frequencies

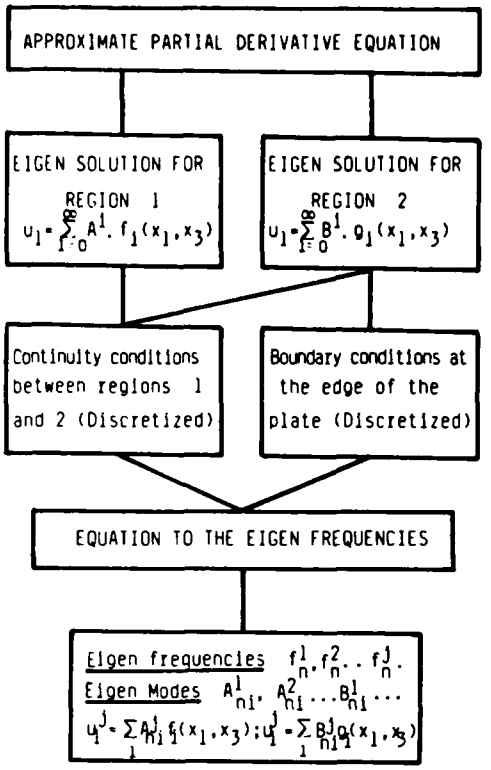



Figure 6: Principle of the semi-algebraical method of solution.

when the number of points at which the continuity and the boundary condition are expressed is increased.
This was done in the case of a plano-convexe resonator with 2 ho = 1.321 mm ; $R_c$ = 175 mm ; $2l_1 = 2h_e$ = 8 mm ; $2l_2$ = 15 mm ; n = 5 ; Quartz AT cut. The results are given in table I :

TABLE I

| $q_1 = q_2$ | $f^{3.0}$ (5*overtone) (Hz) |
|---|---|
| 4 | 6334 078.12 |
| 5 | 6334 030.13 |
| 6 | 6334 012.30 |

We can observe a quite fast convergence. A more detailed consideration of the eigen modes, computed here using the $y_1$ parabolic cylinder function for the external region, has revealed that this function gives a better representation (more rapidly convergent) for the electroded region than for the external region. We expect that the use of the $D_1^j$ function for the unelectroded region will give a faster convergence.

We have computed, for this resonator, the influence of the choice of the type of boundary condition at the plate edge (table 2), as expected, we cannot observe any significative difference in the computed eigen frequencies.

TABLE II

| TYPE | $f^{3.0}$ (Hz) |
|---|---|
| Neumann | 6334 012.30820 |
| Dirichlet | 6334 012.30820 |

## III FINITE ELEMENT SOLUTION

The homogeneous form of the approximate equation is solved by a finite element technique (15) (16) when the parameter h has any slowly variation with $x_1$ and $x_3$ The natural boundary condition implies the normal derivative of displacement equal to zero, so that the continuity relations between electroded and unelectroded regions are automatically satisfied and the condition at the edge of the plate is of Neumann type. This later one can also be made of Dirichlet type. The circular domain is divided into triangular elements with the displacement at the corners as unknown (100 x 4 to 200 x 4 nodes per resonator). The problem is solved as an eigenvalue one. The matrices resulting from the assemblage of these elements are stored on a "skyline" basis. The dynamic capacitance of the resonator can also by calculated.

The overal accuracy can be established by considering the variation of the frequency when one adds a few elements.

The size of the domain has to be kept to the minimum necessary to give a fixed accuracy (unless it is enough subdivided, which leads to prohibitive computing time) ; otherwise spurious oscillations and wrong frequencies appear.

## IV COMPARISON OF COMPUTED AND EXPERIMENTAL RESULTS

Computer programs were set up using the previously described semi-algebraical and finite element methods. The computed eigen frequencies and eigen modes are compared with experimental ones in the case of a bevelled resonator and of a plano-convex resonator.

Comparison of eigen frequencies : This comparison is made for a bevelled resonator designed for filter application to have very reduced anharmonic responses.

Table III : Resonance frequency of a bevelled resonators : 2ho = 1,094 mm ; $2l_1$ = 8,93 ; $2l_2$ = 15.00 mm ; $R_c$ = 60 mm ; R = 0,0028 ; AT QUARTZ

| | | |
|---|---|---|
| Fundamental mode | Fr. experimental | 1556. KHz |
| | Fr. finite element method | 1557. |
| 3 rd overtone | Fr. experimental | 4563. KHz |
| | Fr. finite element method | 4558. |

In table III, we can see that a good agreement, compatible, with the uncertainty on the thickness, mass loading measurements and the material constants, is obtained.

A comparison was then made in the case of a plano convex resonator with circular electrodes. In table IV we compare the experimental frequency with the eigen frequencies computed for the Gauss-Hermite approximation with two approximations for $1/h^2$ (given (in paragraph II) and with with the results of the semi algebrical model and the results of the finite element calculation.

Table IV : Resonance frequency of a plano convex resonator : [5th overtone, 2 ho = 1.321 ± 1 mm ; R = 0.0017 ; $2l_2$ = 15 mm ; $R_c$ = 175 mm ; AT cut ; circular electrode $2l_1$ = 8 mm (Fr in Hz)]

TABLE IV

| | |
|---|---|
| Fr (experimental) | 6338.560 |
| Fr (approx. 1 for $1/h^2$; Gauss Hermite) | 6333.988 |
| Fr (approx. 2 for $1/h^2$; Gauss Hermite) | 6340.312 |
| Fr (approx. 1 for $1/h^2$ ; Semi Algebraical) | 6334.012 |
| Fr (approx. 2 for $1/h^2$ ; Semi Algebraical) | 6339.907 |
| Fr (Exact $1/h^2$ ; Finite Element Method) | 6334. |

In table IV no thickness adjustments were made to allow for the uncertainty on the resonators parameters and the material constants. During the computation, made in the last three cases with Neumann conditions at the external edge of the plate, it was observed that the displacement near this edge is so small that it is of comparable magnitude as the numerical noise.

Considering the approximations of $1/h^2$, the first one leads to a minorant of the eigen frequency (figure 3) the second to a majorant, but for this type of resonator the first one is much more accurate.

Comparison of the vibration modes : The vibration modes corresponding to the eigen frequencies of table III and IV were computed and are compared to experimental ones in figures 5 and 6. In these figures the variations, computed by the finite element method, of $u_1$ along the $x_1$ axis of the resonators are compared to experimental ones obtained by X-ray measurements (17) (18). This method gives a good approximation of the value of $u_1$ ($x_1$, $x_3$) except for the small values. This can be observed in figure 7.

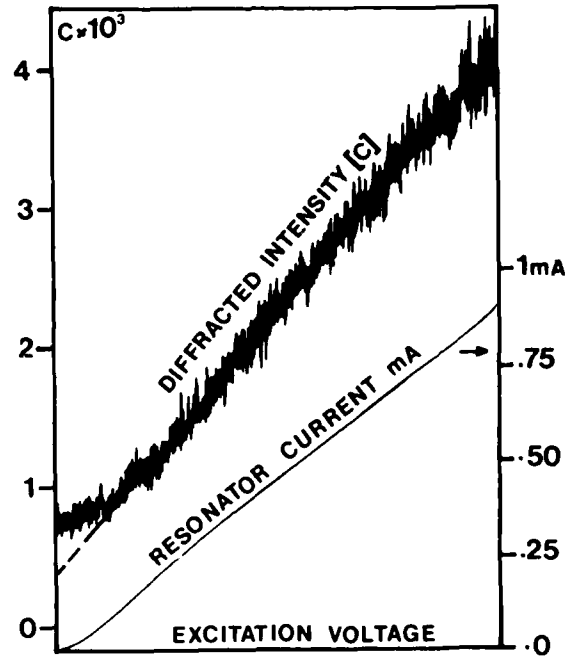



Figure 7: Relation between the diffracted intensity and the displacement (measured by the current)

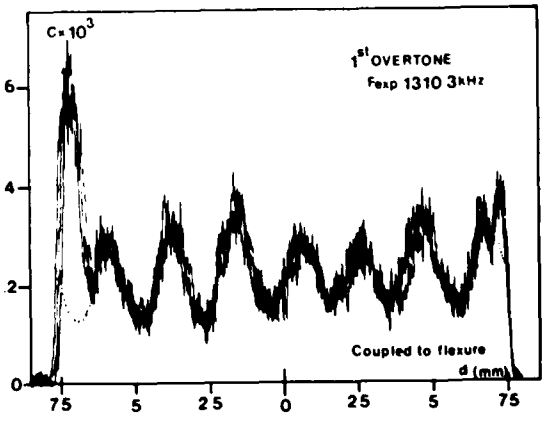



Figure 8: Fundamental mode (plano-convexe resonator)

In figure 8 to 10 are represented the computed eigen modes and the experimental modes of the plano convex resonator of table IV. It can be observed that a good agreement is obtained except for the 1st mode which is strongly coupled to flexure (Probably due to a non negligible value of $u_t$ at the edge of the resonator).

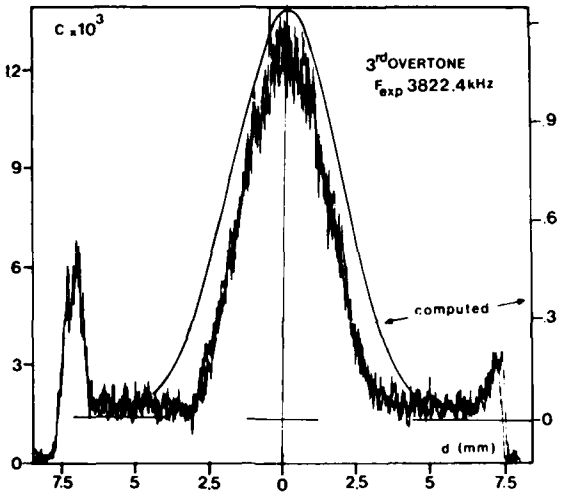



Figure 9: 3rd overtone of the same resonator

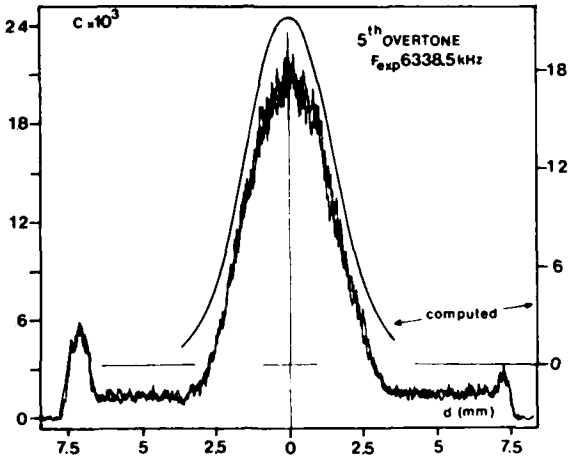



Figure 10: 5th overtone of the same resonator

In figure 11a are given the level curves computed for a plano-convexe resonator with the same parameters as for table III except the electrode diameter that was 10 mm. In figure 11b, are given the experimentally observed Xray interference pattern obtained by stroboscopic topography (18). As discussion reference (18) the fringes existing in figure 11 b, are also level curves of the $u_t$ component of the displacement, we can observe that a good agreement is obtained between the computed and the observed mode.

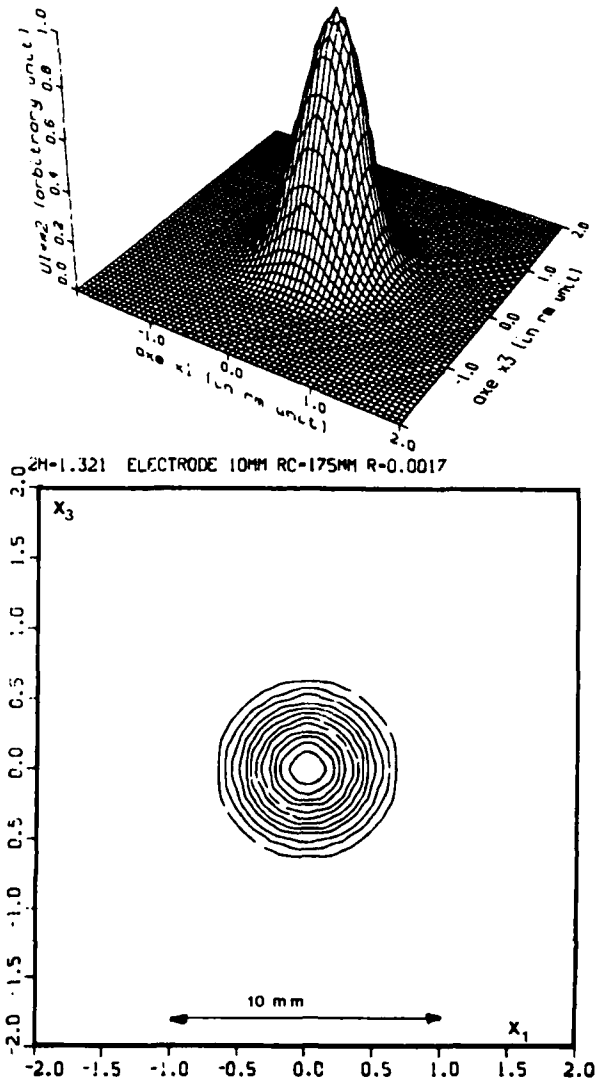



Figure 11a: Computed mode and contour curves

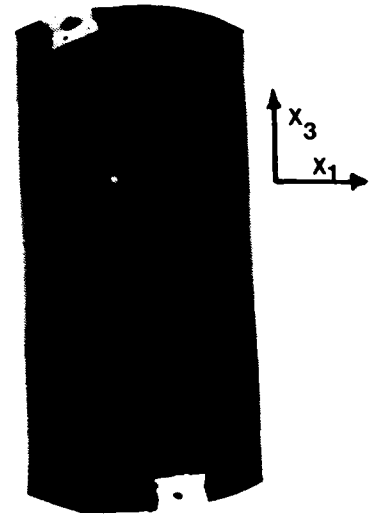



Figure 11b: Observed contour lines obtained by stoboscopic X-ray topography

In figure 12 and in figure 13 are represented the computed and the observed modes considered in table III for a bevelled resonator. In figure 14 is displayed the shape of the second anharmonic of the fundamental mode.

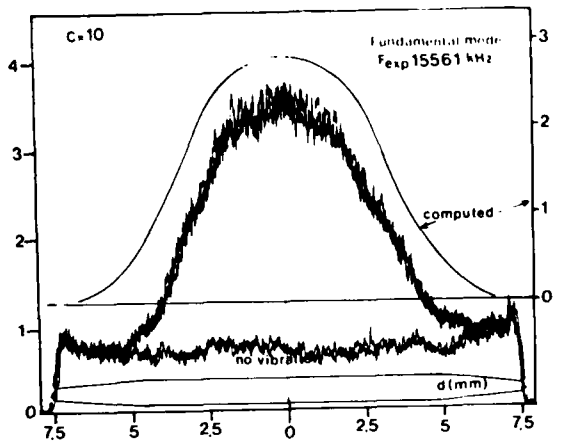



Figure 12 : Fundamentai mode (bevelled resonator.

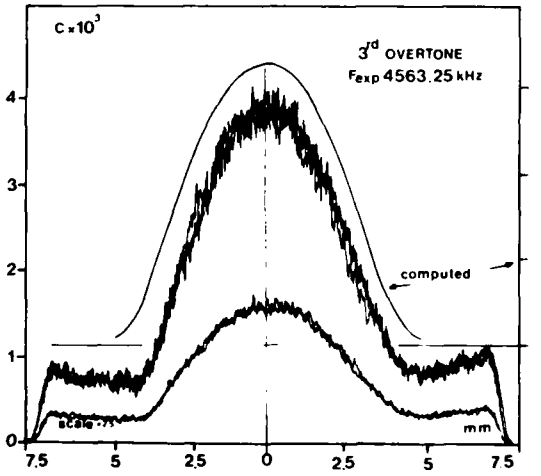



Figure 13 : 3rd overtone (bevelled resonator)

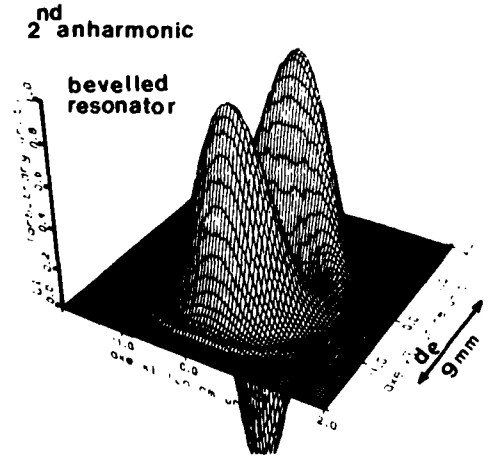



Figure 14 : 2nd anharmonic (bevelled resonator)

## V CONCLUSION

The semi algebraical method of resolution of the approximate equation established by Professor Tiersten and coworkers permits to compute the eigen frequencies and the modes of partially or totally contoured resonators with finite lateral dimensions and quite arbitrary electrode shape. This method has the advantage to require a moderate amount of computations to obtain a high precision.

The finite element method of solution of the approximate equation permits to consider the exact thickness variations together with arbitrary geometry. This method gives accurate results and has the possibility to analyse the thickness vibration of complex piezoelectric structures.

The already obtained results are a new confirmation of the fundamental interest of the approximate equation in the description of many types of resonators.

Extensive application of these models together with other experiments are needed to study the influence of several parameters on the properties of the plano-convexe and bevelled resonators. Further developments of the calculation and experiments concerning the non spherical resonators are also to be done. Several interesting properties are expected from these resonators.

Aknowledgments : The authors wish to aknowledge the skillful assistance of R. Chenebault, A. Daniel, for experimental work and sample preparation. Special thanks are due to M.P. Louis for preparing this manuscript.

## REFERENCES

(1) H.F. Tiersten
J. Acoust. Soc. Am. Vol. 59 p 879 (1976)

(2) H.F. Tiersten
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) H.F. Tiersten ; D.S. Stevens
Proc. 34th Annual Frequency Control Symp p 37 (1982)

(5) D.C.L. Wangheluwe
Proc. 34th Annual Frequency Control Symp p 412(1980)

(6) B.K. Sinha ; D.S. Stevens
J. Acoust. Soc. Am. Vol 66 n° 1 p 192 (1979)

(7) J. Détaint, J. Schwartzel, C. Joly, E. Philippot
Proc. 41st Annual Frequency Control Symp. p 314 (1987)

(8) B. Dulmet, H.F. Tiersten
Proc. IEEE Ultrasonics Symp. p 355 (1986)

(9) R.C. Peach
Proc. 42nd Annual Frequency Control Symposium (1988)

(10) A.V. Apostolov ; S.H. Slavov
Appl. Phys. Vol. 139 p 33-37 (1982)

(11) M. Abramowitz ; I. Stegun
Handbook of Mathematical functions ; Dover (1972)

(12) N.N. Lebedev Special Functions and their applications ; Dover 1972

(13) E.T. Whittaker ; G.N. Watson
A course of Modern Analysis Cambridge Press 1927

(14) K. Nakamura ; H. Shimizu
Proc. IEEE Ultrasonics Symp. p. 606 (1976)

(15) O.C. Zienkiewicz
The finite Element Method Mc Graw Hill 1977

(16) H. Carru, R. Lefèvre,
Proc. 42nd Annual Frequency Control Symp. (1988)

(17) 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 Symposium
(1986)

(18) A. Zarka, B. Capelle, B. Zheng, J. Détaint,
J. Schwartzel
Proc. 42nd Annual Frequency Control Symposium
(These proceedings) 1988.

42nd Annual Frequency Control Symposium - 1988

# VIBRATIONS OF Z-CUT RESONATOR-STRUCTURE BY FINITE ELEMENT ANALYSIS.

Y.K Yong
Department of Civil and Enviromental Engineering
Rutgers University
P.O. Box 909
Piscataway, NJ 08855-0909

P.C.Y. Lee
Department of Civil Engineering
Princeton University
Princeton, NJ 08544

S.S. Chuang
Statek Corporation
512 North Main Street
Orange, CA 92668

## SUMMARY

Vibrations of a miniature Z-cut, third overtone, extensional quartz resonator were studied. The resonator-structure consisted of a long vibrating beam, a short supporting beam, and a tuning fork. The vibrating beam and supporting beam intersected each other perpendicularly at their middle points. The ends of the supporting beam were in turn connected to the ends of tines of the tuning fork. Two corners near the base of the tuning fork were mounted on supports which had a fixed width but variable lengths. The vibrating beam is mainly excited at the third overtone of extensional mode. The resonator was produced by chemical etching of a Z-cut quartz plate. An anisotropic, four-node quadrilateral, Mindlin plate element model was employed for the study. Incompatible modes were included in the stiffness matrix to prevent shear locking. The frequency spectrum of the resonator as a function of the mounting length and its mode shapes were discussed. The resonator motional resistances for different mounting lengths were measured. Normalized strain energy ratios in the base area and tine area were computed and compared with the normalized motional resistance at various mounting lengths. Good correlations were found: the peaks in strain energy ratio corresponded with the peaks in motional resistance. These peaks occurred at regions in the frequency spectrum which reflected strong coupling with other modes of vibration. Large scatterings of the measured motional resistances were also observed at these regions. The strain energy ratio in the base and tines was proposed as a qualitative criterion for evaluating the relative changes in motional resistance of resonators with different mounting lengths and geometries.

## I. INTRODUCTION.

Practical applications for low frequency quartz resonators include those found in watches, clocks, microcomputers, pressure transducers and temperature transducers. The geometric shapes of these low frequency quartz resonators are generally more complicated than those of the thickness shear resonators. The most commonly employed modes of vibration in these resonators are flexure, extension, torsion and face shear. Classical solutions for these modes of vibration exist only for the simplest shapes, such as, a prismatic beam or rod. In reality, the vibrational modes are quite complex, as they involve not only the thickness, width and length of the vibrating element, but also the supporting and mounting structure. The quality factor, Q, of the resonator can decrease dramatically with small changes in either the vibrating element or mounting structure. For example, Tomikawa, Oyama and Konno[1] reported the use of small modifications to the basewidth of a quartz tuning fork to achieve higher resonator Q.

We study, in this paper, the vibrations of a miniature Z-cut, third overtone, extensional quartz resonator. A four node, quadrilateral, Mindlin plate element model which accommodates three translational degrees of freedom and three shear displacements is employed for the study. As is usual in a plate element, a scheme is needed to preventing shear locking. We chose to include incompatible modes in the computation of the element stiffness matrix to prevent shear locking. The main objectives of the study are to develop a fully anisotropic plate element using Mindlin's[2] first order plate equations, and to analyse the vibrations of the resonator and its relative changes in motional resistance with respect to changes in the mounting structure. The motional resistance of a resonator is inversely proportional to its Q. Hence, if the motional resistance, which is measured experimentally, could be related to the numerical results, the elastic, finite element model could be used to evaluate the relative merits of a resonator design in terms of its Q.

## II. FINITE ELEMENT FORMULATION OF THE ANISOTROPIC, MINDLIN'S FIRST ORDER PLATE EQUATIONS.

### a) MINDLIN'S FIRST ORDER PLATE EQUATIONS.

The finite element model is based on the anisotropic, Mindlin's[2] first order plate equations of motion. These equations are adequate for frequencies up to the fundamental thickness shear and its overtones. The interested reader is referred to reference 2 for the details in the development of these equations. The stress equations of motion, stress-strain relations and strain-displacement relations are stated below.

CH2588-2/88/0000-029 $1.00 C 1988 IEEE

Stress equations of motion.

$$t^{(0)}_{ij,j} + F^{(0)}_i = 2b\rho\ddot{u}^{(0)}_i, \tag{1a}$$

$$t^{(1)}_{ij,j} - t^{(0)}_{i2} + F^{(1)}_i = \frac{2b^3}{3}\rho\ddot{u}^{(1)}_i, \tag{1b}$$

$$i,j = 1,2,3.$$

The terms $t^{(0)}_{ij}$ and $t^{(1)}_{ij}$ are, respectively, the two-dimensional, zeroth and first order stresses. Partial differentiations of the stresses with respect to the thickness coordinate, $x_2$, are defined to be zero, since the stresses are functions of $x_1$, $x_3$ and t only. $F^{(n)}_i$, n = 0,1, are the two-dimensional, surface tractions. The zeroth order displacements $u^{(0)}_i$ are the translational displacements, while the first order displacements $u^{(1)}_i$ are the shear, or, rotational displacements. The plate thickness and density are 2b and $\rho$, respectively.

Stress-strain relations.

$$t^{(0)}_p = 2b\kappa_{(p)}\kappa_{(q)}C_{pq}e^{(0)}_q, \tag{2a}$$

$$t^{(1)}_p = \frac{2b^3}{3}(C_{pq} - C_{p2}C_{2q}/C_{22})e^{(1)}_q, \tag{2b}$$

$$p,q = 1,2,3,4,5,6.$$

where $\kappa_{(p)}$ are the shear correction factors with the usual values,

$$\kappa^2_2 = \kappa^2_4 = \kappa^2_6 = \frac{\pi^2}{12} \text{ and } \kappa_1 = \kappa_3 = \kappa_5 = 1. \tag{2c}$$

The parenthesis in the terms $\kappa_{(p)}$ signify suspension from summation. $C_{pq}$ are the anisotropic elastic constants referred to the plate axes. The values of the elastic constants, referred to the crystal axes, were measured and reported by Bechmann, Ballato and Lukaszek[3]. Equation (2b) is obtained by setting $t^{(1)}_2$ to zero and solving for the strain $e_2{}^{(1)}$ to yield

$$e^{(1)}_2 = -C_{2q}e^{(1)}_q/C_{22}. \tag{2d}$$

Therefore, while the stress $t^{(1)}_2$ is neglected and set to zero, the corresponding strain $e^{(1)}_2$ actually has a definite value given by Eq.(2d), and is accommodated in the first order stress-strain relations of Eq.(2b).

Strain-displacement relations.

$$e^{(0)}_{ij} = \tfrac{1}{2}(u^{(0)}_{j,i} + u^{(0)}_{i,j} + \delta_{2i}u^{(1)}_j + \delta_{2j}u^{(1)}_i), \tag{3a}$$

$$e^{(0)}_{ij} = \tfrac{1}{2}(u^{(1)}_{j,i} + u^{(1)}_{i,j}), \tag{3b}$$

where $\delta_{2i}$ is the Kronecker delta.

Boundary conditions.

Unique solutions to the set of Eqs.(1) to (3) can be obtained by specifying the essential or natural boundary conditions on C, the intersection of the right cylindrical plate boundary with the middle plane of the plate. Namely,

specify $u^{(0)}_i$ or $p^{(0)}_i = n_j t^{(0)}_{ij}$ on C, and (4a)

specify $u^{(1)}_i$ or $p^{(1)}_i = n_j t^{(1)}_{ij}$ on C, (4b)

where $p^{(n)}_i$, n = 1,2, are the edge traction vectors, and $n_j$ are the components of the vector normal to the edge of the plate.

b) VARIATIONAL EQUATION.

An integral equation suitable for finite element implementation may be derived from Eqs.(1) by employing variational principles[4,5]. This is done by multiplying Eqs.(1a) and (1b) respectively by the variational displacements $\delta u^{(0)}_i$ and $\delta u^{(1)}_i$, adding the two resulting equations, integrating with respect to the plate area and applying the two-dimensional divergence theorem to yield

$$\int_A (\delta e^{(0)}_{ij}t^{(0)}_{ij} + \delta e^{(1)}_{ij}t^{(1)}_{ij})\,dA = \int_C (\delta u^{(0)}_i p^{(0)}_i + \delta u^{(1)}_i p^{(1)}_i)\,dC - \int_A (\delta u^{(0)}_i m^{(0)}_{ij}\ddot{u}^{(0)}_j + \delta u^{(1)}_i m^{(1)}_{ij}\ddot{u}^{(1)}_j)\,dA, \tag{5}$$

where $\delta e^{(n)}_{ij}$ (n=0,1) are the variational strains,

$$m^{(0)}_{ij} = 2b\rho\delta_{ij},$$

$$m^{(1)}_{ij} = \frac{2b^3}{3}\rho\delta_{ij},$$

A is the plate area,

and C is the intersection of the right cylindrical plate boundary with the middle plane of the plate.

$u^{(n)}_i$ and $p^{(n)}_i$, n=0,1, in the line integral on the right-hand side form the essential and natural boundary conditions, respectively. Eq.(5) is employed in the finite element formulation.

c) FINITE ELEMENT MATRIX EQUATIONS.

The stresses, strains, displacements and edge tractions in Eq.(5) can be written in a vector form, that is,

$$\underline{t} = [t^{(0)}_1, t^{(0)}_2, t^{(0)}_3, t^{(0)}_4, t^{(0)}_5, t^{(0)}_6, t^{(1)}_1, t^{(1)}_3, t^{(1)}_4, t^{(1)}_5, t^{(1)}_6]^T, \tag{6a}$$

$$\underline{e} = [e^{(0)}_1, e^{(0)}_2, e^{(0)}_3, e^{(0)}_4, e^{(0)}_5, e^{(0)}_6, e^{(1)}_1, e^{(1)}_3, e^{(1)}_4, e^{(1)}_5, e^{(1)}_6]^T, \tag{6b}$$

$$\underline{u} = [u^{(0)}_1, u^{(0)}_2, u^{(0)}_3, u^{(1)}_1, u^{(1)}_2, u^{(1)}_3]^T, \tag{6c}$$

and

$$\underline{p} = [p^{(0)}_1, p^{(0)}_2, p^{(0)}_3, p^{(1)}_1, p^{(1)}_2, p^{(1)}_3]^T. \tag{6d}$$

For the finite element implementation, it is convenient to write the variational equation, Eq.(5), in matrix notation,

$$\int_A \delta\underline{e}^T \underline{t}\, dA = \int_C \delta\underline{u}^T \underline{p}\, dC - \int_A \delta\underline{u}^T \underline{m}\, \ddot{\underline{u}}\, dA, \tag{7a}$$

where $\underline{m} = \begin{bmatrix} m_{ij}^{(0)} & 0 \\ 0 & m_{ij}^{(1)} \end{bmatrix}_{6x6}$. (7b)

The constitutive relations from Eqs.(2a) and (2b) in matrix form are,

$$\underline{t} = \underline{C}\,\underline{e} \tag{8a}$$

where

$$\underline{C} = \begin{bmatrix} \bar{C}_{11} & \bar{C}_{12} & \bar{C}_{13} & \bar{C}_{14} & \bar{C}_{15} & \bar{C}_{16} & 0 & 0 & 0 & 0 & 0 \\ \bar{C}_{12} & \bar{C}_{22} & \bar{C}_{23} & \bar{C}_{24} & \bar{C}_{25} & \bar{C}_{26} & 0 & 0 & 0 & 0 & 0 \\ \bar{C}_{13} & \bar{C}_{23} & \bar{C}_{33} & \bar{C}_{34} & \bar{C}_{35} & \bar{C}_{36} & 0 & 0 & 0 & 0 & 0 \\ \bar{C}_{14} & \bar{C}_{24} & \bar{C}_{34} & \bar{C}_{44} & \bar{C}_{45} & \bar{C}_{46} & 0 & 0 & 0 & 0 & 0 \\ \bar{C}_{15} & \bar{C}_{25} & \bar{C}_{35} & \bar{C}_{45} & \bar{C}_{55} & \bar{C}_{56} & 0 & 0 & 0 & 0 & 0 \\ \bar{C}_{16} & \bar{C}_{26} & \bar{C}_{36} & \bar{C}_{46} & \bar{C}_{56} & \bar{C}_{66} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \tilde{C}_{11} & \tilde{C}_{13} & \tilde{C}_{14} & \tilde{C}_{15} & \tilde{C}_{16} \\ 0 & 0 & 0 & 0 & 0 & 0 & \tilde{C}_{13} & \tilde{C}_{33} & \tilde{C}_{34} & \tilde{C}_{35} & \tilde{C}_{36} \\ 0 & 0 & 0 & 0 & 0 & 0 & \tilde{C}_{14} & \tilde{C}_{34} & \tilde{C}_{44} & \tilde{C}_{45} & \tilde{C}_{46} \\ 0 & 0 & 0 & 0 & 0 & 0 & \tilde{C}_{15} & \tilde{C}_{35} & \tilde{C}_{45} & \tilde{C}_{55} & \tilde{C}_{56} \\ 0 & 0 & 0 & 0 & 0 & 0 & \tilde{C}_{16} & \tilde{C}_{36} & \tilde{C}_{46} & \tilde{C}_{56} & \tilde{C}_{66} \end{bmatrix} \tag{8b}$$

where $\bar{C}_{pq} = 2b\, k_{(p)} k_{(q)} C_{pq}$

and $\tilde{C}_{pq} = \frac{2b^3}{3}\left( C_{pq} - C_{2p}C_{2q}/C_{22} \right)$.

The matrix form of the strain-displacement relations is

$$\underline{e} = \underline{d}\,\underline{u}, \tag{9a}$$

where $\underline{d} = \begin{bmatrix} ,x_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & ,x_3 & 0 & 0 & 0 \\ 0 & ,x_3 & 0 & 0 & 0 & 1 \\ ,x_3 & 0 & ,x_1 & 0 & 0 & 0 \\ 0 & ,x_1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & ,x_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & ,x_3 \\ 0 & 0 & 0 & 0 & ,x_3 & 0 \\ 0 & 0 & 0 & ,x_3 & 0 & ,x_1 \\ 0 & 0 & 0 & 0 & ,x_1 & 0 \end{bmatrix}$. (9b)

Matrix $\underline{d}$ is a linear differential operator, and its elements $,x_i$ ($i$ = 1 or 3) signify partial differentiation with respect to the coordinates of the plate.

The displacement vector $\underline{u}$ is approximated by a set of nodal displacements $\underline{Q}$ and interpolation functions $\underline{N}$,

$$\underline{u} = \underline{N}\,\underline{Q} \tag{10}$$

Hence, the variational displacements in Eq.(7a) may be written in terms of the variational nodal displacements,

$$\delta\underline{u}^T = \delta(\underline{N}\,\underline{Q})^T = \delta\underline{Q}^T\,\underline{N}^T. \tag{11}$$

By substituting Eq.(10) into Eq.(9a), we can express the strain-displacement relations in terms of $\underline{N}$ and $\underline{Q}$, namely,

$$\underline{e} = \underline{d}\,\underline{N}\,\underline{Q} = \underline{B}\,\underline{Q}. \tag{12}$$

where $\underline{B}$ is generally known as the strain-displacement matrix, since its elements represent strains values due to a unit vector of nodal displacements. The variational strain $\delta\underline{e}^T$ on the left-hand side of Eq.(7a) is, therefore,

$$\delta\underline{e}^T = \delta(\underline{B}\,\underline{Q})^T = \delta\underline{Q}^T\,\underline{B}^T. \tag{13}$$

By putting Eq.(12) into Eq.(8a), we obtain the stress-displacement matrix equation,

$$\underline{t} = \underline{C}\,\underline{B}\,\underline{Q}. \tag{14}$$

We use Eqs.(11), (13) and (14) to express the integral relation of Eq.(7a) in terms of the nodal displacements and interpolation functions,

$$\delta\underline{Q}^T\left(\int_A \underline{B}^T\underline{C}\,\underline{B}\, dA\right)\underline{Q} = \delta\underline{Q}^T\left(\int_C \underline{N}^T\underline{p}\, dC - \int_A \underline{N}^T\underline{m}\,\underline{N}\, dA\, \ddot{\underline{Q}}\right). \tag{15}$$

After dropping the common variational displacements on both sides of the equation, we observe that Eq.(15) is actually a matrix equation of integral functions,

$$\underline{K}\,\underline{Q} = \underline{P} - \underline{M}\,\ddot{\underline{Q}}. \tag{16}$$

where $\underline{K} = \int_A \underline{B}^T\underline{C}\,\underline{B}\, dA$, $\underline{P} = \int_C \underline{N}^T\underline{p}\, dC$ and $\underline{M} = \int_A \underline{N}^T\underline{m}\,\underline{N}\, dA$ are the element stiffness matrix, nodal force vector and element consistent mass matrix, respectively.

The method for assembling the element equations into one global matrix equation is a standard procedure which can be found in textbooks on the finite element method. The global finite element matrix equation has a similar form to Eq.(16),

$$\underline{K}_G\,\underline{Q}_G = \underline{P}_G - \underline{M}_G\,\ddot{\underline{Q}}_G. \tag{17}$$

where the subscript G denotes a global matrix. Since we are concerned with only free vibration problems, the term $\underline{P}_G$ in Eq.(17) may be dropped with the stipulation that the boundary conditions are satisfied. For a simple harmonic motion with a radial frequency of $\omega$,

$$\ddot{\underline{Q}}_G = -\omega^2\,\underline{Q}_G \tag{18}$$

Hence, using Eq.(18), we arrive from Eq.(17) the general eigenvalue problem for free vibrations,

$$\underline{K}_G\,\underline{Q}_G = \omega^2\,\underline{M}_G\,\underline{Q}_G \tag{19}$$

d) FOUR-NODE, ISOPARAMETRIC QUADRILATERAL ELEMENT.

The 4-node, isoparametric quadrilateral element is one of the simplest element for a plate. Their shape functions[6] are

$$N_1(\xi,\eta) = \tfrac{1}{4}(1-\xi)(1-\eta),$$

$$N_2(\xi,\eta) = \tfrac{1}{4}(1+\xi)(1-\eta),$$

$$N_3(\xi,\eta) = \tfrac{1}{4}(1+\xi)(1+\eta),$$

and $N_4(\xi,\eta) = \tfrac{1}{4}(1-\xi)(1+\eta)$, (20)

where $\xi$ and $\eta$ are the local coordinates. The matrix of shape functions $\underline{\mathbf{N}}$ in Eq.(10), therefore, takes the form,

$$\underline{\mathbf{N}} = [\underline{n}_1, \underline{n}_2, \underline{n}_3, \underline{n}_4], \tag{21}$$

where $\underline{n}_i = \underline{\mathbf{I}}\, N_i, \quad i=1,2,3,4. \tag{22}$

The matrix $\underline{\mathbf{I}}$ is an identity matrix with a size that is dependent on the number of degrees of freedom per node in the plate element. We have six displacement degrees of freedom per node in our element model, hence, the size of matrix $\underline{\mathbf{N}}$ is 6x24. The nodal displacement vector $\underline{\mathbf{Q}}$ could now be expressed in terms of four component vectors, namely,

$$\underline{\mathbf{Q}} = [\underline{q}_1, \underline{q}_2, \underline{q}_3, \underline{q}_4]^T. \tag{29}$$

where $\underline{q}_i$ represent the vectors of six displacements at each of the four nodes of the quadrilateral.

We choose to use an incompatible modes scheme[7] to "soften" an otherwise overconstrained quadrilateral plate element and prevent shear-locking. In the incompatible modes scheme, two extra shape functions are included in the formulation of the element stiffness matrix. They are

$$N_5(\xi,\eta) = 1-\xi^2,$$
$$\text{and } N_6(\xi,\eta) = 1-\eta^2. \tag{30}$$

They do not belong to any nodes. The element stiffness terms associated with these functions are removed by static condensation. As long as the the opposite sides of the quadrilateral elements are parallel, that is, form a rectangle or parallelogram, the finite element solutions will converge satisfactorily[7]. A finite element computer program[8] was developed and an efficient eigenvalue solver for sparse matrices was added to the program.

## III. RESULTS.

### a) FINITE ELEMENT MODEL FOR THE EXTENSIONAL RESONATOR.

Figure 1 shows the finite element mesh generated for the miniature, Z-cut third overtone extensional resonator. The resonator-structure consists of three parts: (1) a long vibrating beam, (2) a short supporting beam and (3) a tuning fork. The vibrating beam and supporting beam intersect each other perpendicularly at their middle points. The ends of the supporting beam are in turn connected to the ends of tines of the tuning fork. The two corners near the base of the tuning fork are mounted on two supports. The resonator was produced by chemical etching of a Z-cut quartz plate with a thickness of 178 $\mu$m.

The width of the mounts at the two supports are kept constant at 292 $\mu$m, while the length along the $x_1$ axis is variable. In the numerical computations, the entire resonator-structure is free of traction and specified displacement, except at the mounts, where

$$u_1^{(0)} = 0,$$
$$u_2^{(0)} = 0,$$
$$\text{and } u_3^{(0)} = 0. \tag{31}$$

Hence, the strain energies at the mounting areas are due solely to shear displacements since Eq.(31) specifies that the translational displacements are zero.

### b) FREQUENCY SPECTRUM OF THE RESONATOR WITH RESPECT TO MOUNTING LENGTH.

For applications, the vibrating beam of the resonator-structure is mainly excited at the third overtone of extensional mode. The finite element model of Section III(a) is used in the computer program to calculate resonant frequencies in the frequency range of 1.81 to 1.83 MHz for different mounting lengths, $D_p$. Figure 2 shows the resulting frequency spectrum. We are primarily interested in the characteristics of the flat branches along the horizontal, mounting length axis because the upper flat branch describes the third overtone extensional mode. Its six displacement mode shapes at $D_p$ = 180 $\mu$m are exhibited in fig.3. The vertical axis represents the magnitude of a normalized displacement component. We observe that the predominant feature of these mode shapes is the extensional displacement of the vibrating beam. There is a small coupling with displacements $u_2^{(0)}$ and $u_1^{(1)}$.

The lower flat branch gives the frequencies for the fifth overtone torsional mode of the vibrating beam. Its mode shapes are shown in fig.4 for $D_p$ = 180 $\mu$m. Although the torsional vibration is predominantly made up of $u_2^{(0)}$ and $u_3^{(1)}$ displacements, it is also highly coupled to $u_1^{(0)}$, $u_3^{(0)}$ and $u_1^{(1)}$.

The flat branches are intersected by vertical branches at $D_p$ values of, approximately, 110, 225, 390 and 650 $\mu$m. For convenience, these vertical branches are identified by the symbol $V_n$, where n = 1,2,3,4,5,6. Their mode shapes are shown in figs.5 to 10. We provide a brief verbal description of the vibrational patterns,

Branch $V_1$: Antisymmetric, 6th overtone flexure of the tines, coupled with torsion of the vibrating beam (fig.5),

Branch $V_2$: Symmetric, 6th overtone flexure of the tines, coupled with extension of the vibrating beam (fig.6),

Branch $V_3$: 10th overtone flexure of the vibrating beam, coupled with flexure of the tines (fig.7),

Branch $V_4$: 12th overtone, face shear vibration of the vibrating beam (fig.8),

Branch $V_5$: Antisymmetric, 5th overtone flexure of the tines, coupled with torsion of the vibrating beam (fig.9),

and Branch $V_6$: Symmetric, 5th overtone flexure of the tines, coupled with extension of the vibrating beam (fig.10).

### c) COMPARISION OF NUMERICAL RESULTS WITH THE MEASURED, MOTIONAL RESISTANCE OF RESONATOR.

Using the left-hand side of Eq.(16), we could calculate the strain energy in each element of the

model for any given mounting length. The sum of strain energies in the supporting structure and vibrating beam were computed and divided by the total strain energy in the resonator to obtain strain energy ratios. As expected, we observed that the strain energy ratio in the vibrating beam was the highest. The supporting structure of the vibrating beam could be separated into two regions: the base area and tine area. The base area could be further divided into the mounting base area, which is the area mounted on the supports, and the free base area, which consists of the unmounted base area.

Figure 11 shows the measured motional resistance of the resonator for changing mounting lengths. The operating pressure was 1 torr. For comparison purposes, the resistances were normalized by dividing them by a normalizing factor of 3 kΩ. The data were compared against the strain energy ratios in the base and tines.

Figure 12 gives a plot of the strain energy ratio in the mounting base versus the mounting length, along with the normalized motional resistance. The energy ratio was normalized by dividing it by a normalizing factor of 0.012. What appeared to be random scattering of resistances in fig.11 now seemed to have more structure; three peaks in the data correlated well with the peaks in strain energy ratio. Figure 13 shows the strain energy ratio in the base with a normalizing factor of 0.1. We observe again three peaks which coincide with the peaks in measured resistance. Another comparison was made in fig.14 against the strain energy ratio of the supporting structure (base and tines) with a normalizing factor of 0.3. Among the energy ratios considered in fig.12 to 14, the strain energy ratio of the base and tines in fig.14 captures best the relative changes in motional resistance.

Upon comparison of the frequency spectrum in fig.2 with figs.11 to 14, we note immediately that the peaks in strain energy ratios and scattering of resistances occur at mounting lengths where there are strong couplings with other modes of vibration. Hence, the mounting lengths which induce strong couplings should be avoided, and the strain energy ratio in the base and tines could be used as an indicator of the relative changes in motional resistance with respect to mounting lengths and resonator geometries.

### d) CHANGING THE RESONATOR THICKNESS TO DECREASE MOTIONAL RESISTANCE.

One possible way of reducing the strain energy in the base and tines, and, therefore, decreasing the motional resistance is to change the thickness of the resonator. We have seen from fig.2, where the thickness to width ratio of the vibrating beam is 1.98, the extensional mode frequency is fairly close to the fifth overtone torsion. By changing the thickness to width ratio, but keeping the cross-sectional area constant, the torsional frequency can be shifted away from the extensional frequency. This may reduce vibrations in the base and tines.

The strain energy ratio of the base and tines in a traction-free resonator was computed for different thickness to width ratios of vibrating beam, and plotted in the bar graph of fig.15. When we compare two resonators with thickness to width ratios of 1.98 and 1.02, we observe that the latter has a base and tines energy ratio which is less than one half of the former. The two resonators are further compared in fig.16 for mounting lengths of 150, 300 and 450 μm. The latter resonator (thickness to width ratio of vibrating beam = 1.02) has a lower base and tines energy ratio at mounting lengths of 150 and 450 μm. At the mounting length of 150 μm, the energy ratio is about one fifth of the former resonator. Hence, using a mounting length of 150 μm, the latter resonator would have significantly lower motional resistance and higher resonator Q.

## IV. SUMMARY AND CONCLUSIONS.

A finite element program incorporating Mindlin's first order plate equations was developed. Calculations were performed for a miniature, third overtone extensional, Z-cut resonator. Its frequency spectrum as a function of the mounting length and displacement mode shapes were discussed. The resonator motional resistances for different mounting lengths were measured.

Normalized strain energy ratios in the base area and tine area were computed and compared with the normalized motional resistance at various mounting lengths. Good correlations were found: the peaks in strain energy ratio corresponded with the peaks in motional resistance. These peaks occurred at regions in the frequency spectrum which reflected strong coupling with other modes of vibration. Large scattering of measured motional resistance were also observed at these regions. The mounting lengths which introduced large increases in motional resistance also created greater vibrations of the tines and base. Since the motional resistance is inversely proportional to the quality factor, Q, of the resonator, the strain energy ratio in the base and tines (the supporting structure) could be used as a relative criterion for judging the Q of a certain resonator design.

## REFERENCES


1) Tomikawa,Y., Oyama, S. and Konno, M. "A quartz crystal tuning fork with modified basewidth for a high quality factor: finite element analysis and experiments," IEEE Transactions on Sonics and Ultrasonics, Vol.SU-29, No.6, July 1982, pp 217-223.

2) Mindlin, R.D., "Introduction to the Mathematical Theory of Vibrations of Elastic Plates," Monograph, U.S. Army Signal Corps Engineering Laboratories, Fort Monmouth, NJ, 1955.

3) Bechmann,R., Ballato,A.D. and Lukaszek,T.J., "Higher-Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha-Quartz," Proc. IRE 50, 1962, pp 1812-1822.

4) Yong, Y.K., "On the use of 1-D finite elements for the temperature behavior of a contoured and partially plated SC-cut resonator," Proceedings of the IEEE Ultrasonics Symposium, 1987, pp 353-358.

5) Lee, P.C.Y. and Tang, M.S., "Initial stress field and resonance frequencies of incremental vibrations in crystal resonators by finite element method," Proceedings of the 40th Annual Frequency Control Symposium, 1986, pp 152-160.

6) Hughes, T.J.R., The Finite Element Method, Linear Static and Dynamic Finite Element Analysis, Prentice-Hall, Englewood Cliffs, New Jersey, 1987, pp 112-120.

7) Hughes, T.J.R., The Finite Element Method, Linear Static and Dynamic Finite Element Analysis,

Prentice-Hall, Englewood Cliffs, New Jersey, 1987, pp 242-246.

8) Yong, Y.K., "A Lagrangean, High Frequency Plate Element for the Static Temperature Behavior of Low Frequency Quartz Resonators," Proceedings of the 41st Annual Symposium on Frequency Control, 1987, pp 303-310.

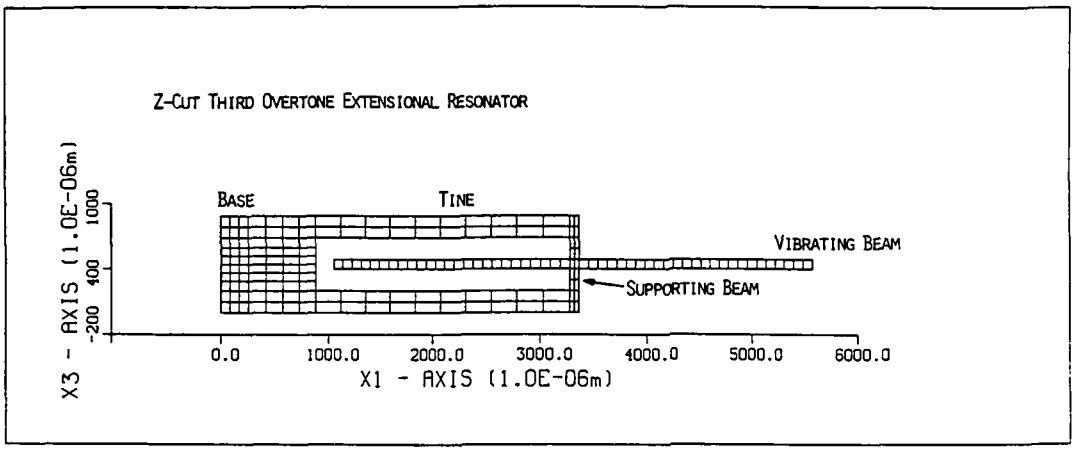



FIG.1 FINITE ELEMENT MESH FOR THE Z-CUT THIRD OVERTONE EXTENSIONAL RESONATOR.

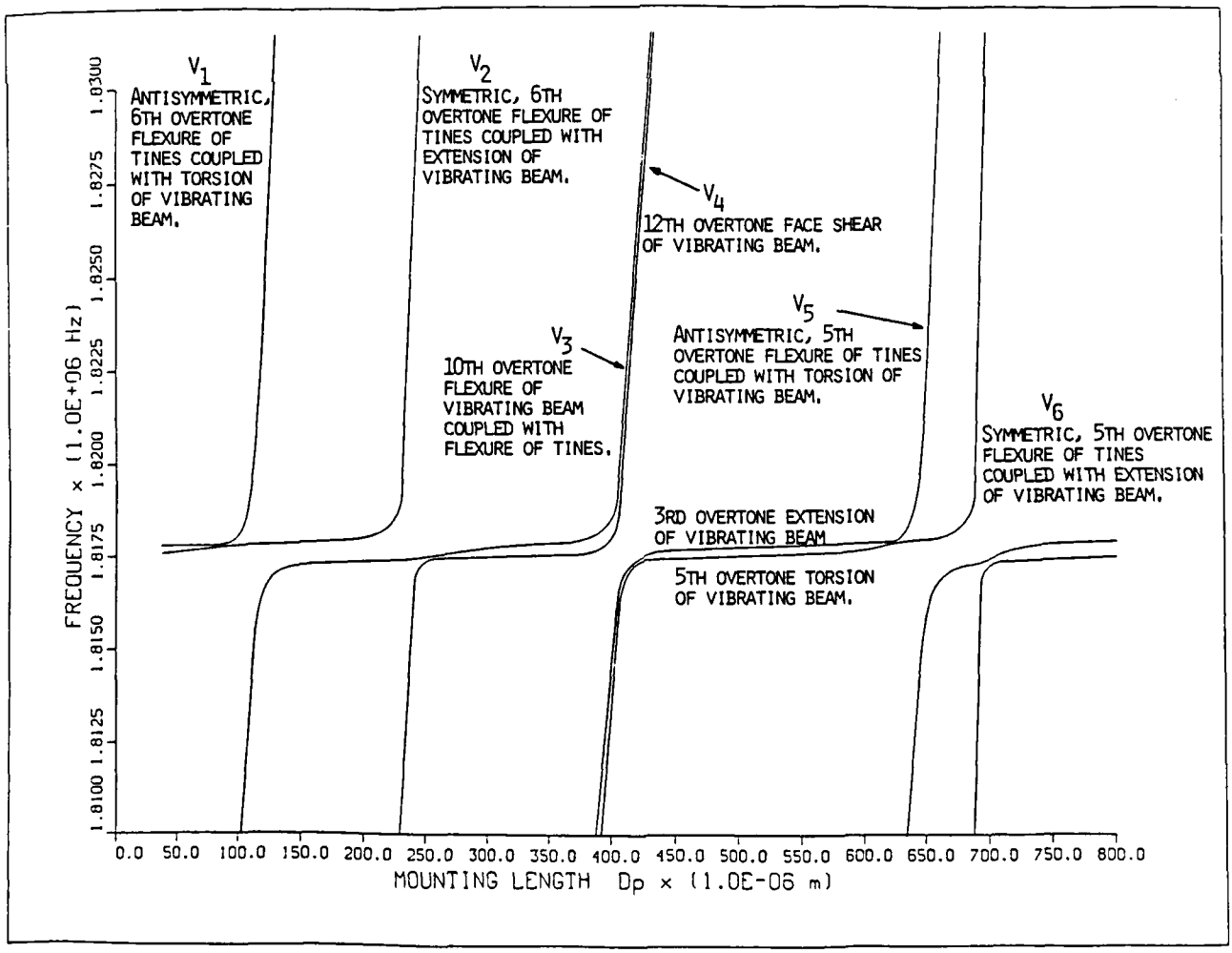



FIG.2 FREQUENCY SPECTRUM OF RESONATOR AS A FUNCTION OF THE MOUNTING LENGTH.

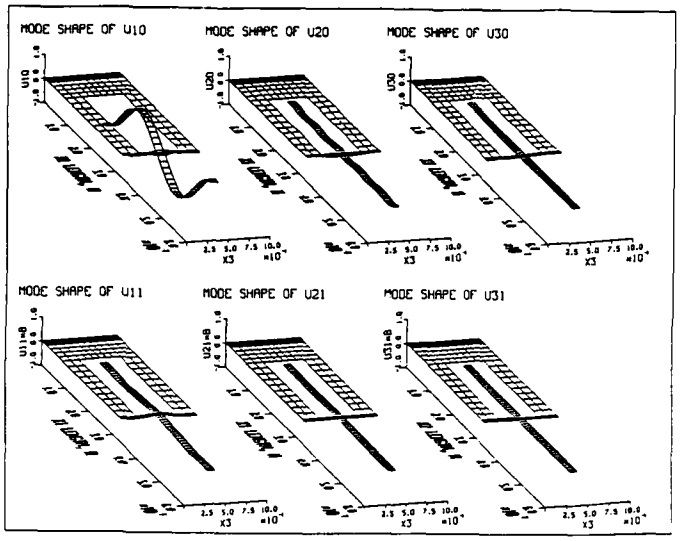



FIG.3 TYPICAL MODE SHAPES OF THE 3RD OVERTONE EXTENSIONAL VIBRATION OF THE VIBRATING BEAM.

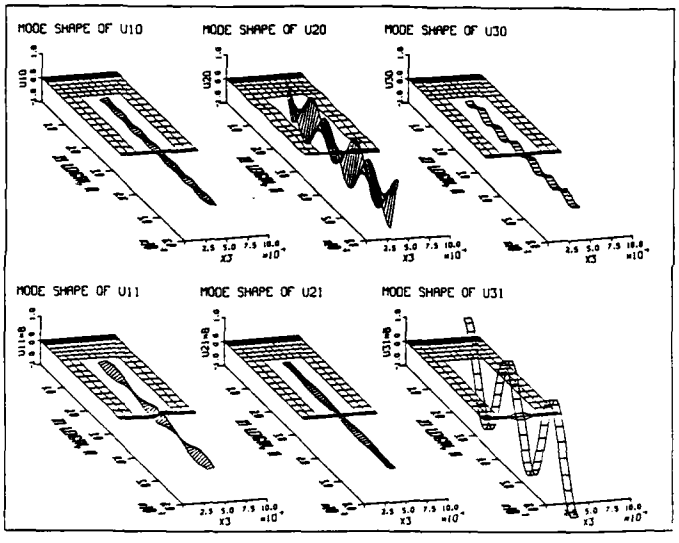



FIG.4 TYPICAL MODE SHAPES OF THE 5TH OVERTONE TORSIONAL VIBRATION OF THE VIBRATING BEAM.

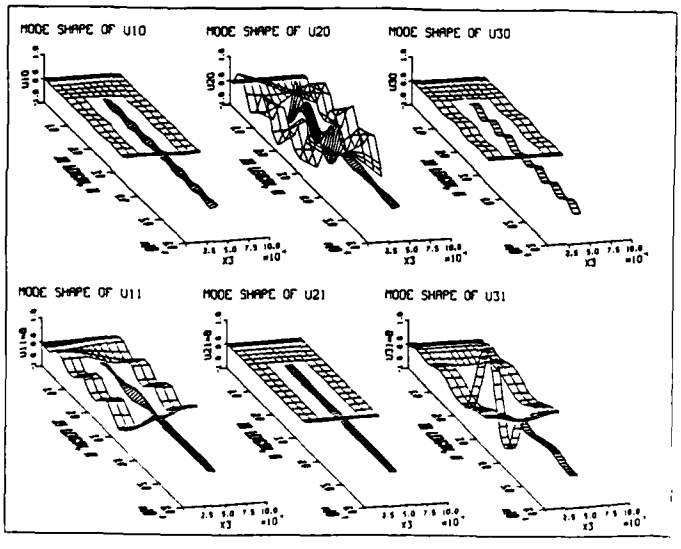



FIG.5 TYPICAL MODE SHAPES OF THE BRANCH $V_1$ OF THE RESONATOR FREQUENCY SPECTRUM.

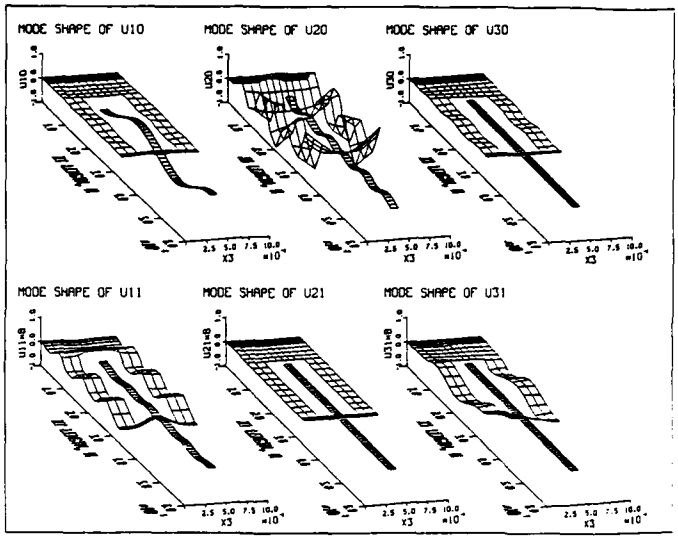



FIG.6 TYPICAL MODE SHAPES OF THE BRANCH $V_2$ OF THE RESONATOR FREQUENCY SPECTRUM.

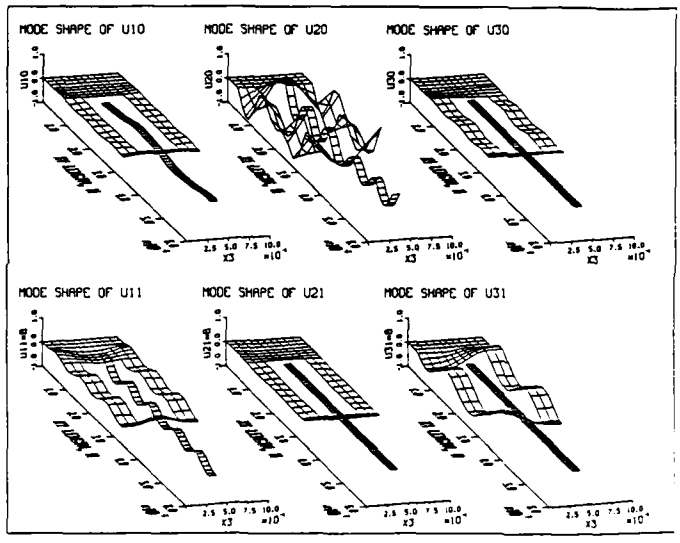



FIG.7 TYPICAL MODE SHAPES OF THE BRANCH $V_3$ OF THE RESONATOR FREQUENCY SPECTRUM.

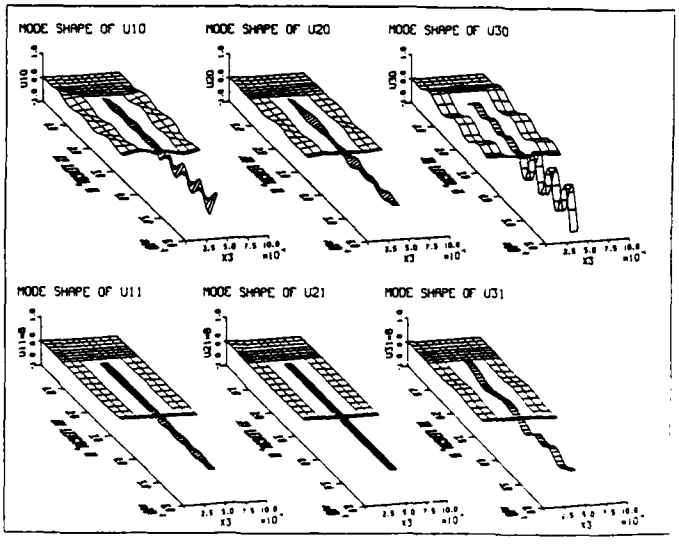



FIG.8 TYPICAL MODE SHAPES OF THE BRANCH $V_4$ OF THE RESONATOR FREQUENCY SPECTRUM.

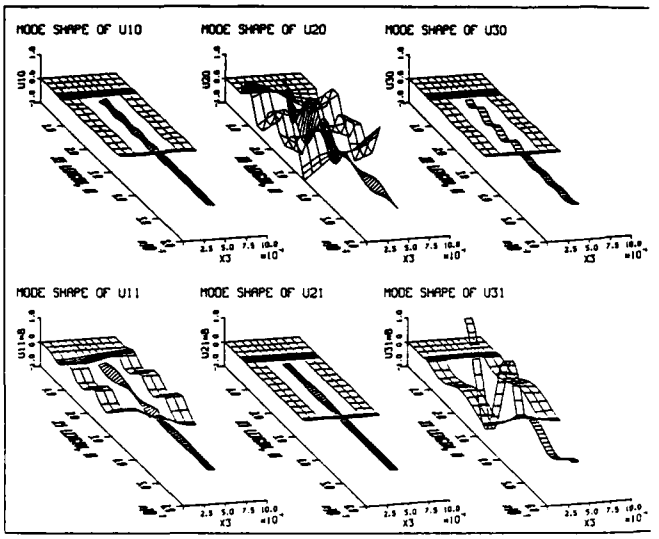



FIG.9 TYPICAL MODE SHAPES OF THE BRANCH $V_5$ OF THE RESONATOR FREQUENCY SPECTRUM.

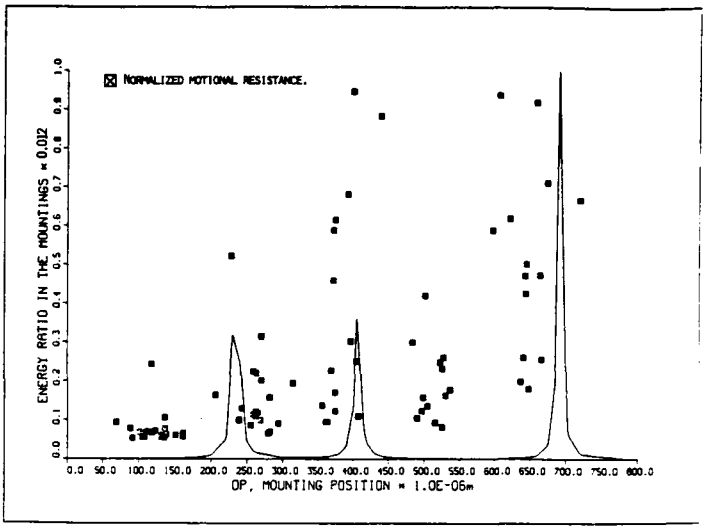



FIG.12 NORMALIZED, STRAIN ENERGY RATIO OF THE MOUNTING AREA AS A FUNCTION OF THE MOUNTING LENGTH.

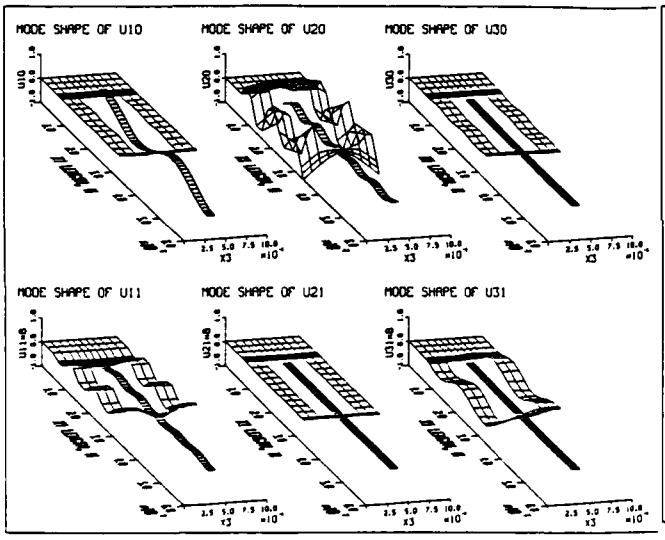



FIG.10 TYPICAL MODE SHAPES OF THE BRANCH $V_6$ OF THE RESONATOR FREQUENCY SPECTRUM.

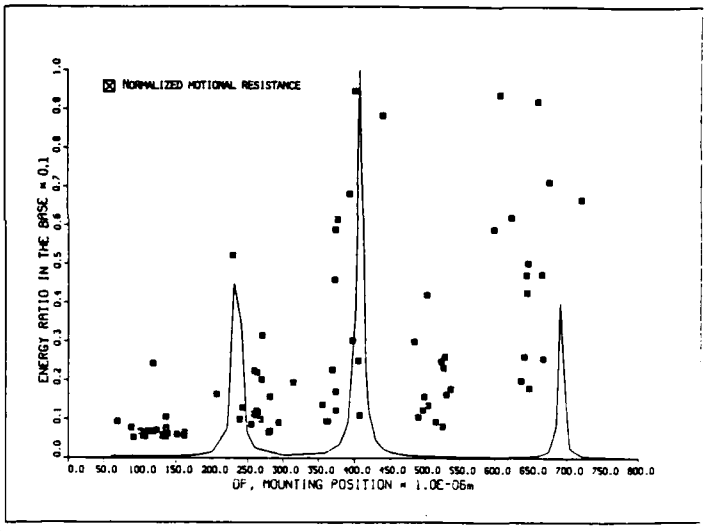



FIG.13 NORMALIZED, STRAIN ENERGY RATIO IN THE BASE AREA AS A FUNCTION OF THE MOUNTING LENGTH.

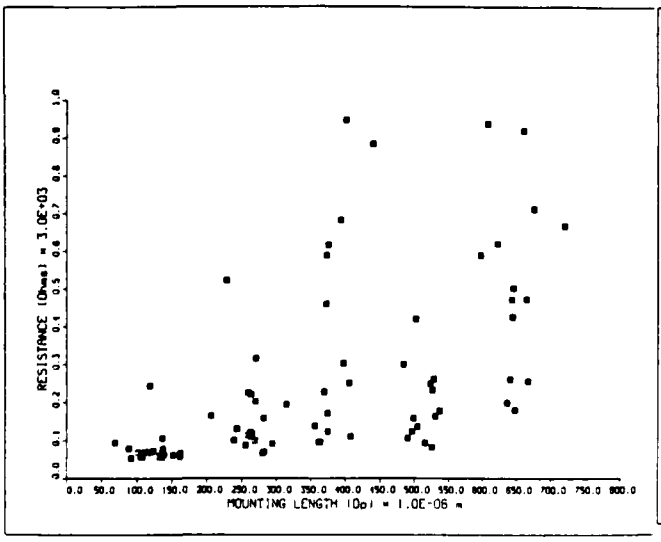



FIG.11 MEASURED MOTIONAL RESISTANCE OF RESONATOR FOR DIFFERENT MOUNTING LENGTHS.

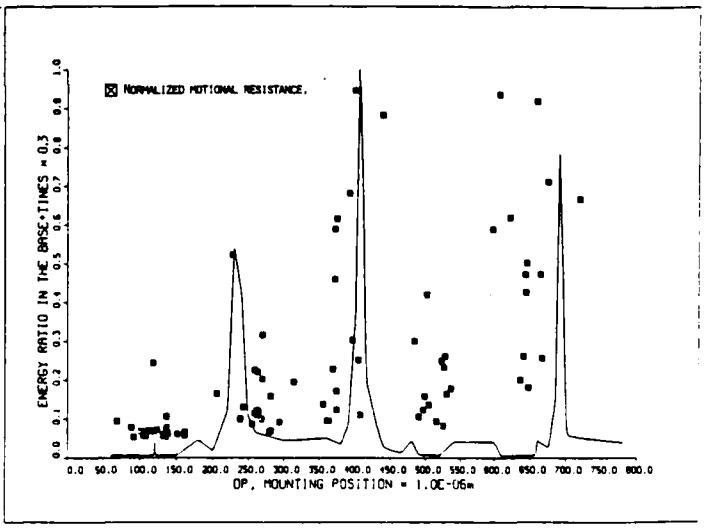



FIG.14 NORMALIZED, STRAIN ENERGY RATIO IN THE BASE AND TINES AREA AS A FUNCTION OF THE MOUNTING LENGTH.

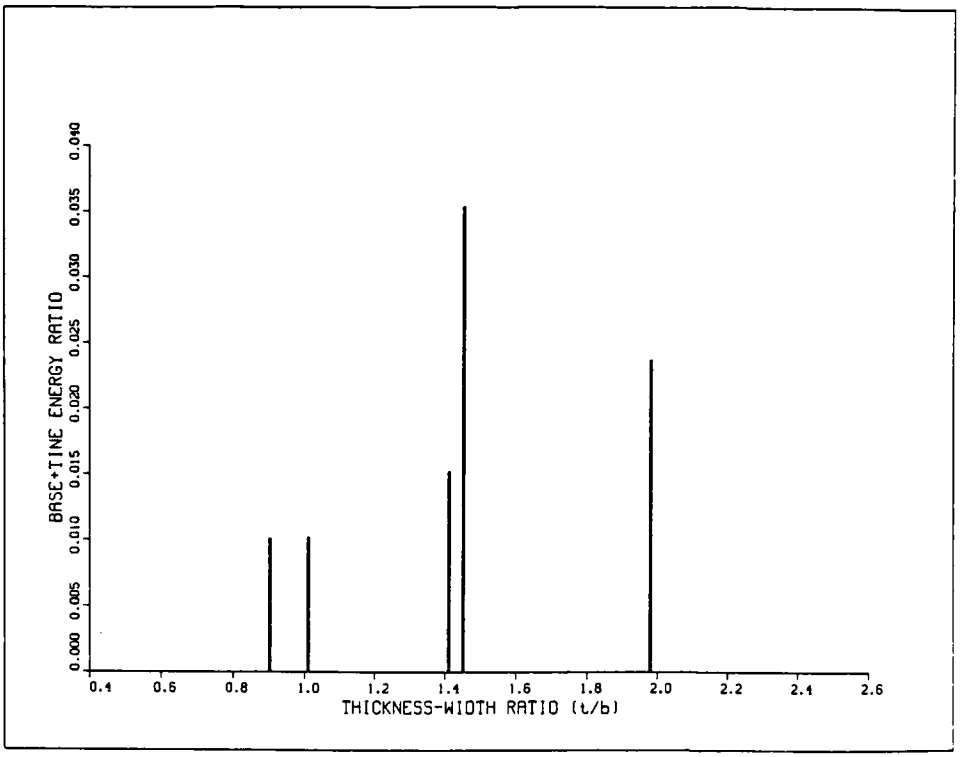



FIG.15 STRAIN ENERGY RATIO IN THE BASE AND TINES AREA OF A TRACTION-FREE RESONATOR FOR VARIOUS THICKNESS TO WIDTH RATIOS OF THE VIBRATING BEAM.

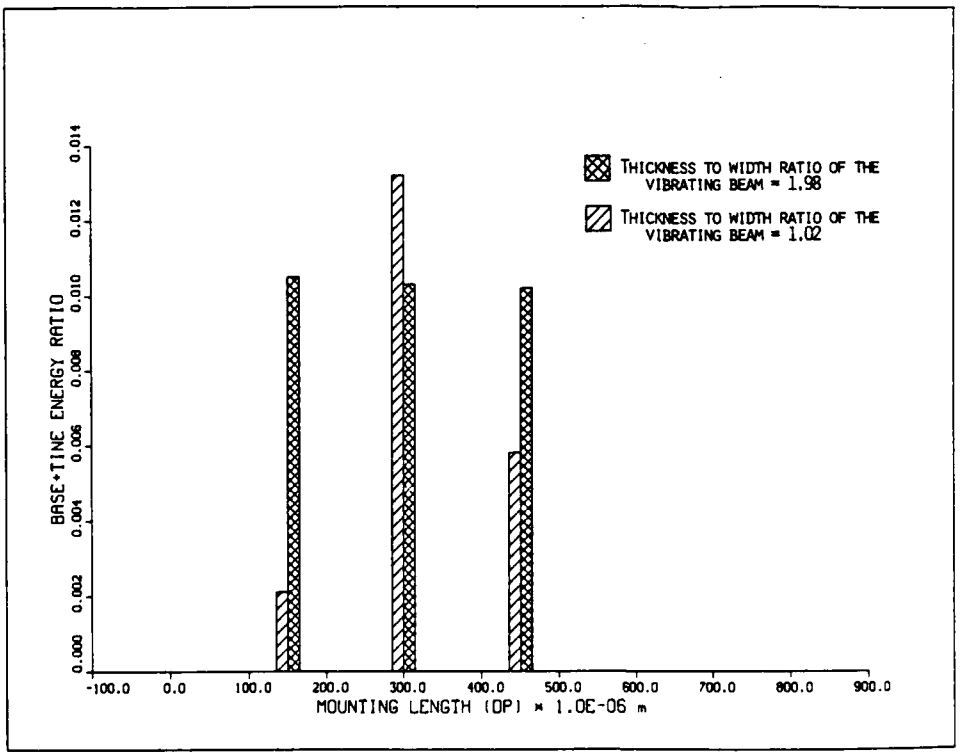



FIG.16 BAR GRAPH OF THE STRAIN ENERGY RATIO IN THE BASE AND TINES AREA FOR TWO RESONATORS.

42nd Annual Frequency Control Symposium - 1988

# RIGOROUS TWO-DIMENSIONAL EQUATIONS FOR THE ANALYSIS OF CONTOURED CRYSTAL RESONATORS

R C Peach

GEC Hirst Research Centre, East Lane, Wembley, Middlesex, HA9 7PP, UK

## SUMMARY

Plano-complex and bi-convex plate geometries are almost invariably employed for high Q quartz resonators at frequencies below 10 MHz. However, despite their importance, the existing theoretical treatments for such resonators are less satisfactory than those for plane parallel devices. The usual approximation assumes "slow" thickness variation, and replaces the constant thickness parameter in the flat plate equations by a varying quantity; this method gives no estimate of the dependence on rate of thickness variation, and makes no distinction between plano-convex and bi-convex plates. In this paper, a generalisation of previous techniques is employed to deduce rigorous two-dimensional equations for a piezoelectric plate resonator with arbitrary contour and crystallographic orientation. Methods are proposed for constructing approximate equations, involving only a finite set of mode amplitudes, from the general two-dimensional equations. The effects of the various mode coupling terms are considered for AT and SC-cut resonators of both plano-convex and bi-convex type. Approximate solutions are deduced for simple contoured plate geometries by a finite element approach. The degree of generality offered by the finite element method is required for the solution of the equations; in addition, it allows realistic boundary conditions to be used at the plate edges, and provides improved estimates of resonator Q factor.

## 1. INTRODUCTION

At frequencies below 10 MHz, resonators are usually manufactured in the form of plano-convex or bi-convex plates so as to provide the required energy trapping. The conventional theory of such resonators assumes that the thickness dependence of the acoustic motion is the same as that of an infinite parallel plate, and describes the transverse amplitude dependence by the approximate parallel plate equation with the constant thickness replaced by a variable parameter [1]. Up to a point, such a theory works quite well, but it is based on an assumption about the rate of thickness variation being 'slow', without providing any indication of how slow it should actually be.

The work described in this paper was motivated by observations on SC-cut resonators, which suggested that performance was very poor for excessively small radii of curvature, even though it was not predicted by the simple theory. In Section 2, general, rigorous equations are derived for an arbitrary contoured plate fabricated from any piezoelectric material; these are in the form of an infinite set of two-dimensional, coupled second order differential equations which depend not only on the plate thickness, but on the thickness gradients as well. The simple theory may be recovered if a single mode approximation is assumed, and terms involving thickness gradients are neglected. In the present work, solutions to the equations are deduced using the finite element method (FEM) as described in Section 3. This is in many ways more complex than the classical variational approach previously adopted [2], but it does offer greatly improved flexibility in the handling of resonator geometries and in the treatment of edge boundary conditions. The system, as presently configured, can analyse a general contoured resonator, either plano-convex or bi-convex, with an arbitrarily specified number of modes. In principle, therefore, the analysis may be performed with any chosen degree of rigour, the only limitation being the available computer resources. The possible implications and consequences of the new sytem have, as yet, only received very limited study, and these results are summarised in Section 4.

## 2. DERIVATION OF THE TWO-DIMENSIONAL EQUATIONS

The equations for the contoured plate are based upon the techniques previously developed for flat plates [2,3]. In this earlier work the two-dimensional equations were deduced directly from the three-dimensional equations of motion by expanding the thickness dependence in terms of a complete set of basis functions; however, this is not the only approach, the theory can be developed by using the same expansion to transform the three dimensional Lagrangian for the problem. The Lagrangian, or variational, form is particularly convenient for the finite element analysis, and this method will therefore be used in this paper, even though the large amount of implicit, formal machinery means that it offers less direct insight than the equation of motion approach [4,5].

First, the compact notation used in previous papers will be introduced to simplify the subsequent formulas. The generalised 4-D displacement vector $\tilde{u}_i$, is defined by

$$\begin{aligned} \tilde{u}_i &= u_i, \quad i = 1,2,3 \\ \tilde{u}_4 &= \phi, \end{aligned} \tag{1}$$

where $u_i$ is the normal displacment vector, and $\phi$ is the quasi-static potential. The generalised stiffness tensor $\tilde{c}_{ijkl}$ with dimensions 4x3x4x3 is defined as

$$\begin{aligned} \tilde{c}_{ijkl} &= c_{ijkl} \qquad i,j,k,l = 1,2,3 \\ \tilde{c}_{4jkl} &= e_{jkl} \qquad j,h,l = 1,2,3 \\ \tilde{c}_{ij4l} &= e_{lij} \qquad l,i,j = 1,2,3 \\ \tilde{c}_{4j4l} &= -\varepsilon_{jl} \qquad j,l = 1,2,3, \end{aligned} \tag{2}$$

where $c_{ijkl}$, $e_{jkl}$, and $\varepsilon_{jl}$ are the usual stiffness constants, piezoelectric stiffness constants, and dielectric permittivities respectively. In addition, we may define a generalised stress and a generalised mass tensor by

$$\begin{aligned} \tilde{T}_{ij} &= T_{ij} \qquad i,j = 1,2,3 \\ \tilde{T}_{4j} &= D_j \qquad j = 1,2,3, \end{aligned} \tag{3}$$

and

$$\rho_{ij} = \rho[\delta_{ij}-\delta_{i4}\delta_{j4}] \qquad i,j = 1,2,3,4, \tag{4}$$

where $T_{ij}$ is the usual stress tensor, $D_j$ is the electric displacement, and $\rho$ is the mass density. With

CH2588-2/88/0000-038 $1.00 C 1988 IEEE

these definitions, it is easily shown that the stress equations and the equations of motion are:

$$\tilde{T}_{ij} = \tilde{c}_{ijk\ell}\tilde{u}_{k,\ell} \tag{5}$$

and $$\rho_{ij}\frac{\partial^2\tilde{u}_j}{\partial t^2} = \tilde{c}_{ijk\ell}\tilde{u}_{k,\ell j}\ . \tag{6}$$

To simplify the notation the ~,s will henceforth be omitted, and it will be understood that, unless specified to the contrary, the stresses and displacements etc. refer to the generalised quantities defined above.

In general, the piezoelectric material will satisfy the equations of motion (6), and certain boundary conditions at the surface. These boundary conditions may take one of two forms, either a component of the displacement, $u_i$, is specified (essential boundary condition), or the corresponding normal component of the stress, $T_{ij}n_j$, is specified (natural boundary condition). It is perfectly possible for some components to satisfy essential boundary conditions and others to satisfy natural boundary conditions on a given region of the surface; for example, with stress free short circuit boundary conditions, $T_{ij}n_j=0$ for $i=1,2,3$, but $u_4=0$. The general solution to the problem may be derived by requiring the Lagrangian function, L, to be stationary with respect to variations in the displacement field $u_i$

$$L = \int_{t_1}^{t_2}\int_V \mathcal{L}\,dV dt + \int_{t_1}^{t_2}\int_\Omega T^o_{ij}\,(u_i-u_i^o)dS_j dt, \tag{7}$$

where the Lagrangian density, $\mathcal{L}$ , is given by

$$\mathcal{L} = \frac{1}{2}\rho_{ij}\dot{u}_i\dot{u}_j - \frac{1}{2}c_{ijk\ell}u_{i,j}u_{k,\ell} \tag{8}$$

In equation (7) the material has a volume V and a bounding surface $\Omega$; $dS_j$ is an element of surface area with direction along the outward normal ($dS_j=n_j dS$). The surface integral in (7) requires some interpretation; if a given component, i, satisfies the natural boundary conditions at a particular point, then $T^o_{ij}n_j$ is a constant equal to the specified value, and at this point $u_i^o$ may be put equal to zero; however, if essential boundary conditions hold, then $u_i^o$ is set equal to the fixed value, and $T^o_{ij}$ is simply the stress computed from $u_i$ in the normal way using Eqn (5).

The variation $\delta L$ produced by a variation $\delta u_i$, which is assumed to vanish at the time limits $t_1$ and $t_2$, is easily shown to be

$$\delta L = \int_{t_1}^{t_2}\int_V [-\rho_{ij}u_j + c_{ijk\ell}u_{k,\ell j}]\delta u_i dV dt + \int_{t_1}^{t_2}\int_\Omega (T_{ij}-T^o_{ij})\delta u_i dS_j dt + \int_{t_1}^{t_2}\int_\Omega \delta T_{ij}(u_i-u_i^o)dS_j dt. \tag{9}$$

In this case it has not been assumed that the variation $\delta u_i$ vanishes at the spatial boundaries, even in those regions where essential boundary conditions hold, and, as a result, additional boundary terms generated by the integration by parts are retained. Clearly the equations of motion (6) must be satisfied if the volume integral is to vanish for arbitrary $\delta u_i$. On a region of the boundary where natural conditions hold, $\delta T^o_{ij}n_j=0$, because $T^o_{ij}$ is a fixed quantity, and the first boundary integral in Eqn (9) forces the condition $(T^o_{ij}-T_{ij})n_j=0$, as required. Similarly, if essential conditions hold, the first boundary term is identically zero, and the second integral forces the condition $u_i=u_i^o$. It will be noted that the essential boundary conditions emerge as a direct consequence of the variationl principle, and are not imposed as an additional constraint, which is the more usual approach.

The specific geometry of the contoured plate is shown in Figure 1. The reference plane is taken as being normal to the local z axis, and z is therefore the thickness coordinate. The upper surface is defined by the equation

$$z = \ell_1(x,y), \tag{10}$$

and the lower surface is defined by

$$z = -\ell_2(x,y). \tag{11}$$

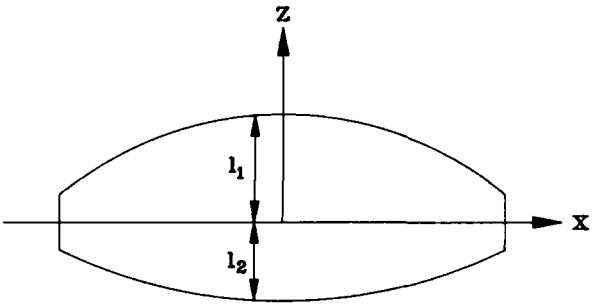



Figure 1: Contoured resonator geometry

There are only two special cases of real significance; plano-convex where $\ell_2=0$ and bi-convex where $\ell_1=\ell_2$. Defining the semi-thickness $\ell$ and the mean centre position h by

$$\ell = \frac{1}{2}(\ell_1+\ell_2), \tag{12}$$

and

$$h = \frac{1}{2}(\ell_1-\ell_2), \tag{13}$$

then a coordinate transformation

$$x' = x \qquad y' = y \qquad z' = (z-h)/\ell \tag{14}$$

may be defined; the upper and lower surfaces now become $z'=+1$ and $-1$ respectively. In the equations that follow the coordinate notations (x,y,z) and $(x_1,x_2,x_3)$ will be used interchangeably, and the indices $\alpha$ and $\beta$ will be used to denote values of 1 or 2, or alternatively implied summations over the indices 1 and 2. From Eqn (14), the equivalences

$$\frac{\partial}{\partial x_\alpha} = \frac{\partial}{\partial x'_\alpha} - \frac{1}{\ell}(h_\alpha+z'\ell_\alpha)\frac{\partial}{\partial z'}, \tag{15}$$

and

$$\frac{\partial}{\partial z} = \frac{1}{\ell}\frac{\partial}{\partial z'} \tag{16}$$

where $\alpha = 1,2$, $h_\alpha = \frac{\partial h}{\partial x'_\alpha}$ and $\ell_\alpha = \frac{\partial \ell}{\partial x'_\alpha}$, are easily derived. In addition, the Jacobian of the transformation is given by

$$J = \frac{\partial(x, y, z)}{\partial(x',y',z')} = \ell \tag{17}$$

It will be assumed, for the moment, that stress free, open circuit boundary conditions hold on both upper and lower surfaces; putting the generalised stress equal to zero on these surfaces eliminates the surface integral in (7) from all regions except the boundary edge. Using Eqns (15), (16) and (17), the Lagrangian L of Eqn (7) may be transformed to the new coordinates; from now on the ',s will be omitted, and all indices will refer to the transformed coordinates.

$$L = \int_{t_1}^{t_2}\int_V \Big[\tfrac{1}{2}\rho_{ij}\dot{u}_i\dot{u}_j - \frac{1}{2\ell^2}c_{i3j3}u_{i,3}u_{j,3} - \frac{1}{\ell}c_{i3j\alpha}u_{i,3}x$$

$$(u_{j,\alpha} - \frac{1}{\ell}(h_\alpha+z\ell_\alpha)u_{j,3}) - \tfrac{1}{2}c_{i\alpha j\beta}(u_{i,\alpha} - \frac{1}{\ell}(h_\alpha+z\ell_\alpha)u_{i,3})x$$

$$(u_{j,\beta} - \frac{1}{\ell}(h_\beta+z\ell_\beta)u_{j,3})\Big]\ell\, d^2x_\alpha\, dzdt + \int_{t_1}^{t_2}\int_E T^o_{i\alpha}u_i\ell\, ds_\alpha dzdt \tag{18}$$

In the final integral E refers to the edge surface, and $ds_\alpha$ is an element of length along the boundary in the x-y plane; $ds_\alpha$ is directed along the outward normal as before. The next step in the procedure is to expand the displacement field in terms of a complete set of functions in z, and to perform the integration over z explicitly.

The expansion takes the form [2,3]

$$u_i = \sum_n A_n(x,y,t)u_i^n(z), \tag{19}$$

where the $u_i^n$ are some set of basis functions. In this particular case, these functions will be the solutions for an infinite flat plate with stress free, open circuit boundary conditions combined with the auxiliary electrostatic modes; the basis functions themselves do not satisfy the boundary conditions of the contoured plate, but this is not necessary as long as they form a mathematically complete set. Specifically, if the three 4-D vectors $b_{ij}$, j=1,2,3, contain the unit polarisation vectors and quasi-static potentials of the three possible z propagating plane waves, and if $b_{i4}$=0, i=1,2,3, and $b_{44}$=1, then the $u_i^n$ comprise the four sets of functions, j=1,2,3,4,

$$b_{ij}\cos\Big[\frac{r\pi}{2}(z+1)\Big],\ r=0,1,2,\ldots \tag{20}$$

Substituting Eqn (19) into Eqn (18) and integrating over z the two dimensional Lagrangian is obtained:

$$L = \int_{t_1}^{t_2}\int_A \frac{1}{2}\sum_n \Big[\rho_n\ell\dot{A}_n^2 - \frac{\nu_n}{\ell}A_n^2\Big] - \sum_{mn}\Big[(a^\beta_{nm} - h_\alpha e^{\alpha\beta}_{nm} - \ell_\alpha f^{\alpha\beta}_{nm})\frac{\partial A_m}{\partial x_\beta}A_n$$

$$+ \frac{1}{2}\ell b^{\alpha\beta}_{mn}\frac{\partial A_m}{\partial x_\alpha}\frac{\partial A_n}{\partial x_\beta} - \Big(\frac{h_\alpha}{\ell}c^\alpha_{mn} + \frac{\ell_\alpha}{\ell}d^\alpha_{mn} - \frac{h_\alpha h_\beta}{2\ell}g^{\alpha\beta}_{mn}$$

$$- \frac{1}{2}(h_\alpha\ell_\beta + h_\beta\ell_\alpha)q^{\alpha\beta}_{mn} - \frac{1}{2}\frac{\ell_\alpha\ell_\beta}{\ell}r^{\alpha\beta}_{mn}\Big)A_mA_n\Big]dxdydt$$

$$+ \int_{t_1}^{t_2}\int_E \sum_n A_n T^n_{i\alpha}ds_\alpha dt,$$

$$\text{where } T^n_{i\alpha} = \int_{-1}^{1}\ell T^o_{i\alpha}u_i^n dz, \tag{21}$$

and the integral is over the area A of the resonator in the x-y plane. The constant coefficients are easily calculated and are given by:

$$\rho_n = \int_{-1}^{1}\rho_{ij}\,u_j^n\,u_i^n\,dz$$

$$\nu_n = \int_{-1}^{1}c_{i3j3}\,u_{j,3}^n\,u_{i,3}^n\,dz$$

$$a^\alpha_{nm} = \int_{-1}^{1}c_{i3j\alpha}\,u_j^m\,u_{i,3}^n\,dz$$

$$b^{\alpha\beta}_{nm} = \int_{-1}^{1}c_{i\alpha j\beta}\,u_j^m\,u_i^n\,dz$$

$$c^\alpha_{nm} = \int_{-1}^{1}c_{i3j\alpha}\,u_{j,3}^m\,u_{i,3}^n\,dz$$

$$d^\alpha_{nm} = \int_{-1}^{1}c_{i3j\alpha}\,zu_{j,3}^m\,u_{i,3}^n\,dz \tag{22}$$

$$e^{\alpha\beta}_{nm} = \int_{-1}^{1}c_{i\alpha j\beta}\,u_j^m\,u_{i,3}^n\,dz$$

$$f^{\alpha\beta}_{nm} = \int_{-1}^{1}c_{i\alpha j\beta}\,zu_j^m\,u_{i,3}^n\,dz$$

$$g^{\alpha\beta}_{nm} = \int_{-1}^{1}c_{i\alpha j\beta}\,u_{j,3}^m\,u_{i,3}^n\,dz$$

$$q^{\alpha\beta}_{nm} = \int_{-1}^{1}c_{i\alpha j\beta}\,zu_{j,3}^m\,u_{i,3}^n\,dz$$

$$r^{\alpha\beta}_{nm} = \int_{-1}^{1}c_{i\alpha j\beta}\,z^2\,u_{j,3}^m\,u_{i,3}^n\,dz.$$

Having deduced the two dimensional Lagrangian, the equations of motion are easily deduced using the standard Euler-Lagrange equations of motion.

$$\rho_n \frac{\partial^2 A_n}{\partial t^2} + \frac{\nu_n}{\ell^2} A_n + \frac{1}{\ell} \sum_m (a^{\alpha}_{nm} - a^{\alpha}_{mn}) \frac{\partial A_m}{\partial x_\alpha} - \sum_m b^{\alpha\beta}_{nm} \frac{\partial^2 A_m}{\partial x_\alpha \partial x_\beta}$$

$$= \sum_m \left[ \frac{h_\alpha}{\ell^2} (c^{\alpha}_{nm} + c^{\alpha}_{mn}) + \frac{\ell_\alpha}{\ell^2} (d^{\alpha}_{nm} + d^{\alpha}_{mn}) - \frac{h_\alpha h_\beta}{\ell^2} g^{\alpha\beta}_{nm} - \right.$$

$$\left. \frac{(\ell_\alpha h_\beta + \ell_\beta h_\alpha)}{\ell^2} q^{\alpha\beta}_{nm} - \frac{\ell_\alpha \ell_\beta}{\ell^2} r^{\alpha\beta}_{nm} - \frac{h_{\alpha\beta}}{\ell} e^{\alpha\beta}_{mn} - \frac{\ell_{\alpha\beta}}{\ell} f^{\alpha\beta}_{mn} \right] A_m$$

$$+ \sum_m \left[ \frac{h_\alpha}{\ell} (e^{\alpha\beta}_{nm} - e^{\alpha\beta}_{mn}) + \frac{\ell_\alpha}{\ell} (f^{\alpha\beta}_{nm} - f^{\alpha\beta}_{mn}) + \frac{\ell_\alpha}{\ell} b^{\alpha\beta}_{nm} \right] \frac{\partial A_m}{\partial x_\beta} \qquad (23)$$

Although the quations of motion appear awesomely complex, the situation is not quite as bad as it appears, The right hand side of Eqn (23) contains only terms proportional to the thickness gradients, and these are the terms which are neglected in a simple treatment; the left hand side is formally identical to the flat plate equations, but the plate thickness parameter is now shown explicitly. Eqn (23) is precisely equivalent to the original three dimensional equations, and its value lies in the fact that it allows approximate equations of arbitrary complexity to be constructed easily and systematically. It is well known from the simple theory that even a single mode approximation may give quite good results, and the obvious procedure is therefore to select those mode amplitudes, $A_n$, which are thought to be significant, and to neglect all other amplitudes and their associated coupling terms. However, it is known that, even for the case of a flat plate, this approximation does not give particularly good results, but that by modifying the constant $b^{\alpha\beta}_{nm}$ the situation is greatly improved [3].

If the chosen mode indices form a set $Q=\{n_1,n_2,n_3\ldots\}$ and if Q is partitioned into subsets $R_1, R_2$ etc, the members of each subset having a common value of $\nu_n$, then the modified $b^{\alpha\beta}_{nm}$ coefficients are defined by

$$s^{\alpha\beta}_{nm} = b^{\alpha\beta}_{nm} - \sum_{\ell \notin Q} \frac{(a^{\alpha}_{n\ell} - a^{\alpha}_{\ell n})(a^{\beta}_{m\ell} - a^{\beta}_{\ell m})}{\nu_\ell - \rho_\ell \nu_n / \rho_n} \qquad (24)$$

if $m,n \in R_1$

$= b^{\alpha\beta}_{nm}$ otherwise.

It will be noted that the correction term involves a summation over the modes not included in the chosen set Q. With this modification the approximate equations are obtained from (21) or (23) by simply neglecting all mode amplitudes and coupling terms involving mode indices outside the chosen set. Such approximate equations have the following properties:

1 All coupling terms due to the plate thickness variation are included.

2 For the special case of a flat plate, the cutoff frequencies of the chosen modes are correctly predicted, and their dispersion curves are also rigorously correct up to terms of order $k^2$.

3 The equations tend to the exact form as the number of mode amplitudes is increased.

The above equations were developed on the assumption of stress free, open circuit boundary conditions, which are not quite correct in the electroded regions of the resonator. This assumption does not play an essential role in the theory, but is a simplifying factor. In this paper the effects of the electrodes are not treated rigorously, and are allowed for by modifications to the $\rho_n$ and $\nu_n$ parameters.

## 3. THE FINITE ELEMENT METHOD

Variational, or Lagrangian, methods offer a powerful technique for deducing approximate solutions to differential equations. Although it is rarely possible to find the true minimum of the Lagrangian, which would, of course, provide an exact solution of the associated equations, it is often possible to find an approximate minimum over a restricted set of functions. In a previous paper [2], a classical variational approach was used, where the solution was approximated by a linear combination of basis functions, each of which was continuous over the whole domain; this is a simple and often extremely effective technique, but it is limited in the way in which it can represent complex domains and boundaries, and, for this reason, the more flexible finite element method forms the basis of most general computational procedures.

The finite element method is a very well developed technique, and is comprehensively described in many standard texts [5]. The basic principle is to divide the domain of the problem into a certain number of sub-regions, or elements; each of these elements has a number of special points, or nodes, associated with it, and the values of the unknown functions at these node points form the variables for the minimisation problem. Within each element the function is approximated by an interpolation of its values at the node points. The overall solution may therefore be viewed as a piecewise continuous approximation to the true solution. For the elements used in this paper, the approximate function itself is continuous between elements, but the first derivative is not; indeed, the use of finite elements with first derivative continuity is very rare.

To be more specific, Figure 2 shows simple two dimensional finite element meshs for a circular resonator with a square electrode and rectangular lead-offs; it can be seen that the elements are simple quadrilaterals and triangles with the node points located on the boundaries. It is possible to treat each element in terms of a single standard element which is then mapped onto the final geometry. Figures 3 and 4 show the two types of standard element used in this paper, a first and a second order quadrilateral; the triangular elements are treated as degenerate cases of these basic elements. For the first order element we may define a set of shape functions, $N_i$, in terms of the local coordinates $(\xi,\eta)$

$$N_1 = \tfrac{1}{4}(1-\xi)(1-\eta)$$

$$N_2 = \tfrac{1}{4}(1+\xi)(1-\eta) \qquad (25)$$

$$N_3 = \tfrac{1}{4}(1+\xi)(1+\eta)$$

$$N_4 = \tfrac{1}{4}(1-\xi)(1+\eta),$$

there being one shape function for each node point, similarly for the second order element,

$$N_1 = -\tfrac{1}{4}(1-\xi)(1-\eta)(1+\xi+\eta)$$

$$N_2 = \tfrac{1}{2}(1-\xi^2)(1-\eta)$$

$$N_3 = -\tfrac{1}{4}(1+\xi)(1-\eta)(1-\xi+\eta)$$

$$N_4 = \tfrac{1}{2}(1-\eta^2)(1+\xi) \qquad (26)$$

$$N_5 = -\tfrac{1}{4}(1+\xi)(1+\eta)(1-\xi-\eta)$$

$$N_6 = \tfrac{1}{2}(1-\xi^2)(1+\eta)$$

$$N_7 = -\tfrac{1}{4}(1-\xi)(1+\eta)(1+\xi-\eta)$$

$$N_8 = \tfrac{1}{2}(1-\eta^2)(1-\xi)$$

The shape functions form basis functions for a Lagrange interpolation over the element; it is easily seen that for each node point, $(\xi_j,\eta_j)$,

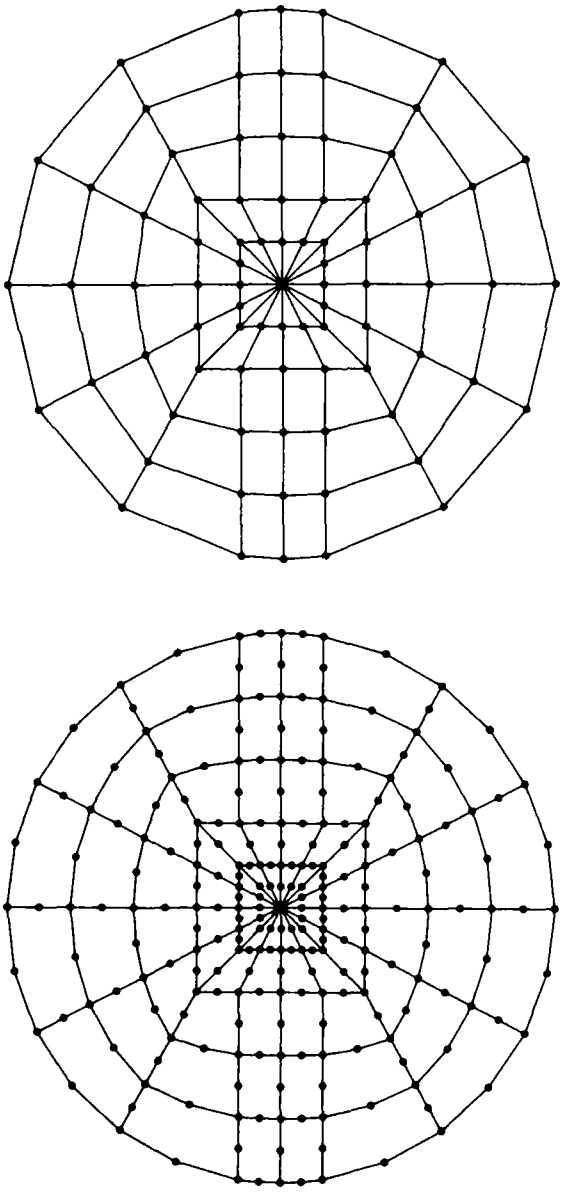


Figure 2: First and second order element meshes

$$N_i(\xi_j,\eta_j) = \delta_{ij}, \qquad (27)$$

and, given values $(A_i)$ of a particular function at the nodes, the function

$$A(\xi,\eta) = \sum_i A_i N_i \qquad (28)$$

forms a polynominal interpolation over the element. It is easily seen that each edge interpolation is a linear or a quadratic function, entirely defined by the nodes on that edge; this ensures function continuity between elements. The $(\xi,\eta)$ coordinates are the local coordinates for the standard prototype element, and a rule must be defined for mapping the element into the final structure. If the actual node coordinates of a chosen element are $(x_i,y_i)$ then a suitable coordinate transformation from the $(\xi,\eta)$ to the $(x,y)$ domain may be defined by

$$x = \sum_i x_i N_i(\xi,\eta)$$

$$y = \sum_i y_i N_i(\xi,\eta) \qquad (29)$$

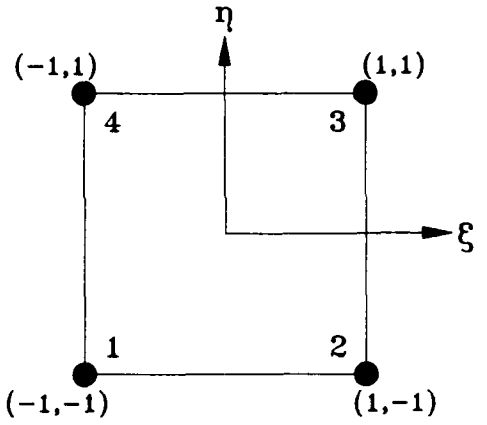



Figure 3: First order (linear) element

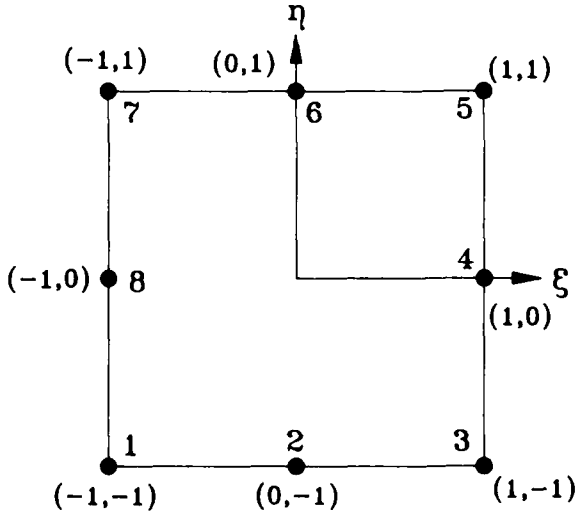



Figure 4: Second order (quadratic) element

This transformation uses the same shape functions as the function interpolation itself; this is not the only possible choice, but it is usually a convenient one; elements generated in this way are usually referred to as isoparametric.

It is fairly obvious that the accuracy of approximation can be improved by using a denser FE mesh, or by using higher order elements; the precise choice is ultimately a matter of engineering judgement, but, up to a point, the use of higher order elements usually pays off.

Having defined the element types and the FE mesh, the transformation of the Lagrangian is a straightforward matter. If the value of the m'th mode amplitude at the i'th node of the structure is $A_m^i$, then the interpolation (28) and coordinate transformation (29) define the approximate form of the mode amplitude function within each element. The integral in the Lagrangian may then be integrated element by element, to provide a formula

$$L = \sum_{nmij} \tfrac{1}{2}\left[ P(m,i,n,j)\dot{A}_m^i\dot{A}_n^j - C(m,i,n,j)A_m^iA_n^j\right], \qquad (30)$$

where stress free boundary conditions on the edges have been assumed. The integral within each element is performed in $(\xi,\eta)$ space; this is far too complex for analytic integration and numerical quadrature is used instead. For the first and second order elements the quadrature is based upon two and three point Gauss rules respectively. There is some considerable art in choosing appropriate elements and corresponding integration rules, but, as yet, this matter has not been considered in any detail.

The problem is therefore reduced to a matrix problem involving the mode amplitudes at the node points; it should be noted that these amplitudes are still functions of t. From Eqn (30) the equivalent eigenvalue problem for deducing the natural frequencies is easily obtained,

$$\sum_{nj} \left[\omega^2 P(m,i,n,j) - C(m,i,n,j)\right]A_n^j = 0. \qquad (31)$$

Even with the simple mesh shown in Figure 2, the number of nodes is considerable, and even though P and C are symmetric

$$P(m,i,n,j) = P(n,j,m,i)$$

and

$$C(m,i,n,j) = C(n,j,m,i), \qquad (32)$$

the eigenvalue problem is still difficult.

The effort needed to solve Eqn (31) can be reduced by a suitable node numbering scheme. It can be seen that P and C will only have a non-zero value for an i,j element if the indices i and j occur together in a particular element. The aim is therefore to minimise the range of node numbers within a given element, as this reduces the bandwidth of the matrices. For the type of structure shown in Figure 2 the node numbers proceed in a roughly spiral pattern, and the matrix bandwidth is proportional to the number of nodes on one ring of the mesh.

Having reduced the problem to a generalised band symmetric eigenvalue problem, the solution may be found using standard eigenanalysis software, such as that provided in the NAG Library [6]. Effects such as the damping caused by the mounting points may then be analysed as a perturbation on the undamped solution.

## 4. RESULTS AND DISCUSSIONS

Although the new contoured resonator theory is well established, it has only recently been given a practical implementation, and the results obtained so far are of a fairly negative form.

Considering the equations of motion (23), and the coefficients (22), certain general observations can be made. In the use of a plano-convex blank, $\ell_2=0$ and $h=\ell$, but for a bi-convex blank $\ell_1=\ell_2$ and $h=0$; the number of coupling terms is therefore considerably less for the bi-convex case, and such a geometry is likely to be theoretically superior. For most resonator designs, the thickness gradients are likely to be small, and terms involving products and higher order derivatives of the gradients will usually be negligible. Many of the coefficients in (22) are identically zero for basis functions with particular symmetries in the z direction.

For m and n both even or both odd,

$$a_{nm}^{\alpha} = d_{nm}^{\alpha} = e_{nm}^{\alpha\beta} = q_{nm}^{\alpha\beta} = 0 \quad .$$

For m and n one even and one odd,

$$b_{nm}^{\alpha\beta} = c_{nm}^{\alpha} = f_{nm}^{\alpha\beta} = g_{nm}^{\alpha\beta} = r_{nm}^{\alpha\beta} = 0.$$

Furthermore, for the chosen set of basis functions, the coefficients $b_{nm}^{\alpha\beta}$, $c_{nm}^{\alpha}$ and $g_{nm}^{\alpha\beta}$ are identically zero unless n and m refer to modes with the same overtone number. In the absence of the thickness gradient terms, only the $a_{nm}^{\alpha}$ coefficients can couple modes of different overtones, and then only if one is even and one is odd. The new coupling terms provide a greater variety of possibilities and allow couplings between pairs of even and odd modes.

It is also possible that certain terms may be identically zero due to some special choice of crystallographic orientation; an obvious case is that of the rotated y-cuts of quartz. However, a comparison of the coefficients for AT-cut and SC-cut quartz for the case of coupling between the fundamental thickness shear mode and the flexural mode showed no differences of obvious significance. It is possible that noticeable differences exist with respect to other modes, but this has yet to be investigated.

To determine the effectiveness of the finite element analysis approach a comparative test was performed using the classical variational method and the finite element method with a range of different meshes. In addition, the FEM results were compared with and without the inclusion of the thickness gradient terms. For the test, a hypothetical SC-cut quartz design was chosen with the following parameters:

| | |
|---|---|
| Frequency | 5 MHz |
| Overtone | 5th |
| Design | Plano-convex |
| Radius of curvature | 50 mm |
| Diameter of blank | 13 mm |
| Diameter of electrode | 8 mm |
| Width of lead off | 4 mm |
| Plating material | Gold |
| Plating thickness | 0.1 μm |

| Mode | f(MHz) classical method | f(MHz) FEM Mesh 1 | f(MHz) FEM Mesh 2 | f(MHz) FEM Mesh 3 | f(MHz) FEM Mesh 1 No grad. terms | f(MHz) FEM Mesh 2 No grad. terms | f(MHz) FEM Mesh 3 No grad. terms |
|---|---|---|---|---|---|---|---|
| (0,0,1) | 4.999997 | 4.999297 | 4.997374 | 4.997858 | 5.004435 | 5.002536 | 5.003015 |
| (1,0,1) | 5.083853 | 5.089299 | 5.083922 | 5.085226 | 5.093865 | 5.088700 | 5.089957 |
| (0,1,1) | 5.089123 | 5.095399 | 5.089730 | 5.091110 | 5.100152 | 5.094476 | 5.095864 |
| (2,0,1) | 5.166364 | 5.183729 | 5.170796 | 5.173841 | 5.187933 | 5.175197 | 5.178182 |
| (1,1,1) | 5.171546 | 5.186857 | 5.176010 | 5.178459 | 5.190549 | 5.180367 | 5.182610 |
| (0,2,1) | 5.176729 | 5.194746 | 5.182500 | 5.185436 | 5.198900 | 5.186838 | 5.189735 |

TABLE 1: Comparison of analysis techniques

The finite element analysis was performed with three different meshes; these were similar to those shown in Figure 2, but with different numbers of sectors and annular rings. The first mesh used 128 first order elements, comprising 8 rings and 16 sectors; the second mesh was identical, but used second order elements instead. The third mesh contained 512 first order elements in a 16 ring, 32 sector pattern; in effect, each second order element in the second mesh was replaced by four first order elements. Table 1 shows a comparison of the first six mode frequencies of the resonator computed using the classical method and the single mode FE method with the three chosen meshes; the modes are classified according to the number of nodes in the (x,y,z) directions, e.g. (0,0,1) for the fundamental. In addition, Table 1 gives the results of the FE analysis when all thickness gradient terms are omitted.

In general, with a variational method, the lower the estimated mode frequencies the better the approximation. On this basis, mesh 2 is clearly the most satisfactory, and better, as well as computationally cheaper, results are obtained using a single second order element rather than four first order elements. The agreement with the classical method is generally very good, and no significant differences in mode frequency are apparent. The thickness gradient terms produce a fairly constant shift in frequency for all the anharmonic modes. It might be supposed that closer agreement would exist between the classical method and the FE method neglecting thickness gradients; however, the classical method uses a parabolic approximation to the blank contour, whereas the FE approach uses the exact spherical shape.

It can be seen that for a single mode analysis the rigorous theory and FE analysis have little advantage over the earlier classical variational method. The new system does offer advantages for analysing mode coupling problems, and some analyses of this sort have been performed, but the results have not been evaluated in any detail. The major limitation is the amount of computer memory and processing time required for elaborate calculations, and the choice of improved eigenvalue programs is currently being considered.

## 5. CONCLUSIONS

A rigorous technique for constructing approximate systems of equations for contoured resonator analysis has been developed, and a computer implementation has been constructed using the finite element method.

Initial results suggest that for a single mode approximation the new method has few advantages over earlier procedures based on a classical variational approach. The finite element method does however provide flexibility in analysing multimode problems, and it is possible that the additional coupling terms included in the new theory will have significant effects in some of these cases.

## ACKNOWLEDGEMENTS

This work was supported by the UK Ministry of Defence.

## REFERENCES

1 H F Tiersten and R C Smythe, 'An analysis of contoured crystal resonators operating in overtones of coupled thickness shear and thickness twist', J. Acoust. Soc. Am., 65, 6, pp 1455-1460, 1979

2 R C Peach, "A variational method for the design of trapped energy resonators", Proc. 39th Annual Freq. Control Symposium., pp. 392-399, 1985.

3 R C Peach, "A normal mode expansion for piezoelectric plates and certain of its applications", To be published in IEEE Trans. UFFC, Sept 1988.

4 H Goldstein, "Classical mechanics", 2nd Ed., Addison-Wesley, 1980.

5 O C Zienkiewicz, "The finite element method", 3rd Ed. McGraw-Hill, London, 1977.

6 NAG Fortran Library Reference Manual (Mark 12), Numerical Algorithms Group Limited, Oxford, England.

42nd Annual Frequency Control Symposium - 1988

# Vibration Analysis of Coupled Flexural Torsional Mode Tuning Fork Type Quartz Crystal Resonator

Hirofumi KAWASHIMA
SEIKO Electronic Components Ltd.
Tochigi-shi Tochigi-ken, Japan 328

## SUMMARY

It is well known that a flexural torsional quartz crystal resonator which couples a torsional mode vibration to a flexural mode vibration, has a cubic curve in frequency temperature characteristics as well as an AT cut quartz crystal resonator. This paper describes the frequency temperature characteristics of a tuning fork-type quartz crystal resonator which couples the torsional mode vibration to the flexural mode vibration and the estimated results of the various constants for the resonator manufactured for trial experiments. First, a secular equation coupled the torsional mode vibration to the flexural mode vibration is calculated. From the secular equation, frequency-temperature characteristics for the flexural-torsional mode tuning fork-type resonator have been analyzed theoretically, specifically concerning a relationship of the resonant frequency, the frequency temperature coefficients vs. thickness, a normaliged frequency and a cut angle. Next, by manufacturing tuning fork type flexural-torsional quartz crystal resonators, the calculated values are compared with the experimental values, at the same time, the various constants of the resonators manufactured for trial experiments are estimated. As a result, it is confirmed that tuning fork type flexural torsional quartz crystal resonators with excellent frequency-temperature characteristics, low crystal impedance, and a high quality factor are realized.

## 1. Introduction

A tuning fork type flexural quartz crystal resonator is used as a resonator for a wristwatch, because of noteworthy features of easy miniaturization, low frequency, high shock resistance, and so on. However, frequency-temperature characteristics of the tuinig fork type flexural quartz crystal resonator which is used widely, have approximately a convex parabolic curve with a temperature coefficient of about -0.04 ppm/°c²; because it is comparatively large, it was impossible to obtain a quartz crystal resonator with a small frequency change over a wide temperature range. Furthermore, though GT cut quartz crystal resonators excellent in frequency temperature characteristics were reported [1-3], as the resonator for [1] has large consumption power, because of high frequency, and as the other is inferior in miniaturization and shock resistance, it is very difficult to utilize the GT cut quartz crystal resonators.

Recently, however, it was reported [4] that a coupling between a flexural mode vibration and a torsional mode vibration of a tuning fork type gives an improvement in frequency-temperature characteristics of the flexural mode vibration. In addition, an analysis of these frequency-temperature characteristics has been made by the finite element method. However, the method has a drawback: it requires many hours of analysis. Therefore, a theoretical analysis method not requiring many hours is extremely necessary in development of a resonator. First, in order to clarify theoretically frequency-temperature characteristics of a flexural-torsional quartz crystal resonator, a secular equation of the flexural-torsional quartz crystal resonator is calculated from partially differential equations including a flexural-torsional coupling, by an elastic theory. It is solved by Galerkin's method. From the obtained frequency equation, a relationship of resonant frequency, frequency-temperature coefficients vs. thickness, a cut angle and a normalized frequency is made clear.

Next, quartz crystal resonators are actually manufactured and the calculated values are compared with the experimental values. In addition, electrical equivalent circuit constants for tuning fork type flexural-torsional quartz crystal resonators manufactured for trial experiments are shown.

## 2. Derivation of the Frequency Equation

### 2.1 Deflection curve

Figure 1 shows a tuning fork-type flexural-torsional quartz crystal resonator and its coordinate system. Since the tuning fork type quartz crystal

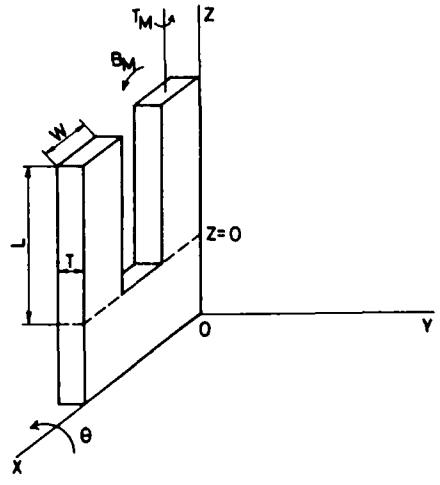



**Fig. 1 Coordinate system of tuning fork type quartz resonator.**

CH2588-2/88/0000-045 $1.00 C 1988 IEEE

resonator can be considered as two cantilevers, a vibration analysis for the resonator is made as this cantilever. As is obvious from Fig. 1, two vibration modes are, respectively, a flexural mode with displacement in the X-axis direction, and a torsional mode with rotation about the Z-axis. Specifically, a deflection curve which is caused by a coupling of the flexural vibration and the torsional vibration, is expressed as follows from [5]:

$$u = -\frac{1}{2I_Y}\left(B_M S'_{33} + T_M \frac{S'_{34}}{2}\right) Z^2 \tag{1}$$

$$\frac{\partial \phi}{\partial Z} = \frac{T_M}{C_t} + \frac{B_M}{2I_Y} S'_{34} \tag{2}$$

where Iy is the moment of inertia; u is the displacement in the X-axis direction, BM is the bending moment; TM is the torsional moment, CT is the torsional riddidity; $\phi$ is the torsional displacement angle; S33', S44' are elastic compliance constants, with a function of cut angle $\theta$.

The last term of Eqs. (1) and (2) show an elastic coupling. Differentiating Eq. (1) with Z twice, Eqs. (1) and (2) are as follows:

$$\left.\begin{aligned} -I_Y E_{ZZ} \frac{\partial^2 u}{\partial Z^2} &= B_M + \frac{E_{ZZ} I_Y}{C_{mt}} T_M \\ C_t \frac{\partial \phi}{\partial Z} &= T_M + \frac{C_t}{C_{mt}} B_M \end{aligned}\right\} \tag{3}$$

where EZZ = 1/S33', Cmt = 2Iy/S34'.

Next, solving Eqs. (3) for BM, TM:

$$\left.\begin{aligned} B_M &= -\frac{I_Y E_{ZZ}}{\eta} \frac{\partial^2 u}{\partial Z^2} - \frac{C_t E_{ZZ} I_Y}{\eta C_{mt}} \frac{\partial \phi}{\partial Z} \\ T_M &= \frac{C_t}{\eta} \frac{\partial \phi}{\partial Z} + \frac{C_t I_Y E_{ZZ}}{\eta C_{mt}} \frac{\partial^2 u}{\partial Z^2} \end{aligned}\right\} \tag{4}$$

where $\eta = 1 - \dfrac{E_{ZZ} I_Y C_t}{C_{mt}^2}$

## 2.2 Equation of motion

Equations of motion for a harmonic flexural vibration and a harmonic torsional vibration are expressed as follows, when taking sectional area A, density $\rho$, polar moment of inertia Io and time t for a resonator:

$$\left.\begin{aligned} \frac{\partial^2 B_M}{\partial Z^2} - \rho A \frac{\partial^2 u}{\partial t^2} &= 0 \\ \frac{\partial T_M}{\partial Z} - \rho I_0 \frac{\partial^2 \phi}{\partial t^2} &= 0 \end{aligned}\right\} \tag{5}$$

In addition, u and $\phi$ are expressed as a function of time as follows:

$$\left.\begin{aligned} u &= U(Z)\exp(i\omega t) \\ \phi &= \Phi(Z)\exp(i\omega t) \end{aligned}\right\} \tag{6}$$

where $i = \sqrt{-1}$, $\omega$ denotes angular frequency.

Further, substituting Eqs. (4) and (6) into Eqs. (5), and differentiating the second equation of Eqs. (5) with Z once,

$$\left.\begin{aligned} -\frac{E_{ZZ} I_Y}{\eta} \frac{d^4 U}{dZ^4} - \frac{C_t}{\eta} \frac{E_{ZZ} I_Y}{C_{mt}} \frac{d^3 \Phi}{dZ^3} + \rho A \omega^2 U &= 0 \\ \frac{C_t}{\eta} \frac{d^3 \Phi}{dZ^3} + \frac{C_t E_{ZZ} I_Y}{\eta C_{mt}} \frac{d^4 U}{dZ^4} + \rho I_0 \omega^2 \frac{d\Phi}{dZ} &= 0 \end{aligned}\right\} \tag{7}$$

## 2.3 Boundary conditions

A cantilever of a flexural-torsional mode has the boundary conditions of "clamped" at one end and "free" at the other end. That is, from [6], for the flexural mode it has the boudary conditions that displacement and rotation become zero at the clamped end (Z = 0), and displacement and rotation are free. In addition, force and moment do not work at the free end (Z = L), while, for the torsional mode, it has the boudary conditions that the displacement angle becomes zero at the clamped end (Z = 0) and the moment becomes zero at the free end (Z = L). Therefore, the boundary conditions at the clamped end and at the free end are expressed as follows:

$$\left.\begin{aligned} Z = 0 \;;\; U = \frac{dU}{dZ} = \Phi &= 0 \\ Z = L \;;\; \frac{d^2 U}{dZ^2} = \frac{d^3 U}{dZ^3} = \frac{d\Phi}{dZ} &= 0 \end{aligned}\right\} \tag{8}$$

Furthermore, the function U, $\phi$ satisfying Eq. (8) are obtained from [6]

$$\left.\begin{aligned} U &= A_1\{(\cosh\lambda L + \cos\lambda L)(\cosh\lambda Z - \cos\lambda Z) \\ &\quad -(\sinh\lambda L - \sin\lambda L)(\sinh\lambda Z - \sin\lambda Z)\} \\ \Phi &= B_1 \sin\nu Z \end{aligned}\right\} \tag{9}$$

Thus, the values of $\lambda$ and $\nu$ are calculated by the eigenequation obtained from the boundary conditions, and $\nu = 1.571/L$, $\lambda = 1.875/L$ for a fundamental vibration.

## 2.4 Galerkin's method

Since it is not easy to solve the simultaneous differential equations of Eq. (7), they are written by means of an approximate solution. There are many ways to obtain approximate solutions. In this case, since the functions U, $\phi$ satisfying the boundary conditions are obtained easily, the simultaneous differential equations are solved approximately by Galerkin's method available for calculating an angular frequency. First, substituting Eq. (9) into Eq. (7), the following equations are obtained:

$$\left.\begin{aligned}&\left(\omega^2-\frac{1}{\eta}\omega_F^2\right)U+\frac{I_Y E_{ZZ} C_t}{\eta C_{mt}\rho A}\nu^3 Y_1=0\\&\frac{I_Y E_{ZZ} C_t}{\eta C_{mt}\rho I_0}\frac{\lambda^4}{\nu}U+\left(\omega^2-\frac{1}{\eta}\omega_T^2\right)Y_1=0\end{aligned}\right\}\quad(10)$$

where $\omega F$ and $\omega T$ are, respectively, angular frequencies for a flexural mode and a torsional mode; in addition,

$$Y_1=B_1\cos\nu Z$$

$$\omega_F^2=\frac{E_{ZZ}I_Y}{\rho A}\lambda^4,\quad \omega_T^2=\frac{C_t}{\rho I_0}\nu^2$$

Next, after multiplying the first equation of Eq. (10) by the eigenfunction U satisfying the boundary conditions, and the second equation by Y1 as well, integration is made along the bar. As a result, the following equations are obtained:

$$\left.\begin{aligned}&\left(\omega^2-\frac{1}{\eta}\omega_F^2\right)\int_0^L U^2dZ+\frac{I_Y E_{ZZ}C_t}{\eta C_{mt}\rho A}\nu^3\int_0^L Y_1U\,dZ=0\\&\frac{I_Y E_{ZZ}C_t}{\eta C_{mt}\rho I_0}\frac{\lambda^4}{\nu}\int_0^L UY_1dZ+\left(\omega^2-\frac{1}{\eta}\omega_T^2\right)\int_0^L Y_1^2dZ\\&=0\end{aligned}\right\}\quad(11)$$

Furthermore, the integration terms of Eq. (11) are as follows:

$$\left.\begin{aligned}&\int_0^L U^2dZ=a_1A_1^2L\\&\int_0^L Y_1UdZ=a_2A_1B_1L\\&\int_0^L Y_1^2dZ=a_3B_1^2L\end{aligned}\right\}\quad(12)$$

where a1, a2 and a3 are different in the order of the vibration modes. They have the following values.

Both of flexure and torsion: Fundamental vibration:

$$\left.\begin{aligned}&a_1=9.23\\&a_2=0.933\\&a_3=0.5\end{aligned}\right\}\quad(13)$$

Flexure: Second overtone, Torsion: Fundamental vibration:

$$\left.\begin{aligned}&a_1=2905\\&a_2=24.9\\&a_3=0.5\end{aligned}\right\}\quad(14)$$

In addition, substituting Eq. (12) into Eq. (11), the following equations are obtained:

$$\left.\begin{aligned}&\left(\omega^2-\frac{1}{\eta}\omega_F^2\right)a_1A_1+\frac{C_tI_YE_{ZZ}}{\eta C_{mt}\rho A}\nu^3a_2B_1=0\\&\frac{I_YE_{ZZ}C_t}{\eta C_{mt}\rho I_0}\frac{\lambda^4}{\nu}a_2A_1+\left(\omega^2-\frac{1}{\eta}\omega_T^2\right)a_3B_1=0\end{aligned}\right\}\quad(15)$$

## 2.5 Secular equation

Equation (15) shows simultaneous equations with respect to A1 and B1. Since A1 and B1 are not zero, the condition that there is a solution in Eq. (15), is as follows:

$$\begin{vmatrix}\left(\omega^2-\frac{1}{\eta}\omega_F^2\right)a_1 & \frac{C_tI_YE_{ZZ}}{\eta C_{mt}\rho A}\nu^3a_2\\ \frac{I_YE_{ZZ}C_t}{\eta C_{mt}\rho I_0}\frac{\lambda^4}{\nu}a_2 & \left(\omega^2-\frac{1}{\eta}\omega_T^2\right)a_3\end{vmatrix}=0\quad(16)$$

Equation (16) is a secular equation for an angular frequency $\omega$. Solving Eq. (16) which shows the fourth-order equation concerning $\omega$, the following relation is obtained:

$$\omega^2=\frac{1}{2\eta}\left\{\omega_F^2+\omega_T^2\pm\sqrt{(\omega_F^2-\omega_T^2)^2+4K_m^2\omega_F^2\omega_T^2}\right\}\quad(17)$$

where $\omega=2\pi f$, and f is frequency:

$$K_m^2=\frac{a_2^2}{a_1a_3}\frac{E_{ZZ}I_YC_t}{C_{mt}^2}$$

Equation (17) gives two angular frequencies. The sign K m is a coupling factor showing coupling intensity between the flexural mode and the torsional mode. As is obvious from Eq. (13) and (14), the coupling intensity varies with the vibration order. If there is no coupling between them, Km=0. Therefore, two angular frequencies show those for the independent flexural and torsional modes, respectively.

## 3. Calculation of Frequency-Temperature Coefficients

The frequencies obtained from Eq. (17) are a function of a dimension and elastic compliance, which are also a function of temperature. Therefore, by taking an arbitrary temperature T1, Taylor's series are expressed at the temperature 25℃ as follows:

$$f(T_1)=f(25)\{1+\alpha(T_1-25)+\beta(T_1-25)^2+\gamma(T_1-25)^3+\cdots\}\quad(18)$$

Where the first-, second- and third-order temperature coefficients $\alpha,\beta,\gamma$ are defined as follows:

$$\left.\begin{aligned}&\alpha=\left(\frac{1}{f}\frac{\partial f}{\partial T_1}\right)_{T_1=25℃}\\&\beta=\left(\frac{1}{2!}\frac{1}{f}\frac{\partial^2 f}{\partial T_1^2}\right)_{T_1=25℃}\\&\gamma=\left(\frac{1}{3!}\frac{1}{f}\frac{\partial^3 f}{\partial T_1^3}\right)_{T_1=25℃}\end{aligned}\right\}\quad(19)$$

For example, the first order temperature coefficients $\alpha$ is obtained as follows from the calculation of Eq. (19):

$$\alpha=\frac{1}{2}\left[-\frac{\dot\eta}{\eta}+\frac{1}{2\eta\omega^2}\left\{2\left(\omega_F^2\frac{\dot\omega_F}{\omega_F}+\omega_T^2\frac{\dot\omega_T}{\omega_T}\right)\pm\dot g\right\}\right]$$

where showing $\cdot \equiv \partial/\partial T1$, the first-order temperature coefficient $\alpha$ is calculated from differentiation of $\eta$, $\omega F$, $\omega T$ and g with respect to temperature (see Appendix). In addition, the constants of [7] are used in the calculation of temperature coefficients.

## 4. Results and Discussion

First, elastic compliance S34' dependent on a cut angle, which exercises great influence on a coupling between a flexural mode vibration and a torsional mode vibration is stated. It is shown that the flexural mode vibrations and the torsional mode vibration are fully coupled with each other, by illustrating frequency spectra at the cut angles $\theta$ where $S34' \neq 0$. Next, since frequency-temperature coefficients of a flexural mode tuning fork-type resonator coupled with the torsional mode vibration vary with the cut angle $\theta$ and a dimension of the resonator, a relatonship between the frequency temperature coefficients $\alpha$, $\beta$ and the cut angle $\theta$ is shown first. Furthermore, a relationship is shown between cut angles $\theta$ where $\alpha = 0$ and $\beta = 0$. Next, a relationship between $\alpha, \beta$ and the dimensions, particularly, thickness, is shown. In addition, since changing the cut angle $\theta$ or the dimension of the coupling resonator is equivalent to changing a frequency of the flexural mode vibration and / or the torsional mode vibration, by defining newly a normalized frequency $\delta f$, a relationship between $\delta f$ and $\alpha$, $\beta$, when taking the cut angle $\theta$ as a parameter, is made clear. Finally, the frequency temperature characteristics obtained from the calculation and the experiments are shown. A concrete description is given below.

### 4.1 Elastic compliance S34'

Figure 2 shows a relationship of elastic compliance, which relates to a flexural vibration and a torsional vibration, vs. a cut angle; S33' relates to the flexural vibration, S55' and S44' relate to the torsional vibration. In addition, S34' is a factor indicating coupling intensity between the flexural vibration and the torsional vibration, which varies with a cut angle and has positive and negative values.

Specifically, as is obvious from Fig. 2, a flexural vibration and a torsional vibration do not couple at the cut angles of about 3°, 70° and 130° at all, because S34' reaches zero at those cut angles and they degenerate.

Furthermore, it was reported in [4] that the flexural mode vibration and the torsional mode vibration couple at the cut angle of about 80°, and, consequently, $\alpha$ and $\beta$ for the flexural mode vibration almost reach zero. On the other hand, it is also understood from the calculated results that S34' has a positive value, thus, both mode vibrations couple. However, the coupling between both mode vibrations are very weak because S34' is much smaller than S33', S55'.

### 4.2 Frequency spectrum

Figure 3 shows a relationship of resonant frequencies fF, fT vs. thickness of a tuning fork type flexural torsional quartz crystal resonator, which performs fundamental vibrations for a flexural mode and a torsional

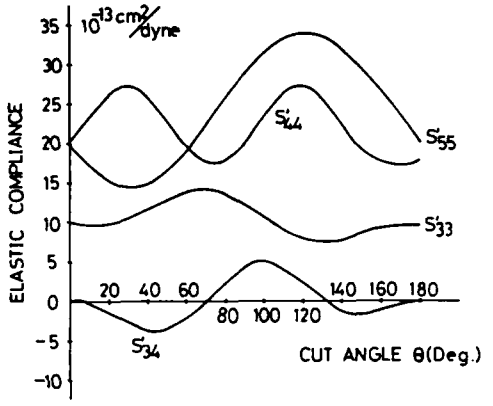



Fig. 2 Relationship between cut angles and elastic compliances.

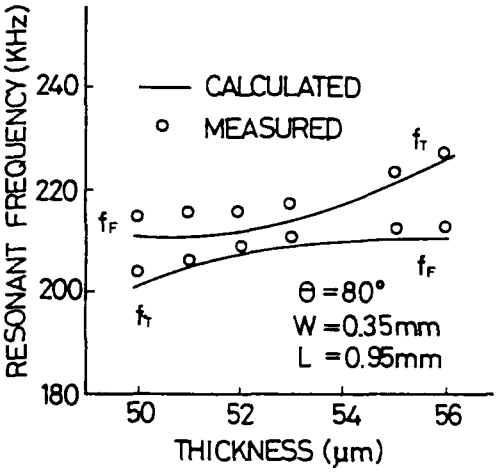



Fig. 3 Frequency spectrum vs. thickness.

mode, which has the cut angle $\theta = 80°$, $W = 0.35mm$, and $L = 0.95mm$. The solid lines are the calculated values, while the dashed lines are the experimental values. Since it is understood from the analysis of the finite element method that the nodal points of a tuning fork type quartz crystal resonator do not exist at the U shape portion, but portions near more mounting portion, effective length $L = 1.183mm (L = L + 2W/3)$ for flexure and effective length $L = 1.3mm (L = L + W)$ for torsion are used in the calculation. Resonant frequency fF calculated at $T = 50\ \mu m$ has a value of about 1.9 percent lower than the experimental value, while resonant frequency fT calculated also has a value of about 1.5 percent lower than the experimental value. Both results agree well. In addition, it is understood from Fig. 3 that according to the calculated values, the flexural vibration and the torsional vibration start to couple strongly at about 50 $\mu m$, and reach the strongest coupling at 52~53 $\mu m$. According to the experimental values, they perform the strongest coupling at about 53 $\mu m$. Thus, it was confirmed from the calculation and the experiments that the flexural vibration and the torsional vibration couple at $\theta = 80°$.

### 4.3 Frequency temperature coefficients and cut angle

Figure 4 shows a relationship of temperature co

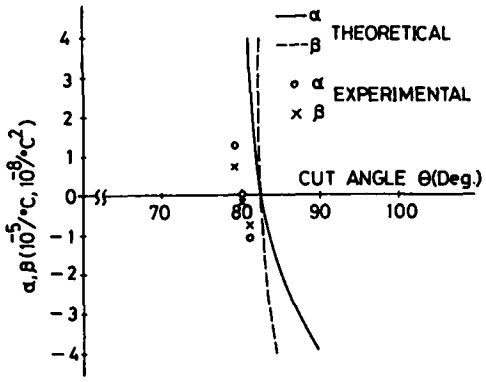



**Fig. 4 Relationship between first and second order temperature coefficients and cut angles.**

efficients and a cut angle for a flexural-torsional quartz crystal resonator which vibrates at fundamental vibrations for both modes. A solid line and a dashed line show the calculated values and the signs O and X, the experimental values. The resonators used in these experiments have a dimension of L=0.95㎜ and W=0.35㎜. A cut angle $\theta$ and thickness T where $\alpha$ and $\beta$ reach zero together, are $\theta = 79.5°$ and T=51$\mu$㎜, respectively, for the experimental values; on the other hand, $\theta = 83.3°$ and T=61.1 $\mu$㎜ for the calculated values. The calculated valueand the experimental value for the cut angle $\theta$ obtained, agree well. However, for the thickness T producing an optimum coupling, there is about a 16 percent difference. It is possible that this difference is due probably to theresonators being calculated as a cantilever approximately or the formation of the resonators by a photolithographic process causing the frequency difference between both mode vibrations. Furthermore, when changing a cut angle $\theta$ at the condition of an optimum coupling, as $\theta$ is decreased, $\alpha$ and $\beta$ obtain more positive values. By contrast, as $\theta$ is increased, they obtain more negative values. The tendency of $\alpha$ and $\beta$ agrees well between the calculated values and the experimental values. In addition, a tuning fork type resonator is designed so that it has novibration leakage at two supporting portions. This is because the vibration leakage exerts an influence on the frequency temperature characteristics.

Figure 5 shows a relationship between the second order temperature coefficient $\beta$ where $\alpha$ 0, and a cut angle $\theta$ for the flexural-torsional quartz crystal resonators which vibrate at the fundamental vibrations for both modes. Solid lines show the values calculated from this theory; a dashed line shows the values calculated by the finite element method of [4], and circles denote the experimental values. The cut angle $\theta$ where $\beta$ 0, is 83.3° in the calculation of this theory, about 80° in the analysis of the finite element method, and 79.5° in the experiments. The results calculated from this theory agree well with those obtained from the finite element method and the experiments. In addition, the calculated results also predict that a cut angle $\theta$ where $\alpha$ $\beta$ 0, is in the vicinity of $\theta$ 40 .

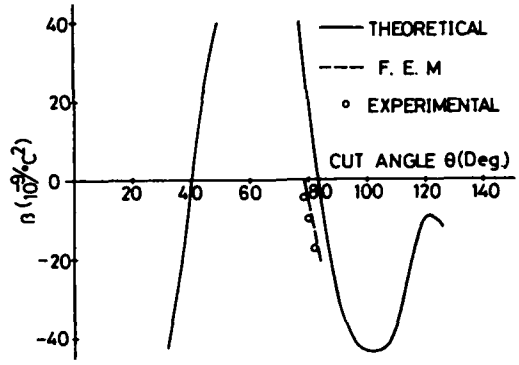



**Fig. 5 Relationship between second order temperature coefficient and cut angles when $\alpha = 0$.**

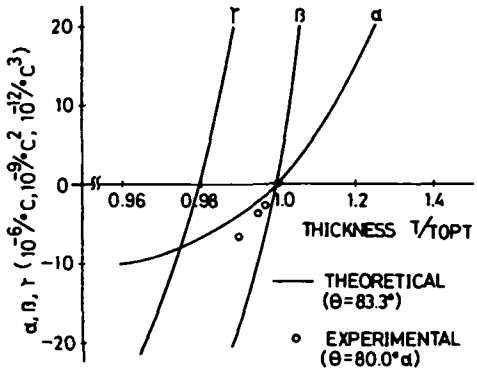



**Fig. 6 Dependence of thickness for temperature coefficients $\alpha$, $\beta$ and $\gamma$.**

### 4.4 Frequency-temperature coefficients and thickness

Figure 6 shows a relationship of $\alpha$, $\beta$, $\gamma$ and thickness T divided by optimum thickness TOPT at which $\alpha$ and $\beta$ reach zero (hereafter, called optimum thickness ratio), of the flexural-torsional quartz crystal resonators which vibrate at the fundamental vibrations for both vibrations. The resonators used here have a dimension of W=0.35㎜ and L=0.95㎜ for both calculation and experiments; $\alpha$ and $\beta$ reach zero at $\theta = 83.3°$ and the optimum thickness TOPT =61.1$\mu$㎜ in the calculation, while the optimum thickness TOPT where $\alpha$ and $\beta$ almost reach zero at $\theta = 80°$ in the experiments, is 50.2 $\mu$㎜. In addition, solid lines in Fig. 6 show the calculated values. A relationship between $\alpha$, $\beta$, $\gamma$ and an optimum thickness ratio at $\theta$ 83.3° is shown also. As the optimum thickness ratio increases, $\alpha$, $\beta$ and $\gamma$ shift to positive values. This is because the coupling between the flexural vibration and the torsional vibration becomes stronger. On the other hand, circles are the experimental values showing a relationship of $\alpha$ vs. an optimum thickness ratio at $\theta$ 80° , whose tendency agrees well with the calculated values.

### 4.5 Frequency temperature coefficients and $\delta$f

Thus far it is shown that frequency temperature

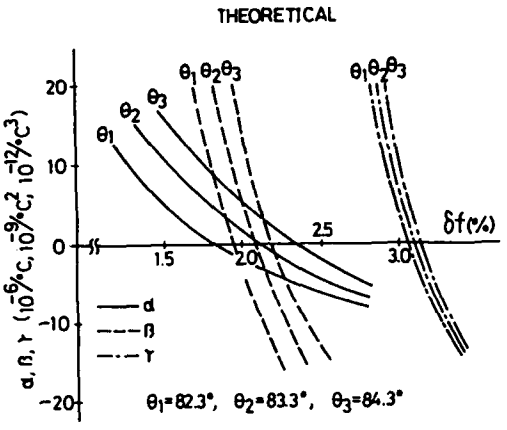



**Fig. 7 Relationship between $\alpha$, $\beta$, $\gamma$ and $\delta f$ when changing cut angles.**

coefficients of a flexural mode resonator coupling to a torsional vibration change according to a dimension, in particular, thickness. This is equivalent to changing the frequency difference between both mode vibrations. That is, the frequency-temperature coefficients can be expressed by the frequency difference of both modes. In addition, in case of forming resonators, it is very convenient that degree of coupling between both modes is expressed by the frequency difference of the flexural mode and the torsional mode, because of the ability to measure these frequencies very accurately. A normalized frequency $\delta f$ between two angular frequencies $\omega_+$ and $\omega_-$ (angular frequency $\omega_+$ for plus, angular frequency $\omega_-$ for minus) is defined as follows:

$$\delta f = \frac{\omega_+ - \omega_-}{\omega_+} \times 100\,(\%) \qquad (2)$$

Next, a relationship between this $\delta f$ and frequency-temperature coefficients is explained below.

Figure 7 shows the calculated values for a relationship between the normalized frequency $\delta f$ and $\alpha$, $\beta$, $\gamma$, when taking a cut angle $\theta$ as a parameter of the flexural torsional quartz crystal resonator which has a dimension of width W 0.35mm and length L 0.95mm, and vibrates at the fundamental vibrations for both flexure and torsion. Three cut angles of $\theta$ 82.3°, $\theta$ 83.3° and $\theta$ 84.3° in Fig. 7 are chosen as the parameters. This theoretical calculation obtains the cut angle $\theta$ 83.3° and $\delta f$ 2.06% at which $\alpha$ and $\beta$ reach zero, thickness T to obtain this $\delta f$ is 61.1$\mu$m, and $\gamma$ has a value of $1.26\times10^{-10}/℃^3$. As $\theta$ increases even at the same $\delta f$, $\alpha$ and $\beta$ become larger as $\delta f$ increases at the same cut angle. In other words, because a coupling between both modes become weak, $\alpha$, $\beta$ and $\gamma$, respectively, become small. Figure 8 shows a relationship between $\delta f$ and $\alpha$, $\beta$ of a flexural torsional quartz crystal resonator which vibrates at the fundamental vibration for torsional vibration, which is cut out at $\theta$ 80°. A solid line and a dashed line show the calculated values of $\alpha$, $\beta$, respectively, while the signs O and X show the experimental values of $\alpha$, $\beta$, respectively. The resonators used in the calculation and the experiments have a dimension of L 2.96mm and W 0.24mm. According to this theoretical calculation and the experiments, $\delta f$ where $\alpha$ reaches zero at $\theta$ 80°, is 2.96 percent for the calculated values, and 3.1 percent for the experimental values. In

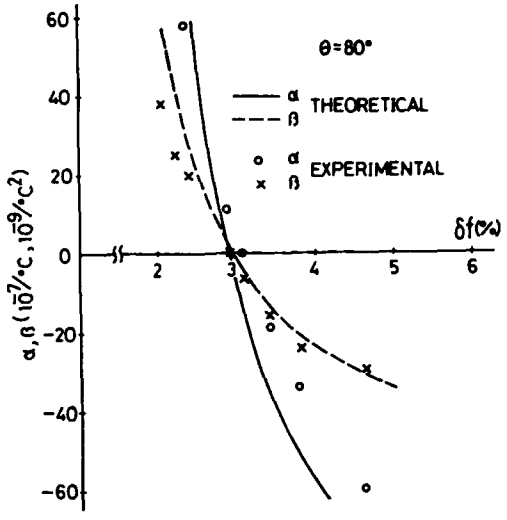



**Fig. 8 Relationship between $\alpha$, $\beta$ and $\delta f$.**

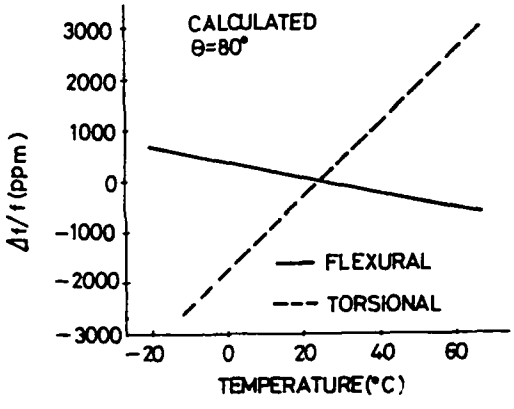



**Fig. 9 Frequeney temperature characteristics.**

addition, $\delta f$ where $\beta$ reaches zero, is 3.0 percent for the calculated values and 2.86 percent for the experimental values. Furthermore, a cut angle $\theta$ and $\delta f$ where both $\alpha$ and $\beta$ almost reach zero ($\alpha = 3.4\times10^{-8}/℃$, $\beta = 1.4\times10^{-8}/℃^2$, $\gamma = 0.95\times10^{-10}/℃^3$), are $\theta$ 80.15° and $\delta f$ 2.98%. On the other hand, $\theta$ 79° and $\delta f$ 2.86 percent in the experimental values ($\alpha = 2\times10^{-10}/℃$, $\beta = 1.5\times10^{-8}/℃^2$, $\gamma = 1.59\times10^{-10}/℃^3$). As is apparent from the above mentioned results, the calculated values and the experimental values agree well.

### 4.6 Frequency temperature characteristics

Figure 9 shows the frequency temperature characteristics of uncoupled flexural and torsional modes for a tuning fork type quartz crystal resonator, which has $\alpha F$ 14.6ppm/℃ and $\alpha T$ 71.0ppm/℃ with $\theta$ 80°. Thus, though the uncoupled resonator has the large value of $\alpha F$, a coupling between both modes improves the value of $\alpha F$.

Figure 10 also shows the frequency temperature characteristics of a tuning fork type quartz crystal resonator with a width of W 0.35mm and a length of L 0.95mm which vibrates at the fundamental vibration for

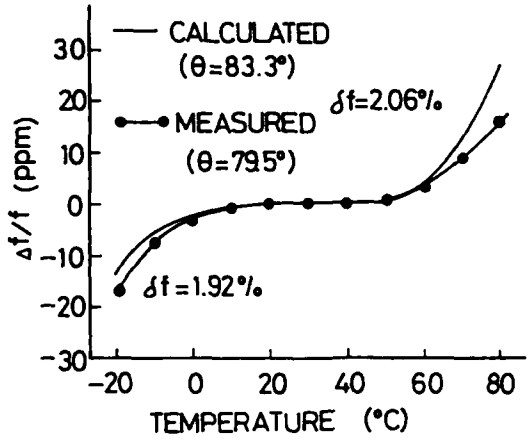



Fig. 10 Frequeney temperature characteristics.

flexure and torsion. The resonator has a value of almost $\alpha = 0$, $\beta = 0$ at $\theta = 83.3°$, $\delta f = 2.06\%$ in the calculation and at $\theta = 79.5°$, $\delta f = 1.92\%$ in the experiments. Thus, the analytical results are in good agreement with the experimental value.

Figure 11 illustrates the frequency-temperature characteristics of a flexural torsional quartz crystal resonator. The resonator vibrates at the second-order harmonic vibration for flexure and at the fundamental vibration for torsion, which has a dimension of L 2.96㎜ and W 0.24㎜. As $\delta f$ decreases, that is, a coupling between both modes of the resonator becomes stronger, a right-down curve shifts to a right-up curve. As was explained in Fig. 8, $\alpha$ and $\beta$ almost reach zero at $\theta = 80.15°$ and $\delta f = 2.98$ percent in the calculation; as is shown in Fig.11, the resonator has excellent frequency temperature characteristics. On the other hand, circles denote the experimental values and show the temperature characteristics at $\theta$ 79° and $\delta f$ 2.86 per cent. Thus, the analysis results of the frequency temperature characteristics obtained by this theoretical analysis agree well with the experimental values.

## 5. Electrical Characteristics

It is essential that a quartz crystal resonator for a wristwatch has low series resistance R1 and a high quality factor Q because of lowering electric current consumption in an oscillating circuit. Therefore, the design for a tuning fork type resonator to obtain small R1 must naturally be made. In addition, it is designed so that the resonator which is small in R1 distribution is obtained. This is because the R1 has a small influence upon a coupling between both mode vibrations, namely, upon frequency temperature characteristics. This applies, for example, for electrodes which excite the fundamental vibrations for flexure and torsion, on a tuning fork type resonator which has dimensions of L 0.95㎜, W 0.35㎜, T 51㎛, and a total width 0.8㎜ and total length 2.85㎜ of the tuning fork type resonator. That is, when disposing the electrodes of length about 80 percent of L, on four surfaces each of the tuning fork arms, the typical values for electrically equivalent circuit constants are shown in Table 1. As is apparent from Table 1, a quartz crystal resonator with low series resistance R1 and a high quality factor Q has been obtained, in addition to the conventional tuning fork type quartz crystal resonator with the resonant frequency f 32 kHz which is used widely in wristwatches. The signs in Table 1 show motional inductance L1, motional capacitance C1, and shunt capacitance C0.

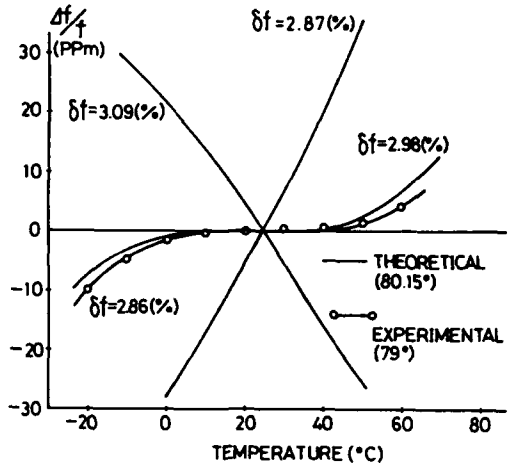



Fig. 11 Frequeney temperature characteristics.

Table 1 Electrical equivalent circuit constants

| カット角 (θ°) | 周波数 (kHz) | $R_1$ (kΩ) | $L_1$ (kH) | $C_1$ (fF) | $C_0$ (pF) | $Q$ (×10⁴) |
|---|---|---|---|---|---|---|
| 76.8 | 196.608 | 11.0 | 1.2 | 0.54 | 0.65 | 13 |

## 6. Conclusions

In this paper, various theoretical and experimental examinations concerning a quartz crystal resonator coupling the torsional mode vibration to the flexural mode vibration have been made for application to a wrist watch. First, a deflection curve including a coupling between the flexural mode vibration and the torsional mode vibration is calculated. Simultaneous differential equations are derived from the equations of motion for the harmonic flexural vibration and the harmonic torsional vibration using this deflection curve. In addition, the frequency equation is obtained by solving the simultaneous differential equations with Galerkin's method. Next, a relationship of resonant frequency, frequency temperature coefficients vs. a cut angle, thickness, a normalized frequency has been clarified. Furthemore, resonators are actually manufactured for trial experiments, as the result of comparing the calculated values with these experimental values. It was confirmed that the analysis results for frequency temperature coefficients obtained from this theoretical analysis agree well with experimental values. In addition, since the obtained resonators have small R1 of about 10 kΩ, a high quality factor and excellent frequency temperature characteristics, they are extremely available for high accuracy wristwatches. Moreover, these

analysis results almost bear comparison with those obtained from the finite element method. However, the advantage of this analysis method is that it is possible to analyze the frequency-temperature characteristics in an extremely short time, that is, the time required in the calculation of the temperature coefficients is approximately one seventieth again as much as that in the finite element method.

Acknowledgement. The author would like to thank M. Akihito Kudo for his help in the experiments of this study.

## REFERENCES

1. H. Kawashima, H. Sato and O. Ochiai, New frequency-temperature charactertistics of miniaturized GT cut quartz resonators, Proceedings of the 34th Annual Symposium on Frequency Control, pp. 131-139 (1980).
2. M. Nakazawa and S. Kozima, A study of GT-type quartz crystal plate, I. E. E. E. Trans. Sonics & Ultrason. SU-29, 3 (May 1982).
3. W.P. Mason, A new quartz crystal plate, designated the GT, which produces a very constant frequency over a wide temperature range, Proc. IRE, pp. 220-223 (1940).
4. Momosaki, Inoue and Kogure, A tuning fork type quartz crystal resonator with the same frequency temperature characteristics as those of an AT plate, Nikkei Electronics, 11 26, 226 (1979).
5. S. G. Lekhnitski, Theory of Elasticity of an Anisotropic Body, San Francisco, California, Holden Day (1963).
6. K. Matsudaira, Fundamental Vibration, Kyoritsu Publishing (1950).
7. R. Bechmann, A. D. Ballato and T. J. Lukaszek, High order temperature coefficients of the elastic stiffness and compliances of alpha quartz, Proc. IRE, 50, pp. 1812-1822 (1962).
8. A. E. H. Love, A Treatise on the Mathematical Theory of Elasticity, Dover-New York (1934).

## APPENDIX

Calculation of $\alpha$

$$\frac{\dot{\omega}_F}{\omega_F}=\frac{1}{2}\left(2\frac{\dot{W}}{W}-\frac{\dot{\rho}}{\rho}-\frac{\dot{S}'_{33}}{S'_{33}}-4\frac{\dot{L}}{L}\right)$$

$$\frac{\dot{\omega}_T}{\omega_T}=\frac{1}{2}\left\{\frac{\dot{C}_t}{C_t}-\frac{\dot{\rho}}{\rho}-2\frac{\dot{L}}{L}-\frac{\dot{W}}{W}-\frac{\dot{T}}{T}-\frac{2}{(W^2+T^2)}\cdot\left(W^2\frac{\dot{W}}{W}+T^2\frac{\dot{T}}{T}\right)\right\}$$

$$\frac{\dot{\eta}}{\eta}=\left(1-\frac{1}{\eta}\right)\left(2\frac{\dot{S}'_{34}}{S'_{34}}+\frac{\dot{C}_t}{C_t}-\frac{\dot{S}'_{33}}{S'_{33}}-3\frac{\dot{W}}{W}-\frac{\dot{T}}{T}\right)$$

$$g=\{(\omega_F^2-\omega_T^2)^2+4C\omega_F^2\omega_T^2\}^{1/2}$$

$$\dot{g}=\frac{2}{g}\left\{(\omega_F^2-\omega_T^2)\left(\omega_F^2\frac{\dot{\omega}_F}{\omega_F}-\omega_T^2\frac{\dot{\omega}_T}{\omega_T}\right)+C\omega_F^2\omega_T^2\left(\frac{\dot{C}}{C}+2\frac{\dot{\omega}_F}{\omega_F}+2\frac{\dot{\omega}_T}{\omega_T}\right)\right\}$$

In addition, if taking the coefficients of linear expansion $\dot{X}/X$, $\dot{Z}/Z$ of quartz in the X axis direction and the Z axis direction, respectivly,

$$\frac{\dot{W}}{W}=\frac{\dot{X}}{X}$$

$$\frac{\dot{T}}{T}=(1-n_2^2)\frac{\dot{X}}{X}+n_2^2\frac{\dot{Z}}{Z}$$

$$\frac{\dot{L}}{L}=(1-n_3^2)\frac{\dot{X}}{X}+n_3^2\frac{\dot{Z}}{Z}$$

$$\frac{\dot{\rho}}{\rho}=-\left(2\frac{\dot{X}}{X}+\frac{\dot{Z}}{Z}\right)$$

$$\frac{\dot{C}}{C}=\frac{\eta}{\eta-1}\frac{\dot{\eta}}{\eta}\qquad\because\ C=K_2^2$$

$$S'_{33}=S_{11}m_3^4+2S_{13}m_3^2n_3^2-2S_{14}m_3^3n_3+S_{33}n_3^4+S_{44}m_3^2n_3^2$$

$$S'_{34}=2S_{11}m_2m_3^3+2S_{13}m_3n_3(m_3n_2+m_2n_3)-S_{14}m_3^2(m_3n_2+3m_2n_3)+2S_{33}n_2n_3^2+S_{44}m_3n_3(m_2n_3+m_3n_2)$$

$$S'_{44}=4S_{11}m_2^2m_3^2+8S_{13}n_2n_3m_2m_3-4S_{14}m_2m_3(m_2n_3+m_3n_2)+4S_{33}n_2^2n_3^2+S_{44}(m_2n_3+m_3n_2)^2$$

$$S'_{55}=S_{66}m_3^2+4S_{14}m_3n_3+S_{44}n_3^2$$

$$m_2=\cos\theta,\quad n_2=\sin\theta$$

$$m_3=-\sin\theta,\quad n_3=\cos\theta$$

(Rotation angle about X axis)

Furthermore, the torsional rigidity Ct is obtained from [8]

$$C_t\fallingdotseq\frac{WT^3}{3S'_{55}}\left(1-\frac{T}{W}\sqrt{\frac{S'_{44}}{S'_{55}}}\frac{192}{\pi^5}\right)=\frac{WT^3}{3S'_{55}}(1-k)$$

When $W/T>>1$, since the term of "tanh" approximately reaches 1, Ct is simplified.

$$C_t=\frac{WT^3}{3S'_{55}}\left\{1-\frac{T}{W}\sqrt{\frac{S'_{44}}{S'_{55}}}\frac{192}{\pi^5}\sum_{n=0}^{\infty}\frac{1}{(2n+1)^5}\cdot\tanh\left(\frac{(2n+1)\pi}{2}\frac{W}{T}\sqrt{\frac{S'_{55}}{S'_{44}}}\right)\right\}$$

where $\quad k=\frac{T}{W}\sqrt{\frac{S'_{44}}{S'_{55}}}\frac{192}{\pi^5}$

From the above relation,

$$\frac{\dot{C}_t}{C_t}=\frac{\dot{W}}{W}+3\frac{\dot{T}}{T}-\frac{\dot{S}'_{55}}{S'_{55}}-\frac{k}{(1-k)}\frac{\dot{k}}{k}$$

$$\frac{\dot{k}}{k}=\frac{\dot{T}}{T}-\frac{\dot{W}}{W}+\frac{1}{2}\frac{\dot{S}'_{44}}{S'_{44}}-\frac{1}{2}\frac{\dot{S}'_{55}}{S'_{55}}$$

From these relating equations, $\alpha$ is calculated as a parameter of a dimension and a cut angle, and also $\beta$ and $\gamma$ are calculated as well.

42nd Annual Frequency Control Symposium - 1988

# TEMPERATURE DERIVATIVES OF THE DYNAMIC PERMITTIVITY AND PERMEABILITY OF THE SIMPLE THICKNESS MODES OF QUARTZ PLATES

J. KOSINSKI*, A. BALLATO*, T. LUKASZEK, M. MIZAN, R. McGOWAN, and K. KLOHN, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ 07703 (*Also with Monmouth College, West Long Branch, NJ 07764).

## ABSTRACT

Measurements of selected quartz material properties are being made as part of a cooperative Army/Air Force/EIA/RPI effort. The quartz material properties are obtained from measurements of the mass-loaded resonance frequencies of various harmonics of a set of carefully fabricated thickness-mode resonators. The measurement system is based on IEC-444, and its performance has been extensively verified. To date, one set of resonators employing thickness excitation of the thickness modes has been measured using this system.

The data thus obtained have been reduced using a one dimensional analysis incorporating the effects of mass-loading and the full effects of piezoelectricity, yielding preliminary information on the temperature coefficients of certain elastic constants and on the temperature coefficients of piezoelectric coupling for rotated Y-cuts. The temperature coefficients of the dynamic permittivity and dynamic permeability of the simple thickness modes of quartz plates are developed from the data as a sample application of the new measurements.

## INTRODUCTION

The dielectric, piezoelectric, and elastic behavior of quartz can be described as tensors of rank two, three, and four, respectively. Owing to crystal symmetry, the number of independent material constants reduces to two dielectric, two piezoelectric, and six elastic constants. In the case of rotated Y-cuts, only three of the six elastic constants are excited when driving simple thickness modes.

Measurements of the piezoelectric, dielectric, and elastic constants of quartz have been made by a large number of researchers [1-16]. The values of the piezoelectric, dielectric, and elastic constants of interest here, measured at constant temperature, are shown in Table I, along with the density and thermal expansion coefficients of quartz. The values of the piezoelectric, dielectric, and isagric elastic constants are from Bechmann [3], the density is from Cady [6], and the thermal expansion coefficients are from Sosman's data as analyzed by Mizan and Ballato [10].

The temperature coefficients of the piezoelectric, dielectric, and elastic constants are not nearly as well known as the values of the constants themselves [17-42]. Table II shows the published data on the temperature coefficients of the piezoelectric and dielectric constants. The values from Sosman cover only a very limited temperature range, while the values by Westphal are the best estimate from an inconsistent data set.

TABLE I. Physical Properties of $\alpha$-Quartz

Density ($\times 10^3 kg/m^3$)

$\rho = 2.648$

Thermal Expansion ($T_o = 25^oC$)

| Order | $\alpha^{(n)}_{11}$ | $\alpha^{(n)}_{33}$ | Units |
|---|---|---|---|
| 1 | 13.71 | 7.48 | $\times 10^{-6}/^oC$ |
| 2 | 6.5 | 2.9 | $\times 10^{-9}/^oC^2$ |
| 3 | -1.9 | -1.5 | $\times 10^{-12}/^oC^3$ |

Permittivity Constants ($\times 10^{-12} F/m$)

| $\epsilon^S_{11}$ | $\epsilon^S_{33}$ |
|---|---|
| 39.21 | 41.03 |

Piezoelectric Stress Constants ($C/m^2$)

| $e_{11}$ | $e_{14}$ |
|---|---|
| 0.171 | -0.0406 |

Elastic Stiffness Constants ($\times 10^9 N/m^2$)

| $c^E_{11}$ | $c^E_{33}$ | $c^E_{44}$ | $c^E_{12}$ | $c^E_{13}$ | $c^E_{14}$ | $c^E_{66}$ |
|---|---|---|---|---|---|---|
| 86.74 | 107.2 | 57.94 | 6.99 | 11.91 | -17.9 | 39.88 |

TABLE II. Temperature Coefficients of the Dielectric and Piezoelectric Constants

Temperature Coefficients of the Dielectric Constants ($\times 10^{-4}/^oC$)

| $T_{\epsilon^S_{11}}(1)$ | $T_{\epsilon^S_{33}}(1)$ | $T_o$ | Source |
|---|---|---|---|
| -9.9 | -11.0 | $10^oC$ | Sosman |
| 0.28 | 0.39 | -- | Westphal |

Temperature Coefficients of the Piezoelectric Stress Constants ($\times 10^{-4}/^oC$)

| $T_{e_{11}}(1)$ | $T_{e_{44}}(1)$ | To | Source |
|---|---|---|---|
| 1.6 | -14.4 | -- | Bechmann |

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT 

TABLE III. First Order Temp. Coefficients of Selected Elastic Constants ($x10^{-6}/^{o}C$)

| $T_{C_{14}^E}^{(1)}$ | $T_{C_{44}^E}^{(1)}$ | $T_{C_{66}^E}^{(1)}$ | $T_o$ | Source |
|---|---|---|---|---|
| 82 | -151 | 144 | -- | Bechmann |
| 107 | -169 | 170.1 | 35°C | Atanasoff & Hart |
| 98 | -169 | 168 | 20°C | GPO Res. Report |
| 105 | -158 | 169 | 50°C | Mason |
| 98 | -172 | 180 | 20°C | Koga |
| 101 | -177 | 178 | 25°C | Bechmann, et al. |
| 89 | -172 | 167 | 25°C | Adams, et al. |
| 93.2 | -184 | 180 | 25°C | Kahan |

TABLE IV. Second Order Temp. Coefficients of Selected Elastic Constants ($x10^{-9}/^{o}C^{2}$)

| $T_{C_{14}^E}^{(2)}$ | $T_{C_{44}^E}^{(2)}$ | $T_{C_{66}^E}^{(2)}$ | $T_o$ | Source |
|---|---|---|---|---|
| -82 | -233 | 193 | 20°C | GPO Res. Report |
| -270 | -212 | -5 | 50°C | Mason |
| -13 | -225 | 201 | 20°C | Koga |
| -48 | -216 | 118 | 25°C | Bechmann, et al. |
| -19 | -261 | 164 | 25°C | Adams, et al. |
| -46.6 | -273 | 172 | 25°C | Kahan |

TABLE V. Third Order Temp. Coefficients of Selected Elastic Constants ($x10^{-12}/^{o}C^{3}$)

| $T_{C_{14}^E}^{(3)}$ | $T_{C_{44}^E}^{(3)}$ | $T_{C_{66}^E}^{(3)}$ | $T_o$ | Source |
|---|---|---|---|---|
| -630 | -65 | -167 | 50°C | Mason |
| -625 | -190 | -777 | 20°C | Koga |
| -590 | -216 | 21 | 25°C | Bechmann, et al. |
| -521 | -194 | 29 | 25°C | Adams, et al. |
| -612 | -247 | 25.4 | 25°C | Kahan |

Tables III - V show the published values of the temperature coefficients of the elastic constants of interest here. While the earliest values listed are from Bechmann [19], determinations were also made independently by Koga [30] and Lack, Willard, and Fair [32] at about the same time. These early measurements identified those loci where the first order temperature coefficient of quartz is zero, enabling the development of a large class of temperature insensitive resonator types. The data from Atanasoff and Hart [18] on through Bechmann, Ballato, and Lukaszek [23,24] all constitute refinements of these original measurements, with Bechmann, et al. representing the most comprehensive work listed.

There are some important limitations to the accuracy of the published values. First, the majority of the measurements were made using the quartz blanks in oscillator circuits, with the attendant addition of oscillator-induced errors. Secondly, the theoretical analysis used to reduce the data did not account for mass loading or the full effects of piezoelectricity. As a result, the values of the temperature coefficients are not sufficiently well known at the present time, and the goal of this cooperative effort is a more accurate determination of the values of these coefficients using modern instrumentation and data reduction techniques.

MEASUREMENT SYSTEM

The measurement system being employed is based on International Electrotechnical Commission Publication 444, "Basic method for the measurement of resonance frequency and equivalent series resistance of quartz crystal units by zero phase technique in a pi-network." [43] This recommendation specifies a simple method of measuring the resonance frequency of quartz crystal units with a fractional frequency accuracy ranging between $10^{-6}$ and $10^{-8}$ depending on the type of crystal. The measurement system is based on a pi-network whose electrical characteristics are fully specified.

A block diagram of the basic measurement system specified in IEC-444 is shown in Figure 1. The phase of the crystal transfer admittance is indicated on a phase meter connected across the pi-network, and the frequency giving zero phase reading is measured.

Zero phase is calibrated by inserting a short-circuit in the pi-network. The addition of a temperature chamber and thermometer to the basic system enable us to use the system to measure temperature coefficients. A block diagram of the system as actually implemented is shown in Figure 2. Pertinent parameters of the auxiliary measurement equipment are as follows:

- o HP8405A Vector Voltmeter
  - repeatability $\pm 1.5^{o}$ phase
  - analog output -0.5 to +0.5V dc (2.8mV/°phase)
  - settling time 10 milliseconds
- o FLUKE 8506A Digital Multimeter
  - resolution $10^{-4}$mV (<40 $\mu$deg phase, <<1.5° repeatability)
  - zero point drift $5x10^{-3}$mV over 90 days
- o HP8642B Signal Generator
  - 1 Hz resolution (better than $10^{-7}$ for frequencies of interest)
  - repeatability $<10^{-9}$/day
  - 0.1 - 2100 MHz range

- o ANZAC H-8-4 Hybrid Bridge
  - broadband in-phase 2-way power divider
  - 50 ohms all ports
  - phase response A relative B 2.5° max
- o FLUKE 4250A BCD Programmable Power Source
  - 0 to 10V
  - 1mV resolution
- o HP2804A Quartz Thermometer
  - 40 millidegree Celsius absolute accuracy
  - NBS traceable
- o Pi-network
  - ceramic hybrid microcircuit
  - fixtures designed and fabricated for this experiment by the authors

The system is controlled by a Hewlett-Packard 9825B computer using software written in the hpl language.

PERFORMANCE VERIFICATION

The test system was subjected to an extensive test regimen to verify its accuracy and precision before any data runs were made. The two main areas of concern were test system thermal performance and calibration of the zero phase reference. Verification of test system thermal performance involved careful characterization of the temperature chamber, while calibration of the zero phase reference involved extensive testing of the pi-network.

Figure 3 shows a typical test made during the verification of test system thermal performance. The figure shows chamber temperature versus time in response to a command to change temperature. The dwell time required to achieve thermal equilibrium is readily observed from this figure. Figure 4 is a close-up of the same data focusing on the final chamber temperature. The equilibrium in the chamber is seen to be characterized by random fluctuations about a residual thermal drift.

Similar tests were made over the entire system operating range of -50°C to +150°C. The results are shown in Figures 5 through 7. Based on these results, a stabilization time of 50 minutes at each step was chosen for the data runs. The combined effect of both the random and linear temperature excursions experienced during the period that measurements are actually being made (approximately one minute per crystal) is less than the NBS traceable accuracy of the thermometer. The accuracy of the temperature data is therefore taken to be 40 millidegrees Celsius.

One further thermal consideration was the maximum rate of temperature change during the random fluctuations. This is of concern due to the dynamic thermal behavior [44,45] of quartz resonators. That is, the first order temperature coefficient is acutally given by

$$T_f^{(1)} = a_o + \tilde{a}\cdot\dot{T}(t) \qquad (1)$$

where $a_o$ is the static temperature coefficient we are interested in and $\tilde{a}$ is the dynamic temperature coefficient, with $\tilde{a}$ typically [46,47] less than $\pm 2\times10^{-7}$ sec/°C$^2$. The maximum rate of temperature change measured while the system was dwelling at equilibrium was 10 millidegrees per second, implying a worst case dynamic thermal contribution of $\pm 2\times10^{-9}$/°C as compared to typical values of $a_o$ on the order of $10^{-6}$/°C.

Characterization of the zero phase reference was performed across the entire range of frequencies and temperatures used in data runs later on. These tests involved placing short-circuits in the pi-networks, and measuring any changes in phase reading across the network as the temperature and frequency were varied. Typical data from a characterization run are shown in Figure 8.

The results of the characterization runs, when combined with the typical frequency versus phase slope of the test samples, showed the zero phase reference to be quite good. At operating frequencies below 50 MHz, the zero phase reference did not vary systematically as the temperature was changed. Above 50 MHz, a slight temperature dependence of $10^{-2}$ degrees of phase per degree Celsius was found. With the typical crystal frequency/phase slope of $10^{-7}$ per degree of phase, this again amounts to a negligible contribution of $10^{-9}$/°C to the first order temperature coefficient. The frequency behavior of the networks was even better. Over the 10 ppm sweep range used in the tests, the zero phase reference was found to vary by less than $10^{-3}$ degrees of phase.

One final consideration was the amount of power which would be used to excite the test crystals. This was a concern since the elastic properties of quartz become strongly non-linear above a certain power level [48,49]. This threshold level was determined as 0.5mW for the units used here. The data runs were therefore all made with a crystal driving power well below this threshold, being measured at the test fixture socket as 20nW.

TEST RESONATORS

Sixty rotated Y-cut resonators were obtained for this experiment. They encompassed six different nominal orientation angles:

| | |
|---|---|
| (YXl) -49° | BT-cut |
| (YXl) -45° | |
| (YXl) 0° | Y-cut |
| (YXl) +35.25° | AT-cut |
| (YXl) +40.8° | |
| (YXl) +45° | |

The test units were manufactured by

PIEZO Crystal Co. of Carlisle, PA to the following unit specifications:

- o Quartz Material: Premium Q, pure Z growth, cultured quartz.
- o Frequency: Approximately 10 MHz fundamental, 1/6mm thick.
- o Blank Geometry: Round,14mm diameter,plano-plano, 1μm surface finish.
- o Electrode Geometry: Au keyhole, 7mm diameter with 2.5mm tab, thickness approx. 1000Å.
- o Mount: HC-6/U, solder seal, Ni ribbon slots, bonding cement.

The orientation angle actually achieved for each blank was determined by the manufacturer to an accuracy of ±1′ of arc on all units. These actual measured values were later used in the data reduction.

Upon receipt from the manufacturer, all units were checked for undesired or interfering modes using an HP8407 Network Analyzer system. In all cases, undesired modes were separated from the desired resonance in question by greater than 1% of the resonance frequency.

EXPERIMENTAL DATA

Measurements were made of the frequencies of all excitable overtones as a function of temperature for the complete set of test resonators. The measurements were made under computer control with the use of an r.f. switch allowing five units to be tested per run. A typical run consisted of a one-hour soak at -50°C, followed by 4°C step increases in the chamber temperature until 150°C was reached. At each temperature step, the measurement sequence was:

1) wait for thermal equilibrium
2) find zero phase for unit #1
3) measure temperature
4) measure unit #1 frequency vs. phase over a narrow range about resonance
5) re-measure temperature
6) repeat steps 2-5 for units #2 - #5
7) record data
8) go to the next temperature

After the run was completed, the frequency versus phase data taken for each particular crystal at each temperature step was curve-fitted by the method of least-squares in order to obtain the frequency versus phase response of each crystal at each temperature. Finding the frequencies of zero phase in this fashion allowed for an order of magnitude improvement in frequency resolution. The frequencies of zero phase and their corresponding temperature measurements were then reduced by the method of least-squares to yield the temperature coefficients of frequency for each of the crystals under test.

Figures 9 through 11 present a sample set of the experimental data obtained. In the lower curve on each figure, crosses represent the measured data while the solid line represents the least-squares fit power series expansion about $T_o = 25^oC$. The temperature coefficients of frequency for orders zero through four are listed at left center, and the curve at the top depicts the deviation between the data and the least-squares fit. The coefficients listed on each curve, along with the measured orientation angles and nominal electrode thicknesses, formed the data set from which the material coefficients were then determined.

The number of significant figures for each of the temperature coefficients of frequency was determined by performing repeated tests on a sample group of 10 units. Half of the units were type (YXl)+35.25°, and half were type (YXl)-45°. Each sub-group was run four times, and the coefficients obtained in each run recorded. The repeatability did not seem to depend on the orientation angle, but did depend on the order of the coefficient. The results are summarized in Table VI. The number of significant figures is seen to be 4,3,2, and 1 for orders 1,2,3, and 4, respectively.

TABLE VI. Measurement Repeatability

| Coeff. | Range | Repeat. | Units |
|---|---|---|---|
| $T_f^{(1)}$ | -45 to +91 | 0.06 | $10^{-6}/^oC$ |
| $T_f^{(2)}$ | -57 to +79 | 0.13 | $10^{-9}/^oC^2$ |
| $T_f^{(3)}$ | -114 to +101 | 1.6 | $10^{-12}/^oC^3$ |
| $T_f^{(4)}$ | -44 to +268 | 13 | $10^{-15}/^oC^4$ |

PRELIMINARY DATA ANALYSIS

Expressions for the resonance and antiresonance frequencies of mass-loaded quartz resonators have been determined by Ballato and Lukaszek [50]. These frequencies are determined under assumptions of linearity, lateral unboundedness, equal masses lumped on the surfaces, and thickness excitation of a single mode. They are given as

$$X = (\pi/2)\ (f_{R_\mu}^{(M)}/f_{A_o}^{(1)})\ , \qquad (2)$$

where X is a root of

$$\tan X = X/(k^2+\mu X^2) \qquad (3)$$

for the resonance frequencies, and

$$X = (\pi/2)\ (f_{A_\mu}^{(M)}/f_{A_o}^{(1)})\ , \qquad (4)$$

where X is a root of

$$\tan X = 1/\mu X \qquad (5)$$

for the antiresonance frequencies.

In these relations, $\mu$ is the normalized mass loading and k is the piezoelectric coupling of the mode in question. The symbol $f_{A_o}^{(1)}$ represents the fundamental (1st harmonic) antiresonance frequency with zero-mass loading, while $f_{R_\mu}^{(M)}$ and $f_{A_\mu}^{(M)}$ represent the resonance and antiresonance frequencies of the $M^{th}$ harmonic in the presence of mass loading. Because of the antisymmetric nature of the driving electric field (electrodes instantaneously at opposite polarities with respect to each other), the harmonic number M can only take on odd values. In the absence of mass loading, the antiresonance frequencies are harmonically related and may be written as

$$f_{A_o}^{(M)} = M \cdot f_{A_o}^{(1)} = \frac{M}{2(2h)} \sqrt{\frac{\bar{c}}{\rho}} . \qquad (6)$$

where 2h is the plate thickness, $\rho$ is the density of quartz, and $\bar{c}$ is the piezoelectrically-stiffened elastic constant for the mode in question.

Ballato and Lukaszek have combined these equations and taken power series expansions to produce an equation set for calculating the temperature coefficients of frequency of simple thickness modes from the published values of the material temperature coefficients. For this initial data analysis, the equation set has been inverted and transformed into a set of equations whereby the material temperature coefficients may be determined from a measured set of frequency versus temperature curves [51].

Obtaining the material temperature coefficients from the data set requires several steps. First, the temperature coefficients of piezoelectric coupling are determined from the differences between the frequency-temperature behavior of multiple overtones of individual resonators. The values of the temperature coefficients of piezoelectric coupling are then used to transform the measured temperature coefficients of the mass loaded resonance frequencies into temperature coefficients of the zero-mass-loading, fundamental, antiresonance frequencies, from which the temperature coefficients of the isagric elastic constants are then obtained. Lastly, the temperature coefficients of the dielectric and piezoelectric constants are determined from the temperature coefficient of coupling data.

## MATERIAL CONSTANTS

The measured frequency-temperature data, along with the nominal orientation angles and electrode thicknesses, were entered into a computer data base and analyzed using the inverted equation set. The analysis of the temperature coefficients of the elastic constants converged quite rapidly. The preliminary values of the temperature coefficients of the isagric elastic constants thus obtained are given in Table VII. The processed data for the first order temperature coefficient of the zero-mass-loading, fundamental, antiresonance frequency for the rotated Y-cut is shown in Figure 12, along with the least-squares curve fit using the values from Table VII.

TABLE VII. Preliminary Results For Selected Temperature Coefficients

| Order | $T_{C_{14}^E}^{(n)}$ | $T_{C_{44}^E}^{(n)}$ | $T_{C_{66}^E}^{(n)}$ | Units |
|---|---|---|---|---|
| n = 1 | 103.9 | -170.4 | 177.0 | $10^{-6}/^oC$ |
| n = 2 | -24.7 | -251 | 167 | $10^{-9}/^oC^2$ |
| n = 3 | -500 | -220 | 47 | $10^{-12}/^oC^3$ |

The first order temperature coefficient of coupling data used to determine the temperature coefficients of the piezoelectric and dielectric constants and the least-squares curve fit using the values thus obtained are shown in Figure 13. The least-squares analysis is seen to be quite good, although some points should be noted regarding the input data.

The data at $\theta = -45^o$, $\theta = 0^o$, and $\theta = +35.25^o$ each include values for the temperature coefficient of coupling based on the differences between both fundamental/third overtone and fundamental/fifth overtone measurement pairs. Both sets of measurement pairs at $\theta = +35.25^o$ yield the same values for the temperature coefficient, as would be expected if the measurements are accurate. At $\theta = 0^o$, both sets yield the same average values for the temperature coefficient even though the fundamental/third overtone pairs show a large scatter in the data, caused by the poor operation of the (YXl)$0^o$ units on the third overtone.

The divergence in temperature coefficient of coupling data at $\theta = -45^o$, along with the similarity between the $\theta = -45^o$ and $\theta = -49^o$ fundamental/third overtone data sets are clearly the weakest parts of the data analysis. It is likely that the actual mass loading on these units differs from the 1000Å specified, which would cause this type of problem in the analysis. Discussions with the manufacturer have confirmed that some units did receive additional plating, although the details of which units received the additional plating could not be ascertained from manufacturing records. Work is underway to obtain actual measurements of the film thicknesses, and values for the temperature coefficients of the dielectric and piezoelectric constants will be published when the mass loading question is satisfactorily resolved.

## DYNAMIC PERMITTIVITY AND PERMEABILITY

The dynamic permittivity and permeability

are normalized design parameters ideally dependent only upon the effective, rotated material constants of the mode being excited [52-56]. In the case of thickness excitation of simple thickness modes (TETM), the dynamic permittivity of overtone M is given by

$$\Gamma_1^{(M)} = \frac{8\,\epsilon^{S}\,k^2}{M^2\,\pi^2} \cdot \Psi^{(M)} \qquad (7)$$

while the dynamic permeability is given by

$$\Lambda_1^{(M)} = \frac{\rho}{8\,e^2} \cdot [\,\Psi^{(M)}\,]^{-1} \qquad (8)$$

The quantity $\Psi^{(M)}$ in equations (7) and (8) describes the distribution of motion in the electroded region, with $\Psi^{(M)}=1$ for all overtones in the ideal case of a uniform distribution of motion. In reality, the distribution of motion is not only not uniform, but the value of $\Psi^{(M)}$ may be different for different overtones. While we shall present only the ideal case here, extension to the non-ideal case is straightforward.

In the ideal case, the fundamental mode dynamic permittivity and permeability are both material constants fixed by the choice of orientation angle. The dynamic permeability is independent of overtone, while the dynamic permittivity decreases as the square of the overtone. By knowing the fundamental mode dynamic permittivity and permeability, the motional capacitance and motional inductance of any overtone of a TETM resonator are readily calculated from the dimensions of the device. The motional capacitance is simply

$$C_1^{(M)} = \frac{\Gamma_1^{(1)}}{M^2} \cdot \frac{A}{2h} \,, \qquad (9)$$

and the motional inductance is simply

$$L_1^{(M)} = \Lambda_1^{(1)} \cdot \frac{(2h)^3}{A} \,, \qquad (10)$$

where 2h is the plate thickness and A is the electrode area.

The temperature coefficients of the dynamic permittivity and permeability follow directly from the normalized power series expansions in temperature of equations (7) and (8). In the ideal case, the temperature coefficients of both are independent of overtone number. The ideal case values of the fundamental mode temperature coefficients for the rotated Y-cut TETM resonator are plotted in Figures 14 and 15 as a function of orientation angle. As before, extension to the non-ideal case is straightforward.


## ACKNOWLEDGMENT

The authors wish to acknowledge the contributions of Mr. Dana Sturzebecher, who assisted in performing a number of the measurements.


## REFERENCES


1. S. Annaka and A. Nemoto, "Piezoelectric Constants of ∝-Quartz Determined from Dynamical X-ray Diffraction Curves," J. Appl. Crystallogr., Vol. 10, pp. 354-355, 1977.

2. R. Bechmann, "Contour Modes of Square Plates Excited Piezoelectrically and Determination of Elastic and Piezoelectric Coefficients," Proc. Phys. Soc. London B, Vol. 64, pp. 323-337, 1951.

3. R. Bechmann, "Elastic and Piezoelectric Constants of ∝-Quartz," Phys. Rev., Vol. 110, No. 5, pp. 1060-1061, June 1958.

4. V. Bottom, "Measurement of the Piezoelectric Coefficients of Quartz Using the Fabry-Perot Dilatometer," J. App. Phys., Vol. 41, No. 10, p. 3941, September 1970.

5. V. Bottom, "Dielectric Constants of Quartz," J. App. Phys., Vol. 43, No. 4, p. 1493, April 1972.

6. W. Cady, Piezoelectricity. New York : McGraw-Hill, 1946.

7. A. Lawson, "Comment on the Elastic Constants of Alpha-Quartz," Phys. Rev., Vol. 59, pp. 838-839, 1941.

8. H. McSkimin, "Measurement of the 25°C Zero-Field Elastic Moduli of Quartz by High Frequency Plane-Wave Propagation," J. Acoust. Soc. Am., Vol. 34, pp. 1271-1274, 1962.

9. H. McSkimin, P. Andreatch, and R. Thurston, "Elastic Moduli of Quartz Versus Hydrostatic Pressure at 25° and -195.3°C," J. App. Phys., Vol. 36, No. 5, pp. 1624-1632, May 1965.

10. M. Mizan and A. Ballato, "Stress Coefficients of Doubly Rotated, Low Temperature Coeffi nt, Quartz Resonators," USAERADCOM Technica eport DELET-TR-84-10, US Army Electronics hnology and Devices Laboratory, Fort Monmouth, NJ, p.76, September 1984.

11. M. Shevel'ko and L. Yakovlev, "Precision Measurements of the Elastic Characteristics of Synthetic Piezoelectric Quartz," Sov. Phys. - Acoust., Vol. 23, pp. 187-188, 1977.

12. Y. Shmin, "Theoretical and Experimental Investigation of the Elasticity Constants of

Quartz," Sov. Phys. J., Vol. 22, pp. 956-961, 1970.

13. R. Sosman, The Properties of Silica, New York : Chemical Catalog Co., 1927.

14. R. Stern and R. Smith, "On the Third-Order Elastic Moduli of Quartz," J. Acous. Soc. Am., Vol. 44, pp. 640-641, 1968.

15. R. Thurston, H. McSkimin, and P. Andreatch, "Third-Order Elastic Coefficients of Quartz," J. App. Phys., Vol. 37, No. 1, pp. 267-275, January 1966.

16. A. Warner, M. Onoe, and G. Coquin, "Determination of Elastic and Piezoelectric Constants for Crystals in Class (3m)," J. Acoust. Soc. Am., Vol. 42, pp. 1223-1231, 1967.

17. C. Adams, G. Enslow, J. Kusters, and R. Ward, "Selected Topics in Quartz Crystal Research," Proc. 24th Annual Frequency Control Symposium, U.S. Army Electronic Components Laboratory, Fort Monmouth, NJ, pp. 53-63, April 1970.

18. J. Atanasoff and P. Hart, "Dynamical Determination of the Elastic Constants and Their Temperature Coefficients for Quartz," Phys. Rev., Vol. 59, pp. 85-96, 1 January 1941.

19. R. Bechmann, "Über die Temperatur-Koeffizienten der Eigen-Schwingungen piezoelektrischer Quarzplatten und Stäbe," Hochfrequenz. und Elektroak., Vol. 44, pp. 145-160, November 1934.

20. R. Bechmann, "Temperaturabhängigkeit von Quarzresonatoren," Archiv der elekr. Übertragung, Vol. 5, pp. 89-90, 1951.

21. R. Bechmann, "Über die Temperaturabhängigkeit der Frequenz von AT- und BT-Quarzresonatoren," Archiv. der elekr. Übertragung, Vol. 9, pp. 513-518, 1955.

22. R. Bechmann and V. Durana, "Variation With Temperature of Quartz Resonator Characteristics," Proc. IRE, Vol. 44, No. 3, p. 377, March 1956.

23. R. Bechmann, A. Ballato, and T. Lukaszek, "Higher-Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha-Quartz," Proc. IRE, Vol. 50, No. 8, pp. 1812-1822, August 1962.

24. R. Bechmann, A. Ballato, and T. Lukaszek, "Higher Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha-Quartz," USAELRDL Technical Report 2261, US Army Electronics R&D Lab., Fort Monmouth, NJ, pp. 1-78, Sept. 1963.

25. R. Cook and P. Weissler, "Piezoelectric Constants of Alpha- and Beta-Quartz at Various Temperatures," Phys. Rev., Vol. 80, No. 4, pp. 712-716, November 1956.

26. G. Graham and F. Pereira, "Temperature Variation of the Piezoelectric Effect in Quartz," J. App. Phys., Vol. 42, No. 7, p. 3011, June 1971.

27. A. Kahan, "Elastic Constants of Quartz," RADC Technical Report 82-117, Rome Air Development Center, Hanscom Air Force Base, MA, pp. 26-32, May 1982.

28. A. Kahan, "Elastic Constants of Quartz and Their Temperature Coefficients," Proc. 36th Annual Frequency Control Symposium, U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, pp. 159-169, June 1982.

29. A. Kahan, "Temperature Coefficients of the Elastic Constants of Quartz," RADC Technical Report 82-224, Rome Air Development Center, Hanscom Air Force Base, MA, pp.18-28, August 1982.

30. I. Koga, "On the Temperature Coefficients of Quartz Plates for Long Waves," ENT, Vol. 12, pp. 1-2, 1935.

31. I. Koga, M. Aruga, and Y. Yoshinaka, "Theory of Plane Elastic Waves in a Piezoelectric Crystalline Medium and Determination of the Elastic and Piezoelectric Constants of Quartz," Phys. Rev., Vol. 109, pp. 1467-1473, March 1958.

32. F. Lack, G. Willard, and I. Fair, "Some Improvements in Quartz Crystal Circuit Elements," Bell Syst. Tech. J., Vol. 13, pp. 453-463, 1934.

33. P. Lee and Y. Yong, "Temperature Derivatives of Elastic Stiffnesses Derived From the Frequency-Temperature Behavior of Quartz Plates," Proc. 37th Annual Frequency Control Symposium, IEEE Catalog No. 83CH1957-0, Institute of Electrical and Electronics Engineers, Inc.,345 East 47th Street, New York, NY 10017, pp. 200-207, 1983.

34. P. Lee and Y. Yong, "Temperature Derivatives of Elastic Stiffness Derived From the Frequency-Temperature Behavior of Quartz Plates," J. App. Phys., Vol. 56, No. 5, pp. 1514-1521, September 1984.

35. P. Lee and Y. Yong, "Frequency-Temperature Behavior of Thickness Vibrations of Doubly Rotated Quartz Plates Affected by Plate Dimensions and Orientations," J. App. Phys., Vol. 60, No. 7, pp. 2327-2342, October 1986.

36. A. Lynch, "The Variation With Temperature of the Piezoelectric Coefficients of Quartz," Proc. Phys. Soc. (London), Vol.63, Part 11, pp. 890-892, November 1950.

37. W. Mason, "Low Temperature Coefficient Quartz Crystals," Bell System Tech. J., Vol. 19, pp. 74-93, January 1940.

38. W. Mason, "Zero Temperature Coefficient Quartz-Crystals For Very High Temperatures," Bell Sys. Tech. J., Vol. 30, pp. 366-380, April 1951.

39. A. Mullen, "Temperature Variation of the Piezoelectric Constant of Quartz," J. App. Phys., Vol. 40, pp. 1693-1696, 1969.

40. V. Soroka, "Effects of Impurities on the Temperature Dependence of the Elastic Constants of Quartz," Sov. Phys. -Solid State, Vol. 10, p. 2241, 1969.

41. B. Sinha and H. Tiersten, "First Temperature Derivatives of the Fundamental Elastic Constants of Quartz," J. App. Phys., Vol. 50, No. 4, pp. 2732-2739, April 1979.

42. B. Sinha and H. Tiersten, "Temperature Derivatives of the Fundamental Elastic Constants of Quartz," Proc. 32nd Annual Frequency Control Symposium, U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, pp. 150-154, May-June 1978.

43. IEC Recommendation: "Basic Method for the Measurement of Resonance Frequency and Equivalent Series Resistance of Quartz Crystal Units by Zero Phase Technique in a Pi-Network," Publication 444, International Electrotechnical Commission, Paris, 1973.

44. A. Ballato and J. Vig, "Static and Dynamic Frequency-Temperature Behavior of Singly and Doubly Rotated, Oven-Controlled Quartz Resonators," Proc. 32nd Annual Frequency Control Symposium, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, pp. 180-188, May-June 1978.

45. B. Sinha and H. Tiersten, "Transient Thermally Induced Frequency Excursions in Doubly-Rotated Quartz Thickness-Mode Resonators," Proc. 34th Annual Frequency Control Symposium, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, pp. 393-402, May 1980.

46. G. Théobald, G. Marianneau, R. Prétot, and J.J. Gagnepain, "Dynamic Thermal Behavior of Quartz Resonators," Proc. 33rd Annual Frequency Control Symposium, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, pp. 239-246, May-June 1979.

47. R. Smythe and P. Morley, "Experimental Evaluation of the Effective Non-Linear Elastic Constant for Trapped Energy and Contoured Resonators," Proc. 39th Annual Frequency Control Symposium, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, pp. 317-324, May 1985.

48. G. Théobald and J. Gagnepain, "Frequency Variations in Quartz Crystal Resonators Due to Internal Dissipation," J. App. Phys., Vol. 50, No. 10, pp. 6309-6315, October 1979.

49. H. Tiersten, "Analysis of Non-linear Resonance in Thickness-Shear and Trapped Energy Resonators," J. Acous. Soc. Am., Vol. 59, No. 4, April 1976.

50. A. Ballato and T. Lukaszek, "Frequency-Temperature Coefficients of Mass-Loaded Plate Vibrators," USAECOM Technical Report ECOM-4356, US Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ, 49 pp., September 1975.

51. J. Kosinski, "Temperature Coefficients of the Butterworth-Van Dyke and Transmission Line Equivalent Electrical Networks for Rotated Y-cut Quartz Crystal Resonators," Master's Degree Thesis, Department of Electronic Engineering, Monmouth College, West Long Branch, NJ 07764, 159 pp., April 1988.

52. R. Bechmann and S. Ayers, "Mechanical and Electrical Behavior of Piezoelectric Crystals (Equivalent Electrical Circuit)," in Piezoelectricity. London : Her Majesty's Stationery Office, 1957, pp. 63-70.

53. R. Bechmann, "Schwingkristalle für Siebschaltungen," Archiv der elekr. Ubertragung, Vol. 18, pp. 129-136, 1964.

54. A. Ballato, "Doubly Rotated Thickness Mode Plate Vibrators," in Physical Acoustics (W. Mason and R. Thurston, eds.). New York:Academic Press, Vol. 13, 1977, pp.115-181.

55. A. Ballato, "Piezoelectric Resonators," in Design of Crystal and Other Harmonic Oscillators (B. Parzen). New York:Wiley , 1983, pp. 66-122.

56. A. Ballato, E. Hatch, and T. Lukaszek, "Acoustic Velocities and Time Constants of the Simple Thickness Modes of Quartz Plates," Proc. 1985 Ultrasonics Symposium, pp. 1143-1146, October 1985.

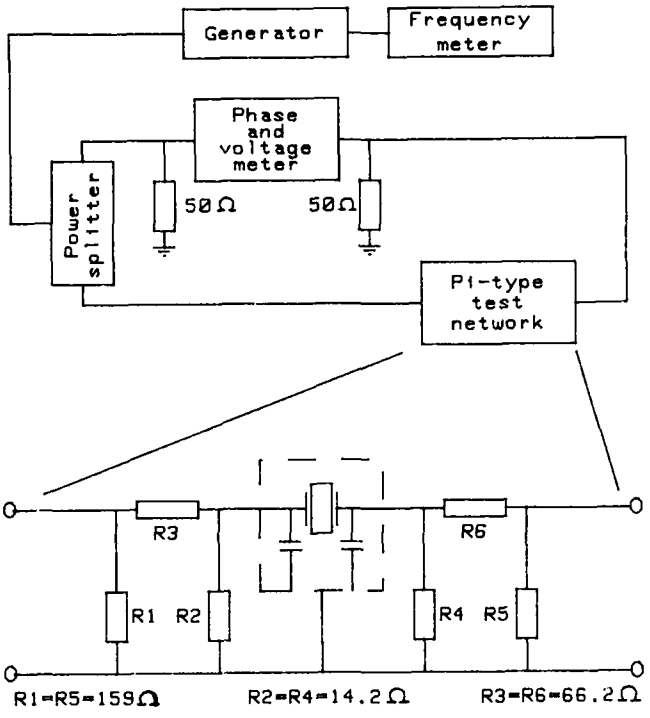



Figure 1. Basic measurement system and pi-type network specified in IEC Recommendation 444.

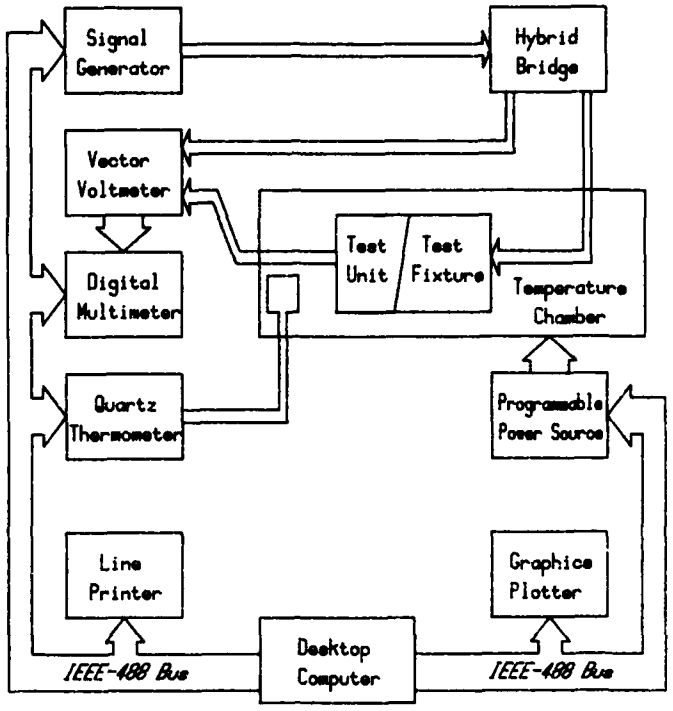



Figure 2. Frequency versus temperature test system as actually implemented.

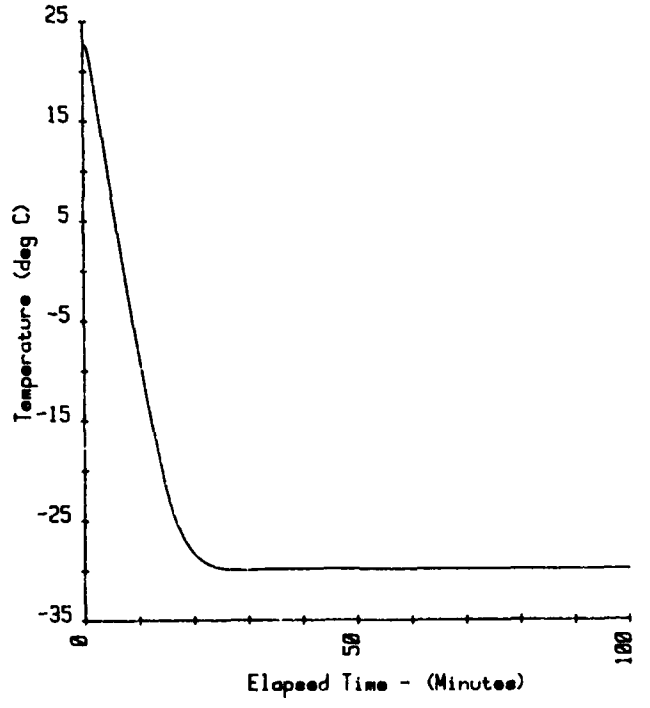



Figure 3. Typical results from a step test used to determine the response time of the Tenney Jr. temperature chamber.

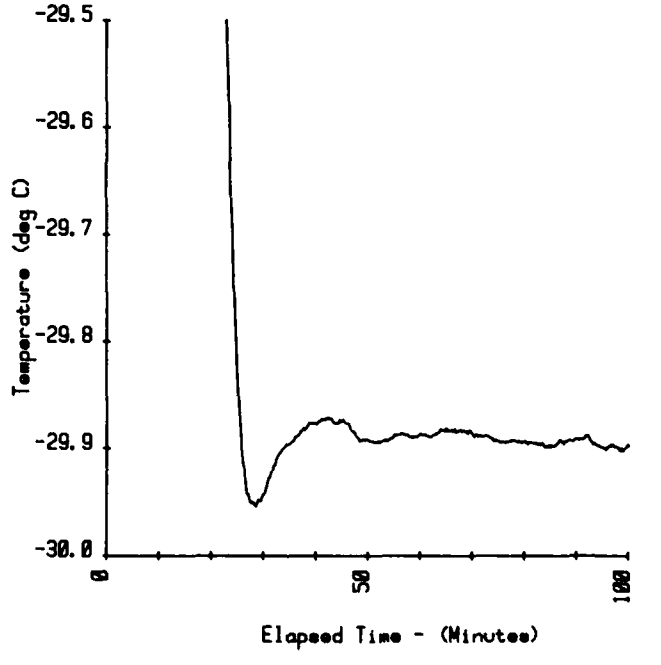



Figure 4. Detail view of step test results showing the residual thermal drift and random fluctuations characterizing the equilibrium condition.

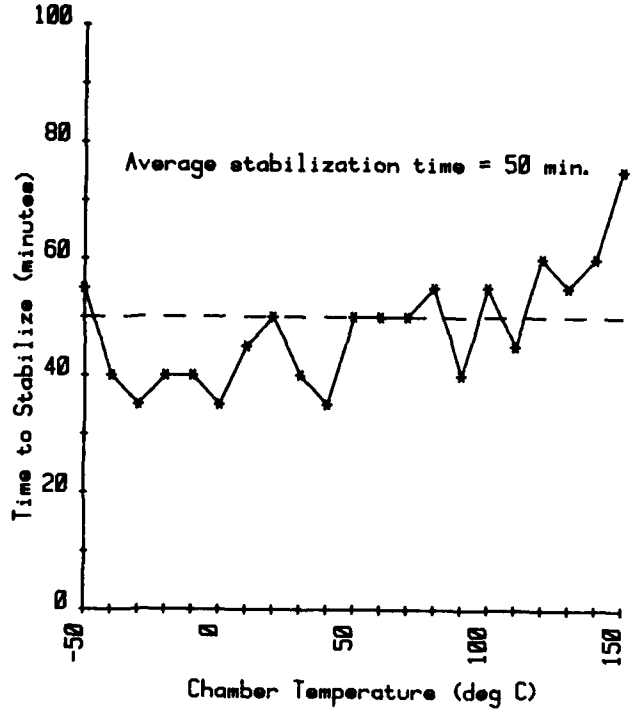



Figure 5. Time required for Tenney Jr. temperature chamber to reach thermal equilibrium as measured over the test system operating range.

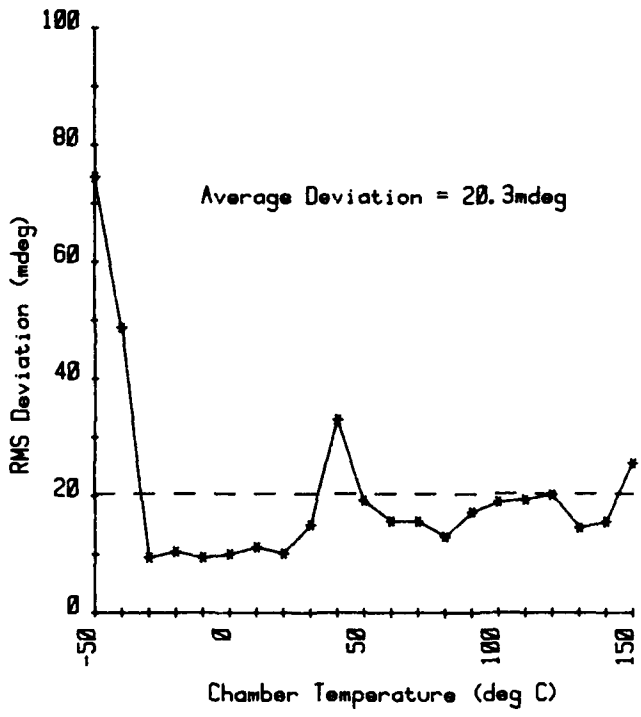



Figure 6. RMS temperature variation within the Tenney Jr. temperature chamber while dwelling at equilibrium as measured over the test system operating range.

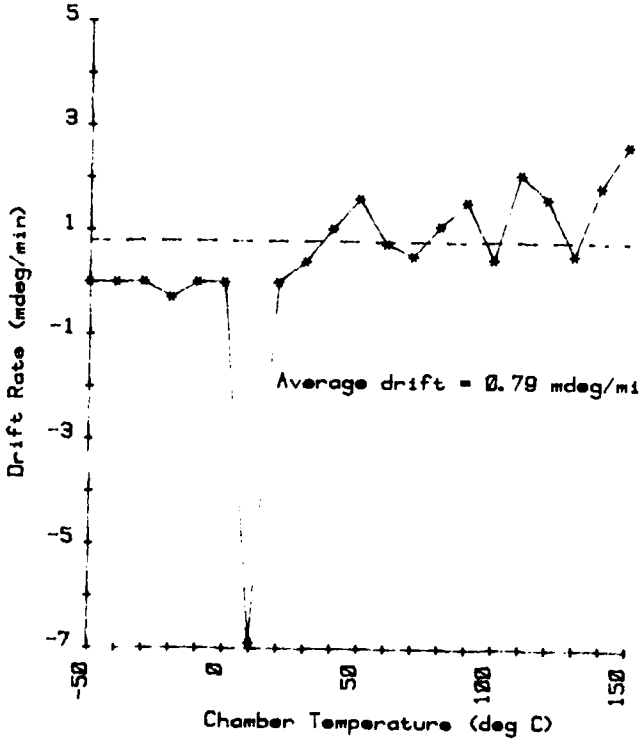



Figure 7. Linear temperature drift within the Tenney Jr. temperature chamber while dwelling at equilibrium as measured over the test system operating range.

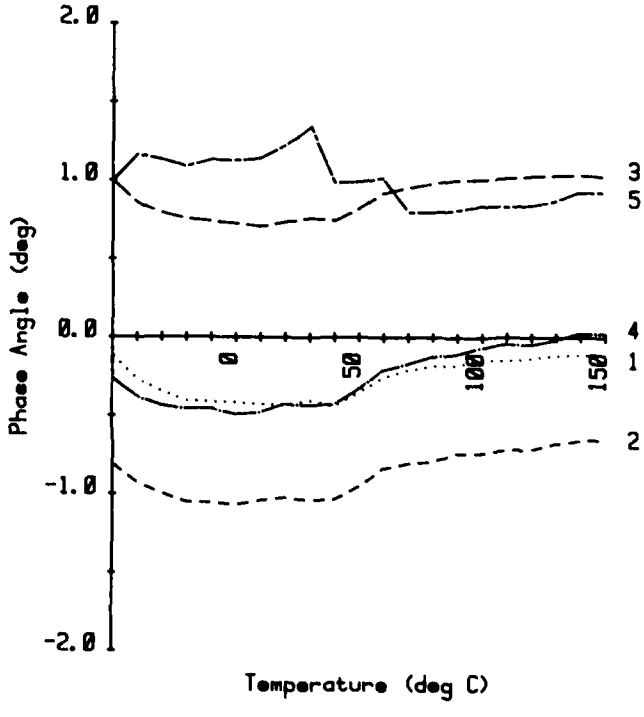



Figure 8. Typical variation with temperature of the zero-phase reference for each of the five pi-networks.

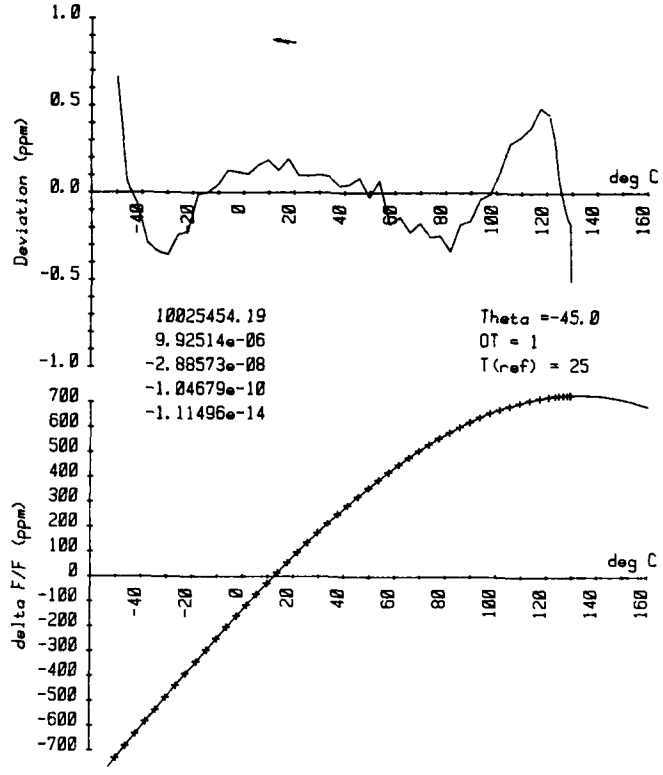



Figure 9. Sample data from a (YXl)-45° resonator operating on the fundamental mode (M=1).

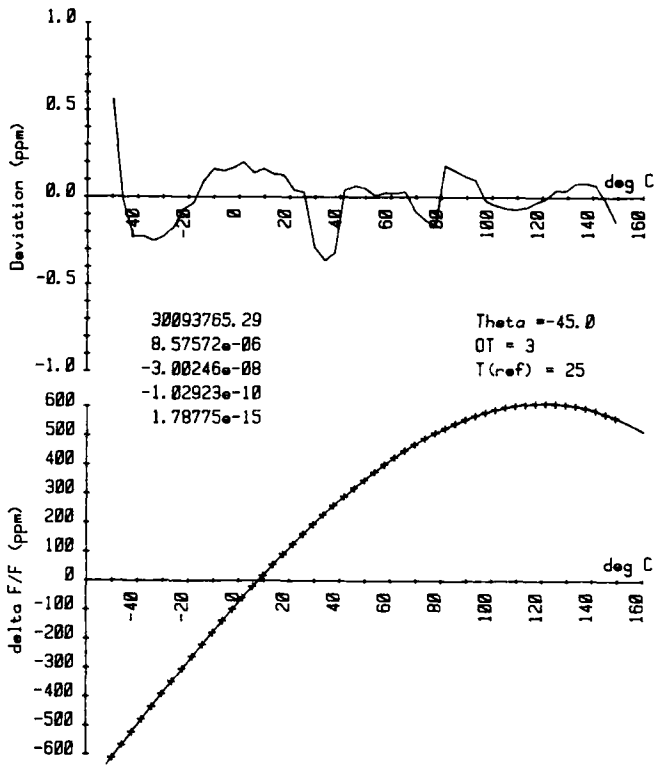



Figure 10. Sample data from a (YXl)-45° resonator operating on the third overtone (M=3).

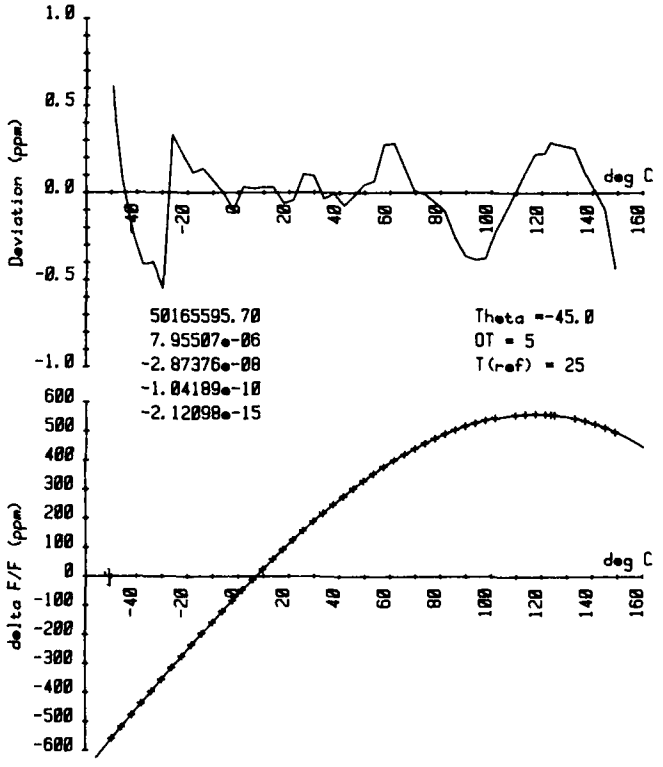



Figure 11. Sample data from a (YXl)-45° resonator operating on the fifth overtone (M=5).

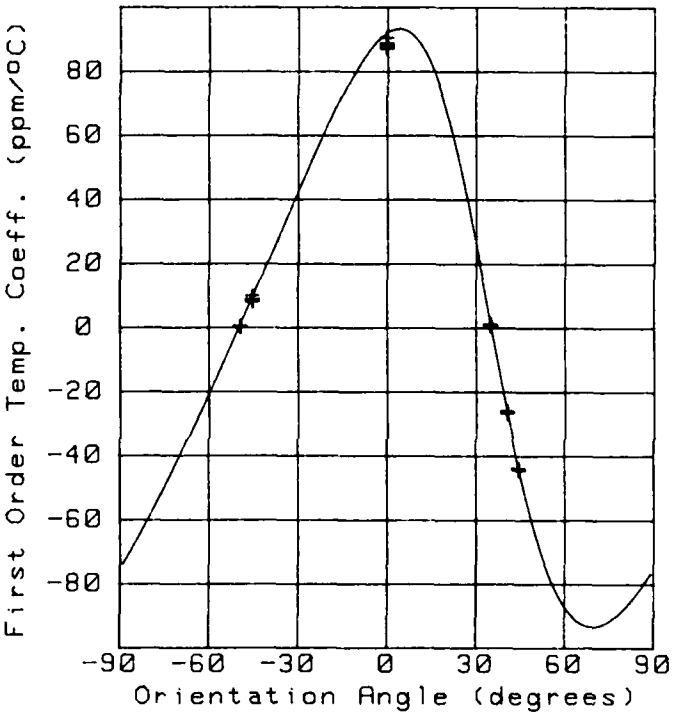



Figure 12. First order temperature coefficient of zero-mass-loading, fundamental, antiresonance frequency $T_{f_a}^{(1)}$ versus orientation angle theta for the rotated Y-cut resonator.

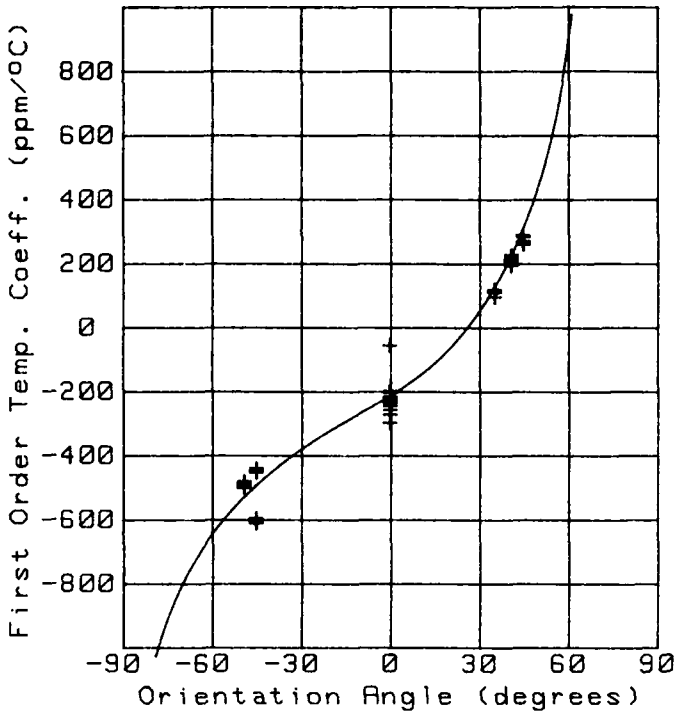



Figure 13. First order temperature coefficient of piezoelectric coupling $T_k^{(1)}$ versus orientation angle theta for the rotated Y-cut resonator.

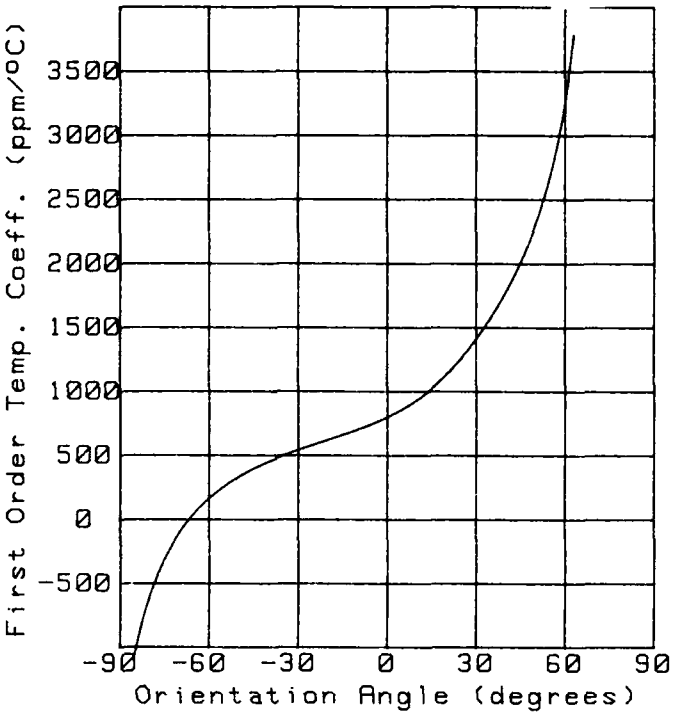



Figure 14. Ideal case, fundamental mode, first order temperature coefficient of dynamic permittivity $T_{\Gamma_1}^{(1)}$ versus orientation angle theta for the rotated Y-cut resonator.

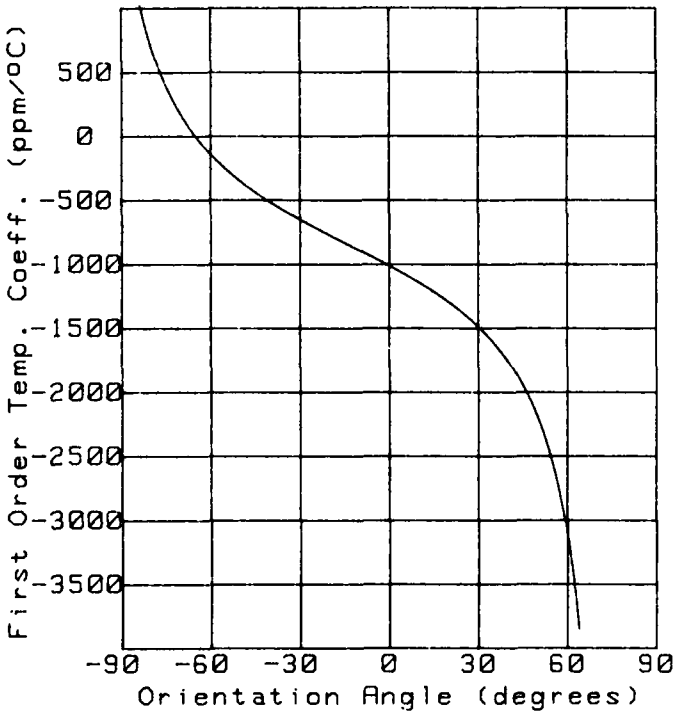



Figure 15. Ideal case, fundamental mode, first order temperature coefficient of dynamic permeability $T_{\Lambda_1}^{(1)}$ versus orientation angle theta for the rotated Y-cut resonator.

42nd Annual Frequency Control Symposium - 1988

# A VARIATIONAL ANALYSIS OF A NEW SHAPE FACE SHEAR MODE QUARTZ CRYSTAL RESONATOR FORMED BY AN ETCHING METHOD

Hirofumi KAWASHIMA, Masaru MATSUYAMA and Mitsuhiro NAKAZATO

SEIKO Electronic Components Ltd.
Hiraicho Tochigi-Shi Tochigi-Ken 328, Japan

## Summary

This paper describes a variational analysis, the frequency temperature characteristics, and the electrical equivalent circuit constants of a face shear mode quartz crystal resonator with two supporting portions at the ends, formed by an etching method. The object of this paper is to propose a new shape face shear mode quartz crystal resonator, and to clarify its frequency temperature characteristics and its electrical equivalent circuit constants in the calculation and in the experiments. A vibration analysis is performed according to the following procedure:

First, a frequency equation is derived by performing a vibration analysis of a face shear mode resonator being incorporated with a vibrational portion and supporting portions with an energy method. From the obtained frequency equation, a relationship of frequency temperature coefficients versus a cut angle or a mass ratio is clarified.

Second, some of the calculated results are compared with the experimental values, as a result, the calculated and experimental results are in good agreement.

Finally, various constants of an equivalent circuit for quartz crystal resonators are examined, so that it is shown that a face shear mode quartz crystal resonator with low series resistance and a high quality factor in the frequency range of 1.5 MHz to 4.2 MHz is obtained, even if the supporting portions are added and connected to the vibrational portion.

## 1. Introduction

A face shear mode quartz crystal resonator exists as a resonator of comperatively low frequency with a zero temperature coefficient by selectio of a cut angle. This quartz crystal resonator is available for consumer products and communication equipment requiring comperatively high precision. Recently, however, as the consumer products and the communication equipment are miniaturized and lightened, the conventional square plate face shear mode quartz crystal resonator formed by a mechanical process is unsatisfactory for them, because of difficulty for miniaturization and support, and also weakness against shock. In addition, the conventional face shear mode quartz crystal resonator is approximately 500 kHz at the highest frequency, therefore, it is impossible to obtain a face shear mode quartz crystal resonator of 1.5 MHz to 4.2 MHz being required at present by using a mechnical process, because of difficulty of the support described above.

In order to overcome the above mentioned drawbacks, the authors propose a new shape face shear mode quartz crystal resonator consisting of a vibrational portion and supporting portions, incorporated by an etching method. From this, the drawbacks of the conventional square plate have been overcome. In analytical procedure, first, the equations of motion of a resonator incorporated with the vibrational portion and the supporting portions are derived from an energy method. From the equations, the frequency equation is calculated by applying Galerkin's method to solve the partial differential equations and using displacement obtained from the experiments of motional inductance L1. Second, resonant frequency, frequency temperature behavior and electrical characteristics of the resonator are analyzed by utilizing the frequency equation and displacement. Finally, these calculated results are compared with the measured values.

## 2. Analysis Procedure

### 2.1 Equations of motion

Figure 1 shows a coordinate system for a face shear mode quartz crystal resonator. This resonator is shown in IRE expression of yxl($\phi$), and CT and DT cut plates manufactured by a mechanical process have cut angles of $\phi = 38°$ and $\phi = -52°$.

Figure 2 shows an actual resonator shape(a) consisting of a vibrational portion and supporting portions and its modified shape(b). The dotted line of Fig.2(a) shows a displacement form of this resonator forecast from that of the conventional square plate. From this, it is possible that the

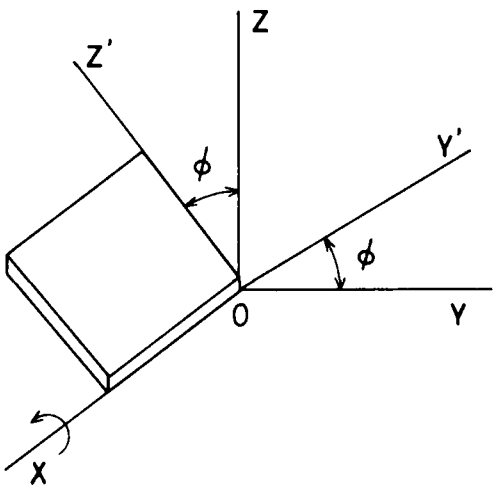



Fig.1 Coordinate system for a face shear mode quartz crystal resonator

CH2588-2/88/0000-065 $1.00 C 1988 IEEE

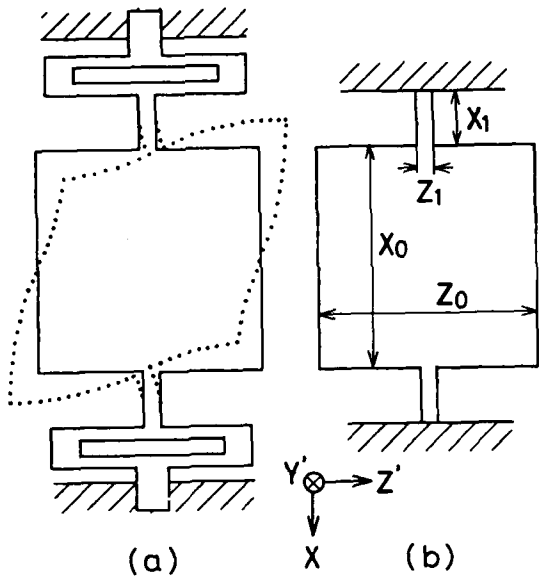



Fig.2 Resonator shape (a) and its modified shape (b)

resonator(a), as shown in Fig.2(b), has the supporting portions vibrating in a flexural mode with each mass m1 at both ends of the vibrational portion and also with the boundary conditions of "clamped end". That is, the supporting portions each have a mass m1 and resistance to a vibration of the vibrational portion ( work done by external force ). Let us take length x0, width z0 and thickness y0 in a dimension of the vibrational portion for this modified shape(b), which correspond to rotated crystal axes x, z' and y' of quartz, respectively. Furthermore, take length x1 and width z1 in a dimension of the supporting portions and the center of the vibrational portion to be the original point o. Now, when taking kinetic energy K1, K2 of the vibrational portion and one supporting, potential energy U1 of the vibrational portion, and work W2 done by external force of the supporting portions, the total kinetic energy K and the total potential energy U are as follows:

$$K = K_1 + K_2\Big|_{+x_0/2} + K_2\Big|_{-x_0/2} \tag{1}$$

$$U = U_1 \tag{2}$$

Therefore, the Lagrangian L is obtained:

$$L = K - U \tag{3}$$

In addition, since resistance caused by the supporting portions can be considered as external force operated on the vibrational portion, the work $\hat{W}$ done by the external force is expressed as follows:

$$\hat{W} = W_2\Big|_{+x_0/2} + W_2\Big|_{-x_0/2} \tag{4}$$

Therefore, when taking time t, from the variational theorem,

$$\delta\int_{t_0}^{t_1} L\,dt + \int_{t_0}^{t_1} \delta\hat{W}\,dt = 0 \tag{5}$$

The kinetic energy K1 of the vibrational portion is expressed as

$$K_1 = \int_{-z_0/2}^{z_0/2}\int_{0}^{y_0}\int_{-x_0/2}^{x_0/2} \frac{1}{2}\rho\left\{\left(\frac{\partial u}{\partial t}\right)^2 + \left(\frac{\partial v}{\partial t}\right)^2 + \left(\frac{\partial w}{\partial t}\right)^2\right\} dx\,dy\,dz \tag{6}$$

where $\rho$ is density of quartz, u, v and w are the displacement in the x, y and z axis directions, respectively. The variation of eq.(6) integrated in time is performed and integrated by parts:

$$\delta\int_{t_0}^{t_1} K_1\,dt = -\int_{t_0}^{t_1} dt\left[\int_{V_1}\rho\left(\frac{\partial^2 u}{\partial t^2}\delta u + \frac{\partial^2 v}{\partial t^2}\delta v + \frac{\partial^2 w}{\partial t^2}\delta w\right)dv\right] \tag{7}$$

where $\int_{-z_0/2}^{z_0/2}\int_{0}^{y_0}\int_{-x_0/2}^{x_0/2} \equiv \int_{V_1}$, $dx\,dy\,dz = dV$

Furthermore, the kinetic energy K2 of one supporting portion is given as

$$K_2 = \frac{1}{2} m_1\left(\frac{\partial w'}{\partial t}\right)^2 \tag{8}$$

where m1 denotes a mass of the supporting portion and w' is the displacement of the supporting portion. Similarly to eq.(6), the variation of eq.(8) is integrated by parts as follows:

$$\delta\int_{t_0}^{t_1} K_2\Big|_{\pm x_0/2} dt = -m_1\int_{t_0}^{t_1} dt\,\frac{\partial^2 w'}{\partial t^2}\,\delta w'\Big|_{\pm x_0/2} \tag{9}$$

Next, consider the potential energy U1 of the vibrational portion. Now, taking stress Ti and strain Si (i=1~6), from the piezoelectric nature and x0, z0 ≫ y0 in a dimension,

$$T_1 = T_2 = T_3 = T_4 = 0 \tag{10}$$

Therefore, the potential energy U1 is calculated from the following equation:

$$U_1 = \int_{V_1}\frac{1}{2}\left(T_5 S_5 + T_6 S_6\right) dV \tag{11}$$

When taking elastic stiffness constants c55, c56, c65, c66, a relationship between stress T5, T6 and strain S5, S6 is given as follows:

$$\begin{aligned} T_5 &= c_{55} S_5 + c_{56} S_6 \\ T_6 &= c_{65} S_5 + c_{66} S_6 \end{aligned} \tag{12}$$

Accordingly, the variation of eq.(11) is given as follows by using eq.(12) [1]:

$$\begin{aligned} \delta\int_{t_0}^{t_1} U_1\,dt = \int_{t_0}^{t_1} dt\,\Big[ &\int_S \left[\left\{T_6\cos(y,n) + T_5\cos(z,n)\right\}\delta u\right. \\ &+ T_6\cos(x,n)\,\delta v \\ &\left. + T_5\cos(x,n)\,\delta w\right] ds \\ &- \int_{V_1}\left\{\left(\frac{\partial T_6}{\partial y} + \frac{\partial T_5}{\partial z}\right)\delta u + \frac{\partial T_6}{\partial x}\delta v\right. \\ &\left. + \frac{\partial T_5}{\partial x}\delta w\right\} dV\Big] \end{aligned} \tag{13}$$

Next, the work W2 done by external force of the supporting portion is given by the following relationship [1]:

$$W_2 = \int_{\pm x_{0/2}}^{\pm(x_{0/2}+x_1)} \frac{1}{2} E I \left(\frac{\partial^2 w'}{\partial x^2}\right)^2 dx \qquad (14)$$

where E is Young's modulus, and I is the moment of inertia. Therefore, the variation of eq.(14) is calculated with the same method mentioned above and obtained as follows:

$$\int_{t_0}^{t_1} \delta W_2 \, dt = \int_{t_0}^{t_1} dt \{ E I \left(\frac{\partial^2 w'}{\partial x^2}\right) \delta \left(\frac{\partial w'}{\partial x}\right) \Big|_{\pm x_{0/2}}^{\pm(x_{0/2}+x_1)} - E I \left(\frac{\partial^3 w'}{\partial x^3}\right) \delta w' \Big|_{\pm x_{0/2}}^{\pm(x_{0/2}+x_1)} + \int_{\pm x_{0/2}}^{\pm(x_{0/2}+x_1)} E I \left(\frac{\partial^4 w'}{\partial x^4}\right) \delta w' \, dx \} \qquad (15)$$

Furthermore, since the operating portion can be considered with the boundary conditions of " one clamped end " at one end, from eq.(15):

$$\frac{\partial w'}{\partial x} = 0, \quad w' = 0 \quad ; \; x = \pm(x_{0/2}+x_1) \qquad (16)$$

Therefore,

$$w' = w_0 \sin px \cos\omega' t \Big|_{\pm x} \quad (|x| \geqq x_{0/2}) \qquad (17)$$

where w0 is amplitude, p is a constant which is obtained from the boundary conditions, and $\omega'$ is angular frequency. From the relationship of eq.(19), p is given as

$$p = \frac{p'\pi}{x_0 + 2x_1} \qquad (p' = 1,3,5,\cdots) \qquad (18)$$

In addition, the external force $F|_{\pm x0/2}$ is calculated from eq.(15) using eq.(17):

$$F\Big|_{\pm x_{0/2}} = \mp \int_{\pm x_{0/2}}^{\pm(x_{0/2}+x_1)} E I \left(\frac{\partial^4 w'}{\partial x^4}\right) dx \qquad (19)$$

$$= \mp w_0 E I \left(\frac{p'\pi}{x_0+2x_1}\right)^3 \cos\frac{x_0}{2(x_0+2x_1)} p'\pi \cos\omega' t \qquad (20)$$

The force given in eq.(20) is the force at $x = \pm x0/2$ and $|z| \leqq z1 \fallingdotseq 0$ (z1 is the width in the z axis direction at the connected portion of the vibrational portion and the supporting portion), and it operates on the vibrational portion in the z axis direction as resistance. Therefore, a distribution of force applied to the vibrational portion is given by $F|_{\pm x0/2} \sin q|x|$ (q is a constant which is determined by the boundary conditions). Furthermore, since the force operating at $\pm x0/2$ is the same, this force can be considered as the force per unit volume $\bar{F}\pm$ operating on the vibrational portion. Accordingly, the following relationship is obtained:

$$\bar{F}_{\pm} = \frac{2 F\Big|_{\pm x_{0/2}} \sin q|x|}{V_1} \qquad (21)$$

Substituting eqs.(7), (9), (13) and (15) into eq.(5), since $\delta u$, $\delta v$, $\delta w$ and $\delta w'$ are, respectively, arbitrary under the condition that the vibrational portion and the supporting portions are connected at $x = \pm x0/2$ and $|z| \leqq z1 \fallingdotseq 0$, the equations of motion are given as follows:

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial T_6}{\partial y} + \frac{\partial T_5}{\partial z}$$

$$\rho \frac{\partial^2 v}{\partial t^2} = \frac{\partial T_6}{\partial x} \qquad (22)$$

$$\rho \frac{\partial^2 w}{\partial t^2} \pm \bar{F}_{\pm} = \frac{\partial T_5}{\partial x}$$

In addition, the boundary conditions are obtained quite similarly from eq.(5):

$$\int_S T_5 \, dS = \mp m_1 \frac{\partial^2 w'}{\partial t^2}\Big|_{\pm x_{0/2}} \; ; \; x = \pm x_{0/2}, \; |z| \leqq z_1 \approx 0$$

$$T_5 = 0 \; ; \; z = \pm z_{0/2} \qquad (23)$$

$$T_6 = 0 \; ; \; x = \pm x_{0/2}, \; y = 0, y_0$$

## 2.2 Displacement form determined by motional inductance L1

The displacement u and w can be written as follows, including a time expression:

$$u = U(x,z) \cos\omega t$$
$$w = W(x,z) \cos\omega t \qquad (24)$$

Furthermore, although a distribution of displacement for the vibrational portion connected to the supporting portions added is very complicated, providing assumption that the displacement U and W have the form:

$$U = A_1 (1 - a_1 \cos\xi_1 x) \sin\eta z$$
$$W = A_2 (1 - a_2 \cos\xi_2 z) \sin\zeta x \qquad (25)$$

where $\eta$ and $\zeta$ are constants which are determined by the boundary conditions of eq.(23) and they have

$$\eta = \frac{m\pi}{z_0} \qquad (m = 1,3,5,\cdots)$$

$$\zeta = \frac{n\pi}{x_0} \qquad (26)$$

n is a solution of $\cos(n\pi/2) = \mu H n$, $\mu = m1/(M1/2)$, H is given by $2\sin\{p'\pi x0/(x0+2x1)\}$ ($\because x0 = z0$). In more detail, $\mu$ shows a ratio of a mass m1 of a supporting portion to a half mass (M1/2) of a vibrational portion, namely, a mass ratio. We can take $\xi 1$ and $\xi 2$ as the following form:

$$\xi_1 = \frac{\pi}{x_0}$$

$$\xi_2 = \frac{\pi}{z_0} \qquad (27)$$

Now, let us calculate constants a1 and a2 from motional inductance L1. The motional inductance L1 is calculated by taking electromagnetic energy equal to kinetic energy. In this resonator, taking into consideration that a mass of a supporting portion is

extremely small as compared with that of a vibrational portion, L1 is given as

$$L_1 = \frac{\rho \int_{V_1} (U^2+W^2)\, dV}{\{ \int_{Se} \varepsilon'_{25} (\frac{\partial U}{\partial z} + \frac{\partial W}{\partial x}) dS \}^2} \qquad (28)$$

where $V_1$: Volume of vibrational portion
S: Area of vibrational portion
Se: Electrode area of vibrational portion
$\rho$: Density of quartz
$\varepsilon 25'$: Piezoelectric constant

In the displacement U, W of eq.(25), replacing A1(1-a1cos $\xi$ 1x) and A2(1-a2cos $\xi$ 2z) by functions of f(x) and g(z), integration of each function and two powers of each function is, respectively, given as the form of

$$\begin{aligned} \int_{-xe/2}^{xe/2} f^2(x)\, dx &= f_1(x_0)x_0 \\ \int_{-ze/2}^{ze/2} g^2(z)\, dz &= g_1(z_0)z_0 \\ \int_{-xe/2}^{xe/2} f(x)\, dx &= f_2(x_0,xe)x_0/\pi \\ \int_{-ze/2}^{ze/2} g(z)dz &= g_2(z_0,ze)z_0/\pi \end{aligned} \qquad (29)$$

where xe and ze are electrode length in the x axis and z axis directions, respectively, f1(x0), g1(z0), f2(x0,xe) and g2(z0,ze) are constants. By utilizing this relation, L1 of eq.(28) is given as

$$L_1 = \frac{\pi^2}{G} \frac{\rho}{\varepsilon'^2_{25}} \frac{y_0 z_0}{x_0} \qquad (30)$$

where G is a constant, e.g. when electrodes on entire surfaces of a vibrational portion are disposed, namely, xe = x0, ze = z0, the constant G has the form:

$$\frac{1}{G} = \frac{f_1(x_0)+g_1(z_0)}{8\{f_2(x_0)+g_2(z_0)\}} \qquad (31)$$

G is easily calculated, when functions f(x) and g(z) are known. In this case, however, since G is unknown, G is obtained from the experiments of motional inductance L1, e.g. when a quartz crystal resonator has dimensions of x0 = z0 = 1.65 mm and y0 = 50 $\mu$m ( f = 1.84 MHz) with a cut angle of $\phi = 37°$, G of eq.(29) has a value of 106.7. Therefore, a constant a ( = a1 = a2) have a value of 0.925, by taking a1 = a2 and A1 = A2 into account, because the quartz crystal resonator is a square plate. Thus, displacement U, W satisfying the boundary conditions from motional inductance obtained by the experiments, is given.

## 2.3 Frequency equation

Let us solve the equations of motion of eq.(22), by using the displacement U, W satisfying the boundary conditions. Since length x0 and width z0 of a face shear mode quartz crystal resonator are much larger than thickness y0, stress T6 approximately becomes zero. In addition, as the vibrational portion and the supporting portions are connected at $x = \pm x0/2$ and $|z| \leqq z1 \fallingdotseq 0$, $\omega$ equals $\omega'$ from eqs.(17) and (24). Furthermore, $\bar{F}\pm$ of eq.(21) is given as follows from eq.(20):

$$\bar{F}\pm = \pm \bar{F}_0 \cos\omega t \qquad (32)$$

where $\bar{F}0$ has the following:

$$\bar{F}_0 = -\frac{2w_0}{V_1} E\, I\, (\frac{p'\pi}{x_0+2x_1})^3 (\cos \frac{x_0}{2(x_0+2x_1)} p'\pi) \sin q|x|$$

The resistance of eq.(32) shows shear elongation of the vibrational portion produced by force $\bar{F}\pm$, By taking the shear elongation quantity $\bar{w}$, the $\bar{w}$ can be expressed as follows:

$$\bar{w} = \bar{W}(x) \cos\omega t$$

where $\bar{W}(x) = \bar{F}0/\rho\omega/\omega$ [1]

Therefore, the resultant displacement w is given as follows, taking the resistance(very weak)into account:

$$w = (W+\bar{W}) \cos\omega t \qquad (33)$$

From eqs.(12), (24) and (33), and taking into consideration T6 = 0, Equation(22) is transformed into the following form:

$$\left.\begin{aligned} -\rho\omega^2 U &= \gamma_{55}(\frac{\partial^2 W}{\partial z \partial x} + \frac{\partial^2 U}{\partial z^2}) \\ -\rho\omega^2(W+\bar{W}) + \bar{F}_0 &= \gamma_{55}(\frac{\partial^2 W}{\partial x^2} + \frac{\partial^2 U}{\partial x \partial z}) \end{aligned}\right\} \qquad (34)$$

where $\gamma_{55} = c_{55} - c_{56}^2/c_{66}$ $(\because c_{56} = c_{65})$

Next, solve eq.(34). Since eq.(34) is not easily solved, an approximate solution by means of Galerkin's method is performed; that is, by multiplying the first term of eq.(34) by the function U satisfying the boundary conditions and the second term by W as well, and by performing integration in the area $(-x0/2 \leqq x \leqq +x0/2,\ -z0/2 \leqq z \leqq +z0/2)$. The calculated result gives the following simultaneous equations:

$$\left.\begin{aligned} (\rho\omega^2 - \gamma_{55}\eta^2)A_1 + \gamma_{55}\xi_2\zeta\Phi_{12}A_2 &= 0 \\ \gamma_{55}\xi_1\eta\Phi_{21}A_1 + (\rho\omega^2 - \gamma_{55}\zeta^2)A_2 &= 0 \end{aligned}\right\} \qquad (35)$$

where product $\Phi$ of $\Phi$12 and $\Phi$21 has a value of 0.4147 when the constant a is 0.925. In addition, since $A1 \neq 0$ and $A2 \neq 0$ in eq.(35), the determinant of eq.(35) must be zero so as to obtain a nontrivial solution. Namely, the following relation is obtained:

$$\begin{vmatrix} \rho\omega^2 - \gamma_{55}\eta^2 & \gamma_{55}\xi_2\zeta\Phi_{12} \\ \gamma_{55}\xi_1\eta\Phi_{21} & \rho\omega^2 - \gamma_{55}\zeta^2 \end{vmatrix} = 0 \qquad (36)$$

From the forementioned relation, and by performing Taylor's series under the condition of

$$(\eta^2+\zeta^2)^2 > 4(\eta^2\zeta^2 - \xi_1\xi_2\eta\zeta\Phi),$$

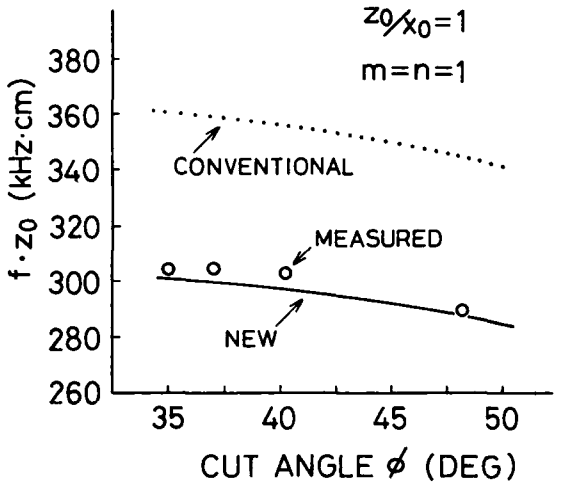



Fig.3 Relationship between frequency constant (f · z0) and cut angle $\phi$

the frequency f is given as

$$f=\frac{1}{2}\sqrt{\frac{\gamma_{55}}{\rho}}\sqrt{\left(\frac{m}{z_0}\right)^2+\left(\frac{n}{x_0}\right)^2}\sqrt{1-\frac{1}{4}k-\frac{1}{16}k^2-\frac{1}{32}k^3} \quad (37)$$

where $k=\frac{4mn(mn-\Phi)x_0^2z_0^2}{(m^2x_0^2+n^2z_0^2)^2}$

## 3. Results and Discussion

Let us show the calculated and measured results, and they are discussed. First, a relationship of a frequency constant versus a cut angle $\phi$ and a mass ratio $\mu$ is shown from eq.(37). Second, a relationship between a piezoelectric constant $\varepsilon 25'$ to excite a face shear mode quartz crystal resonator and a cut angle $\phi$ is shown, which is one of most important factors in designing the resonator. Third, a relationship of frequency temperature coefficients $\alpha$, $\beta$ to a mass ratio $\mu$ and a cut angle $\phi$ is shown and from this, typical frequency temperature characteristics are shown. Next, relationships of motional inductance L1 versus thickness y0 and resonant frequency equivalent to area of a vibrational portion, and of capacitance ratio r, series resistance R1 and quality factor versus resonant frequency are shown. Finally, typical electrical equivalent circuit constants are shown. A detailed description is given below.

### 3.1 Frequency constant

Figure 3 shows a relationship between a frequency constant (f · z0) and a cut angle $\phi$ of a face shear mode quartz crystal resonator. The solid line shows the newly calculated values and the circules show the measured values, while the dotted line shows the conventional result. As is apparent from Fig.3, according to increase of the cut angle $\phi$, the frequency constant (f · z0) decreases. As an example, when a quartz crystal resonator has a cut angle of $\phi=37°$ with so-called " a dimensional ratio z0/x0 = 1 "of width z0 and length x0 of a vibrational portion, a frequency constant (f · z0) becomes 300 kHz · cm and 305 kHz · cm in the calculation and in the experiments.

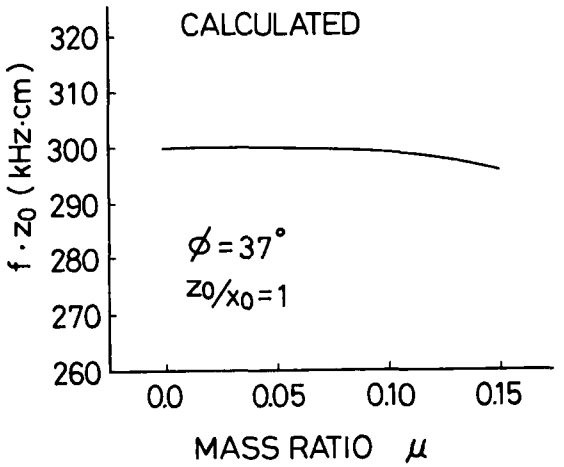



Fig.4 Change of frequency constant (f · z0) to mass ratio $\mu$

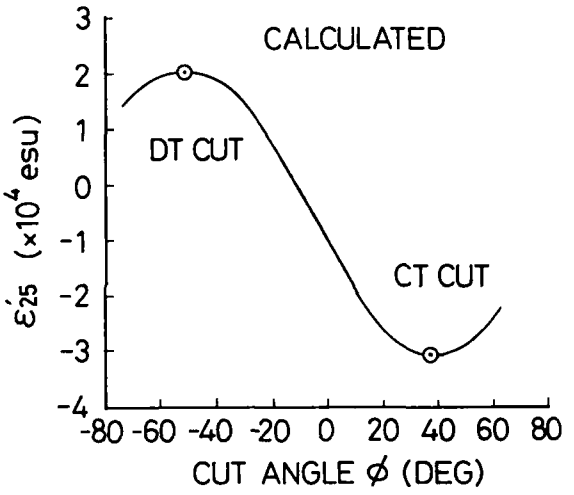



Fig.5 Relationship between piezoelectric constant $\varepsilon 25'$ and cut angle $\phi$

Thus, the calculated result is in good agreement with the measured values with an error of less than three percent between them, as compared with the conventional result with the error of more than fifteen percent.

Figure 4 shows a relationship between the frequency constant (f · z0) and a mass ratio $\mu$ of a face shear mode quartz crystal resonator having the cut angle of $\phi=37°$ and the dimensional ratio of z0/x0 = 1. It is understood from Fig.4 that a variation of the frequency constant (f · z0) is negligibly small in the range of less than the mass ratio of $\mu=0.10$. Therefore, this face shear mode quartz crystal resonator consisting of the vibrational portion and the supporting portions can be considered as the mass ratio $\mu=0$ approximately, because the resonator is such designed as the mass ratio less than $\mu=0.025$.

### 3.2 Piezoelectric constant $\varepsilon 25'$

A piezoelectric constant $\varepsilon 25'$ is one of important factors by which electrical equivalent circuit constants of a face shear mode quartz crystal resonator are determined.

Figure 5 shows a relationship between the cut

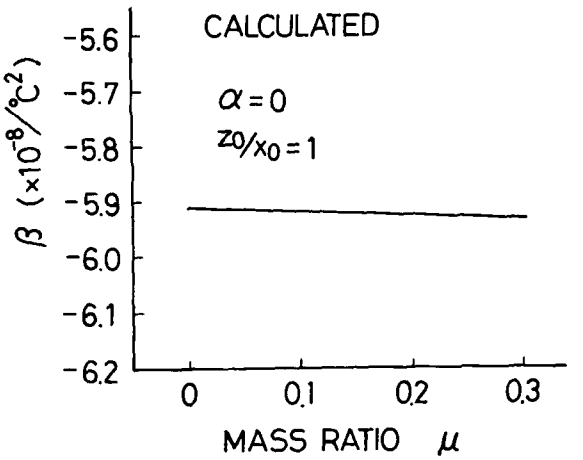



Fig.6 Relationship between mass ratio $\mu$ and the second order temperature coefficient $\beta$ when $\alpha = 0$

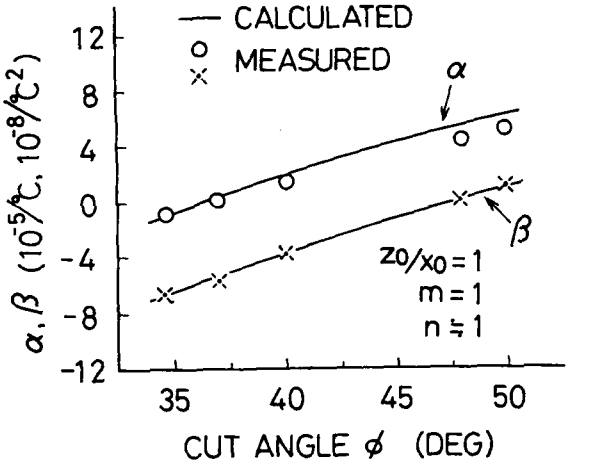



Fig.8 Relationship of the first, second order temperature coefficients $\alpha$, $\beta$ and cut angle $\phi$

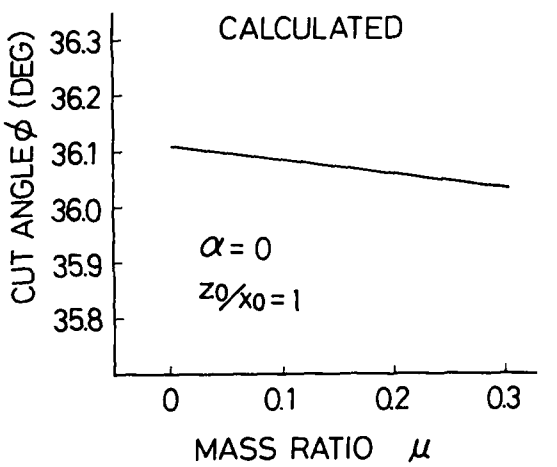



Fig.7 Relationship between mass ratio $\mu$ and cut angle $\phi$ when $\alpha = 0$

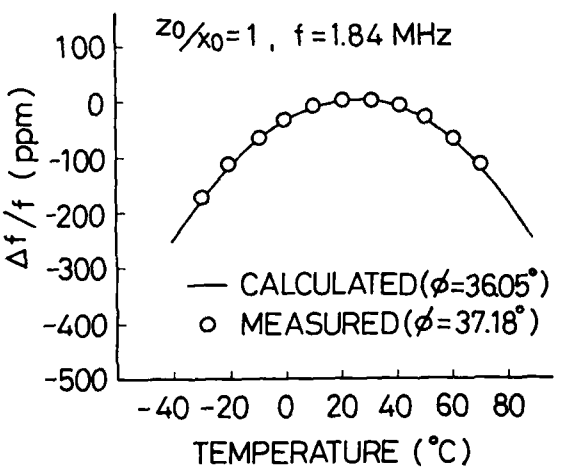



Fig.9 Frequency temperature characteristics

angle $\phi$ and the piezoelectric constant $\varepsilon 25'$. $\varepsilon 25'$ versus the cut angle $\phi$ exhibits the maximum value at DT cut side and the minimum value at CT cut side, the absolute value becomes larger at CT cut side, namely, it is understood that piezoelectric efficiency is more excellent at CT cut side. In addition, it is an interesting result that $\varepsilon 25'$ has the maximum and minimum values in the vicinity of the cut angle $\phi$ giving zero temperature coefficients, even if coincidence may exist.

### 3.3 Frequency temperature coefficients $\alpha$, $\beta$

Figure 6 shows a relationship of the mass ratio $\mu$ and the second order temperature coefficient $\beta$ of a face shear mode resonator with the first order temperature coefficient $\alpha = 0$ and the dimensional ratio of $z0/x0 = 1$. As the mass ratio $\mu$ becomes large, the absolute value of $\beta$ also becomes somewhat large. However, the variational value is so small as to neglect it. Thus, it is understood that adding the supporting portion; hardly influences the frequency temperature characteristics.

Figure 7 shows a relationship between the mass ratio $\mu$ and the cut angle $\phi$ when $\alpha = 0$ and $z0/x0 = 1$. As the mass ratio $\mu$ gets large, the cut angle $\phi$ giving $\alpha = 0$ , gets inversely small. Since the variational quantity is extremly small, it can be considered as within the error of processing accuracy in the range of $\mu < 0.3$.

Figure 8 shows a relationship between the first-order and second-order temperature coefficients $\alpha$, $\beta$ and the cut angle $\phi$ of a face shear mode quartz crystal resonator with dimensions of $x0 = z0 = 1.65$ mm (f = 1.84 MHz) and $\mu = 0.0247$ ($n \fallingdotseq 1$, $m = 1$). As the cut angle $\phi$ increases, $\alpha$ and $\beta$ also increase. The cut angle $\phi$ where gives $\alpha = 0$, is 36.05° in the calculation, and 37.18° in the experiments, and $\beta$ has the values of -0.059 ppm/°C$^2$ and -0.057 ppm/°C$^2$ at the cut angles of $\phi = 36.05°$ and $\phi = 37.18°$, respectively. The calculated values of $\beta$ agree well with the experimental values. On the other hand, concerning $\beta$, the cut angle $\phi$ where gives $\beta = 0$, is approximately 48° in the calculation and in the experiments, and $\alpha$ has the values of 55.2 ppm/°C and 42.5 ppm/°C at each cut angle. The difference of about twenty percent for $\alpha$ between the calculation and the experiments occurs.

### 3.4 Frequency temperature characteristics

Figure 9 shows the frequency temperature characteristics of a face shear mode quartz crystal

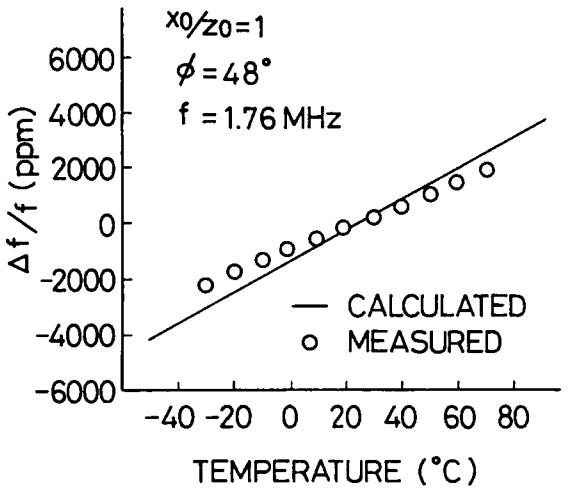



Fig.10 Frequency temperature characteristics

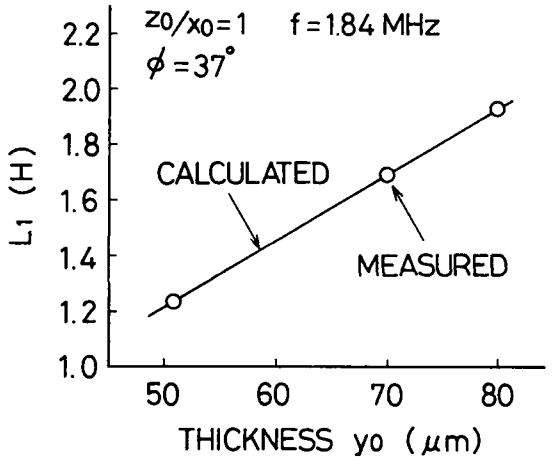



Fig.11 Relationship between motional inductance L1 and thickness y0

resonator with the dimensions of x0 = z0 = 1.65 mm and y0 = 52 μm (f = 1.84 MHz). The circles show the measured values at the cut angle of φ = 37.18°, while the solid line shows the calculated values at φ = 36.05°. A turn over temperature point has approximately 23℃ at both cut angles. Furthermore, the second-order temperature coefficient β has the values of -0.057 ppm/℃$^2$ in the experiments and -0.059 /℃$^2$ in the calculation.

Figure 10 shows another frequency temperature characteristics of a resonator with the same dimensions as those in Fig.9 when the cut angle φ = 48°. β almost reaches zero at this cut angle. In addition, α has the values of 55.4 ppm/℃ in the calculation and 42.5 ppm/℃ in the experiments. This resonator is available for a temperature sensor, because it has high linearity over the wide temperature range of -50 ℃ to +90 ℃.

### 3.5 Motional inductance L1

Figure 11 shows a relationship between motional inductance L1 and thickness y0 of a face shear mode quartz crystal resonator which has dimensions of width z0 = 1.65 mm and length x0 = 1.65 mm (f = 1.84 MHz), and a cut angle of φ = 37°. The constant G of

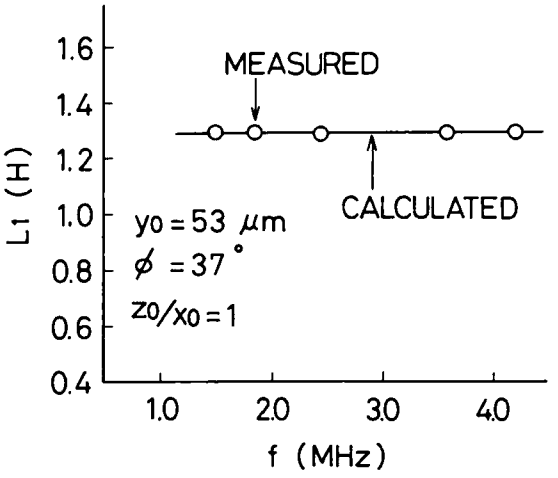



Fig.12 Relationship of motional inductance L1 to resonant frequency f

eq.(30) is obtained from the experimental value of a resonator with the above-mentioned resonator's dimension by utilizing thickness of y0 = 52 μm. Therefore, the solid line showing the calculated results are obtained by using the G derived from the experimental values, taking thickness y0 as a parameter. Judging from good agreement between the calculation and the experiments at thickness of y0 = 70 μm and 80 μm, it is conceivable that the constant G probably has the same value at arbitrary thickness. Equation(30) shows that motional inductance L1 is proportional to thickness y0, whose result agrees well with the measured values.

Figure 12 shows a relationship of motional inductance L1 to a resonant frequency f of a quartz crystal resonator with thickness of y0 = 53 μm, a cut angle of φ = 37° and a dimensional ratio of z0/x0 = 1. It is understood, as is apparent from eq.(30), that motional inductance L1 has a constant value when changing resonant frequency, because to change the resonant frequency f is equivalent to changing area of the vibrational portion with the dimensional ratio of z0/x0 = 1.

### 3.6 Capacitance ratio r

Figure 13 shows a relationship between a capacitance ratio r and a resonant frequency f of the resonator described in Fig.12. The capacitance ratio r is also independent on the resonant frequency, because it is essentially given as functions of a cut angle φ and electrode area disposed on the vibrational portion [2] and has a value of about r = 350 which is sufficiently satisfactory as a resonator.

### 3.7 Series resistance R1 and quality factor Q

Figure 14 shows a relationship of series resistance R1 and a quality factor Q versus a resonant frequency f of the same resonator described in Fig.12. The resonator is such an excellent resonator that it has values of R1 = 100 Ω ~ 180 Ω and Q = 70000 ~ 260000 in the frequency range of f = 1.5 MHz to f = 4.2 MHz. In addition, it is conceivable that a distribution of the series resistance R1 is caused by influence of the supporting portions.

## 4. Electrical Characteristics

Table 1 shows typical values of electrical

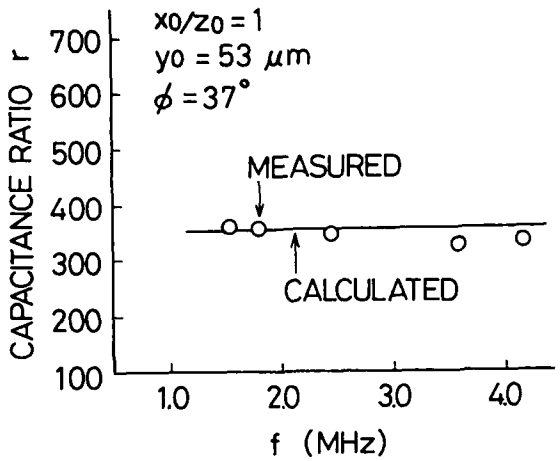



Fig.13 Relationship between capacitance ratio r and resonant frequency f

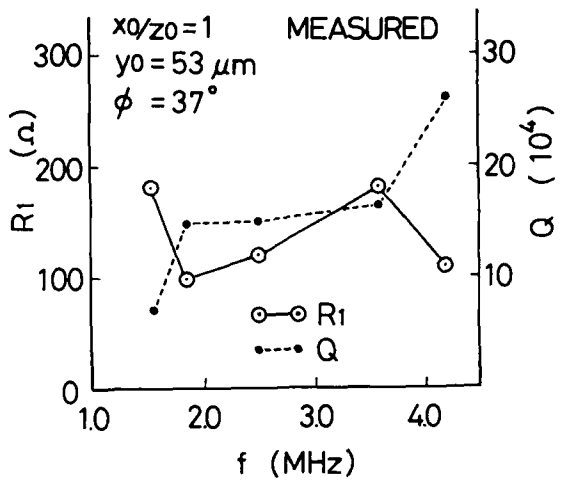



Fig.14 Series resistance R1 and quality factor Q versus resonant frequency f

equivalent circuit constants for face shear mode quartz crystal resonators which have frequencies of f = 1.84 MHz and f = 1.76 MHz, and cut angles of $\phi = 37$ and $\phi = 48°$ . Also, the electrodes are disposed on the entire surfaces of the vibrational portions so as to execute a face shear vibration mode. As is obvious from Table 1, a face shear mode quartz crystal resonator with low series resistance and high quality values can be obtained, even though the supporting portions are added. At the same time, the quartz crystal resonator of this study is excellent even in various electrical characteristics compared with an AT cut quartz crystal resonator of the same frequency.

Table 1 Electrical characteristics

| φ(Deg) | f (MHz) | $L_1$(H) | $R_1$(Ω) | Q ($10^3$) | r |
|---|---|---|---|---|---|
| 37 | 1.84 | 1.29 | 100 | 150 | 350 |
| 48 | 1.76 | 1.12 | 135 | 90 | 330 |

## 5. Conclusions

In this paper, various theoretical and experimental examinations on a variational analysis, frequency temperature characteristics and electrical equivalent circuit constants have been made for a face shear mode quartz crystal resonator with the supporting portions at both ends, formed by an etching method for application to consumer products and communication equipment. In analysis procedure, first, the equations of motion of a resonator incorporated with the vibrational portion and the supporting portions are derived from an energy method. From the equations, the frequency equation is calculated by applying Galerkin's method to solve the partial differential equations. Second, analyzing resonant frequency and frequency temperature characteristics, a relationship of frequency temperature coefficients versus a cut angle and a mass ratio has been clarified. As a result, it is essentially understood that cut angles where give a zero temperature coefficient exist, even if adding the supporting portions such as this resonator, and also good frequency temperature characteristics are obtained. Furthermore, since the second order temperature coefficient reaches zero by selection of a cut angle, it has been shown that this resonator is an optimum resonator as a temperature sensor of quartz. Finally, as a result of examining various electrical characteristics, it has been confirmed that this resonator with the supporting portions at both ends has low crystal impedance and a high quality factor. Thus, this resonator is excellent in various points, and available for a surface mount type, because of ability to design an arbitrary resonator's shape by an etching method.

## References

[1] H.Kawashima:" VARIATIONAL ANALYSIS OF GT CUT QUARTZ CRYSTAL RESONATORS WITH THE SUPPORTING PORTIONS AT THE ENDS ",Proceedings of the 40th Annual symposium on frequency Control, pp 193-200 (1986)

[2] H.Kawashima:" An analysis of electrical equivalent circuit constants of GT cut quartz crystal resonator formed by an etching method ", IECE US87-14 pp51-55 (1987) (in Japanese)

[3] R.Bechmann, A.D.Ballato and T.J.Lukaszek:" High-Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha-Quartz" Proceedings of the IRE August (1962)

42nd Annual Frequency Control Symposium - 1988

# Chemically-Milled UHF SC-Cut Resonators

R. C. Smythe and R. B. Angove
Piezo Technology Inc.
Orlando, Florida 32854

## Abstract

Fundamental mode SC-cut resonators with potential for oscillator and filter applications have been developed at 333 MHz, 500 MHz, and 1 GHz using chemically milling to form a ring-supported structure. A novel electrode configuration which eliminates registration problems is used.

At 333 MHz, Q's are typically 30 x $10^3$. The best Q obtained at this frequency is 34.9 x $10^3$; at 1 GHz, 10.2 x $10^3$ has been measured. Equivalent circuit parameters, frequency-temperature characteristics, nonlinear resonance data, and acceleration sensitivity measurements are reported.

## Introduction

In the fabrication of high frequency resonators, the maximum fundamental frequency is determined by the minimum practical wafer thickness. For wafers fabricated by conventional lapping and polishing techniques, this is typically 30 to 35 microns, corresponding to 50 to 60 MHz for the SC-cut. To overcome this restriction, ring-supported thickness-shear resonators were proposed by Guttwein, Ballato, and Lukaszek [1] and others. Chemical milling [2-6] has been used to fabricate AT-cut [7] and SC-cut [8] ring-supported resonators. In this paper we describe chemically-milled fundamental mode SC-cut resonators at 333 MHz, 500 MHz, and 1 GHz.

Figure 1 shows the resonator configuration. The wafer diameter is 0.25 inch (6.35 mm); its thickness is 53 micrometers, corresponding to a frequency of 34 MHz. A central portion 0.05 inch in diameter is etched to the required thickness, which is about 5.4 micrometers for a resonator frequency of 333 MHz and 1.8 microns for 1 GHz. The double-sided etching process uses Allied Superwet buffered oxide etchant, BOE-1235 [5,9,10]. The crystal blanks are mechanically polished, premium-Q swept quartz. Sweeping was found necessary to reduce the incidence of etch pits and channels [10].

A novel feature of the resonator design is the electrode arrangement, which is simply a pair of crossed strips. From the fabrication standpoint, registration problems are essentially eliminated. Since the electrode width is only a few thousandths of an inch, this is a significant advantage when conventional aperture masks are used in the electrode deposition process. As another advantage, the tab propagation [11] effect allows the use of thicker electrode films than would otherwise be permissible. Furthermore, because of symmetry, odd-order anharmonic modes are well suppressed, and there is no net lateral field excitation of the principal resonance.

The electrodes are aluminum. Nominal electrode widths are 5 x $10^{-3}$ inch for the 333 MHz resonators, and 2.5 x $10^{-3}$ inch for the 500 MHz and 1 GHz units, but measured dimensions are typically 1 x $10^{-3}$ inch larger due to aperture mask shadowing. Standard HC-35 (TO-5 coldweld) enclosures have been used.

## Resonator Performance

Resonators have been fabricated at fundamental frequencies of 333 MHz, 500 MHz, and 1 GHz. In Fig. 2 the best Q's achieved are plotted as a function of frequency. It would be useful to compare these values against the material Q for the SC cut, if this were known. Since it is not, we have made a rough estimate, which is believed to be an upper bound, based on the widely accepted value of the most probable internal friction set forth for the AT cut by Bömmel, Mason, and Warner [12]. This so-called intrinsic Q is described by the function

$$Q_{intrinsic} = Q_i / f,$$

where f is the resonance frequency (MHz) and $Q_i$ is 16 x $10^6$. Bömmel, Mason, and Warner's value was based on measurements of very carefully made resonators at frequencies up to 100 MHz. Similar data are not available for SC-cut resonators, but experience indicates that the intrinsic Q is slightly higher than for the AT-cut. Since in general resonator Q tends to be proportional to volume, we have increased $Q_i$ by the SC/AT thickness ratio, which gives a value of 17.3 x $10^6$. On this basis the estimated intrinsic Q is plotted in Fig. 2. At frequencies up to about 500 MHz we have achieved about 2/3 of this limit.

Typical equivalent circuit parameters are given in Table 1. In this frequency range, electrical losses in the electrodes and tabs, as well as in the leads and mounting ribbons [13] should be included in the equivalent circuit model. We have used the simplest possible representation, Fig. 3, in which a shunt conductance element, Go is added.

Representative spur plots are shown in Fig. 4. For these resonators, the equivalent circuit parameters of the principal mode and the strongest spur are given in Table 2. For oscillator applications, Smith chart plots are frequently helpful. These are also shown in Fig. 4.

CH2588-2/88/0000-073 $1.00 C 1988 IEEE

Table 1. Typical Equivalent Circuit Parameters.

| Freq. (MHz) | 333 | 500 | 950 |
|---|---|---|---|
| $L_1$ (mH) | 0.76 | 0.60 | 0.11 |
| $C_1$ (fF) | 0.30 | 0.14 | 0.25 |
| $R_1$ (ohms) | 53 | 110 | 85 |
| Q x $10^{-3}$ | 30 | 19 | 8 |
| Co (pF) | 0.90 | 0.80 | 1.20 |
| Go (µS) | 65 | 150 | 750 |

Table 2. Equivalent Circuit Parameters of Principal Mode and Strongest Unwanted Mode.

| Unit No. | 74 | 31 | 25 |
|---|---|---|---|
| Principal Mode | | | |
| $f_s$ (MHz) | 333.3 | 507.5 | 954.2 |
| $L_1$ (mH) | 0.80 | 0.69 | 0.10 |
| $R_1$ (ohms) | 48.5 | 106 | 58.4 |
| Q x $10^{-3}$ | 34.7 | 20.5 | 10.1 |
| Co (pF) | 0.84 | 0.78 | 1.19 |
| Go (µS) | 64 | 127 | 764 |
| Unwanted Mode | | | |
| $f_s$ (MHz) | 335.5 | 512.4 | 961.4 |
| $L_1$ (mH) | 4.34 | 1.65 | 0.23 |
| $R_1$ (ohms) | 243 | 346 | 140 |
| Q x $10^{-3}$ | 37.6 | 15.4 | 9.7 |
| Mode Plot | | | |
| Figure | 4a | 4c | 4e |
| Smith Chart Plot | | | |
| Figure | 4b | 4d | 4f |

All parameter measurements were made using a H-P 4191A RF Impedance Analyzer and a one-port coaxial test fixture constructed in accordance with the EIA design. The measurement method has been described earlier [14]. The Smith chart plots were made using the same fixture and a H-P 8753A Network Analyzer with S-Parameter test set. Spur plots were made using the network analyzer and a PTI hybrid network. In accordance with EIA and IEC standards, the effective lead length was 2 mm for all measurements.

Frequency-temperature behavior is about as expected. Figure 5 shows a measured f-T curve for a 333 MHz unit. The theta angle is -33°50' ± 2', which was calculated to give a turnover temperature of about 25°C.

Nonlinear resonance curves for a typical 333 MHz resonator are shown in Fig. 6. The general character of these curves is typical of SC-cut resonators, with the shift in resonance frequency being proportional to the square of the motional arm current and with no observable indication of nonlinear coupling to other modes.

For a group of 333 MHz resonators, the acceleration sensitivity was measured under 150 Hz sinusoidal vibration at 10 g's peak acceleration. The maximum sensitivity in the plane of the wafer and the sensitivity normal to the wafer were measured. The resonators were in TO-5 enclosures with four relatively stiff ribbon mounts. Data for five units are given in Table 3.

Table 3. Dynamic Acceleration Sensitivity, 333 MHz ($f_v$ = 150 Hz, $A_{pk}$ = 10 g's)

| Unit No. | Acceleration Sens. ($10^{-10}$/G) In-Plane | Normal | Resultant |
|---|---|---|---|
| 71 | 6.8 | 4.3 | 8.0 |
| 75 | 5.5 | 3.5 | 6.5 |
| 79 | 6.6 | 3.0 | 7.3 |
| 82 | 5.8 | 2.9 | 6.5 |
| 85 | 12.8 | 1.8 | 12.9 |

## Discussion

These resonators have been developed primarily for use in UHF and microwave frequency generation, where they offer the possibility of eliminating frequency multiplier stages. For many such applications quiescent and vibration-induced phase noise are important performance specifications. For best close-in oscillator noise performance, resonator loaded Q should be maximized. For lowest noise floor, cryst. current should be maximized. In both areas, the SC cut offers advantages over the AT cut. At 333 MHz, a resonator Q of 30 thousand corresponds to a corner frequency of 11 kHz in the Leeson model [15], assuming the ratio of loaded to unloaded Q is 1/2. The maximum crystal current, and hence the noise floor, will be determined by the allowable shift in resonance frequency. For the resonator of Fig. 6, the shift is approximately 1.0 $I^2$ kHz, where is the motional arm current in milliamperes. This shift can be reduced, at the expense of unwanted mode performance, by increasing the electrode area.

In fixed frequency oscillator applications, the primary requirement regarding resonator unwanted modes is that the oscillator consistently start on the principal resonance, and that mode jumping not occur. Typically, it is required that the motional resistance of unwanted modes be 2 to 4 times that of the principal mode. That requirement is met by the resonators developed on this program.

For filter applications, unwanted mode requirements are much more severe. A few resonators have been made at 333 MHz to demonstrate the possibilities. For these units, the nominal electrode width was reduced by a factor of two to 2.5 x $10^{-3}$ inch, resulting in increased unwanted mode frequency offsets and fewer trapped unwanted modes. Figure 7 shows a representative spur plot. The Q of this unit is 27.8 x $10^3$, corresponding to a resistance of 137 ohms, and the motional inductance is 1.8 mH. The motional resistance of the strongest spur is 7600 ohms.

## Conclusions

UHF fundamental mode SC-cut resonators have been developed at frequencies up to 1 GHz with Q-frequency products of $10^7$. Further improvements would result from improved control of parallelism and improved methods of electrode deposition. These would lead to greater uniformity and better unwanted mode performance.

## Acknowledgment

This work was supported by the U.S. Army Laboratory Command, Electronic Technology and Devices Laboratory, under Contract DAAL01-86-C-0006.

## References

[1] G.K. Guttwein, A.D. Ballato, and T.J. Lukaszek, "VHF-UHF Piezoelectric Resonators," U. S. Pat. 3,694,677; 26 September, 1972.

[2] J.R. Vig, J.W. LeBus, and R.L. Filler, "Chemically Polished Quartz," Proceedings of the 31st Annual Symposium on Frequency Control, 1977, pp. 131-143.

[3] J.R. Vig, R.J. Brandmayr, and R.L. Filler, "Etching Studies on Singly and Doubly Rotated Quartz Plates," Proceedings of the 33rd Annual Symposium on Frequency Control, 1979, pp. 351-358.

[4] J.R. Vig and R.J. Brandmayr, U. S. Pat. 4,274,907, June, 1981.

[5] R.J. Brandmayr and J.R. Vig, "Chemical Polishing in Etching Solutions that Contain Surfactants," Proceedings of the 39th Annual Symposium on Frequency Control, 1985, pp. 276-281.

[6] R.J. Brandmayr and J.R. Vig, "Further Results on the Use of Surfactants in Chemically Polishing Quartz Crystals," Proceedings of the 40th Annual Symposium on Frequency Control, 1985, pp. 86-90.

[7] J.R. Hunt and R.C. Smythe, "Chemically Milled VHF and UHF Resonators," Proceedings of the 39th Annual Symposium on Frequency Control, 1985, pp. 292-300.

[8] W.P Hanson, "Chemically Polished High Frequency Resonators," Proceedings of the 37th Annual Symposium on Frequency Control, 1983, pp. 261-264.

[9] Allied Chemical Co., Electronic Chemical Products, 20 Peabody Street, Buffalo, New York, 14210.

[10] J.R. Hunt, "Etch Pits and Etch Channels in Swept AT- and SC-Cut Quartz," Proceedings of the 41st Annual Symposium on Frequency Control, 1987, pp. 181-191.

[11] W.H. Horton and R.C. Smythe, "Theory of Thickness-Shear Vibrators with Extensions and Application to VHF Acoustically-Coupled Resonator Filters," Proceedings of the 21st Annual Symposium on Frequency Control, 1967, pp. 160-178.

[12] H.E. Bömmel, W.P. Mason, and A.W. Warner, Jr., "Experimental Evidence for Dislocations in Crystalline Quartz," Phys. Rev., v. 99, 1955, pp. 1894-1896.

[13] E. Hafner, " The Piezoelectric Crystal Unit -- Definitions and Methods of Measurement," Proc. IEEE, v. 57, no. 2, Feb., 1969, pp. 179-201.

[14] R.C. Smythe, "An Automated Resonator Measurement System Using a Reflection Coefficient Bridge," Proceedings of the 35th Annual Symposium on Frequency Control, 1981, pp. 280-285.

[15] D.B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," Proc. IEEE, v. 54, no. 2, Feb., 1966, pp. 329-330.

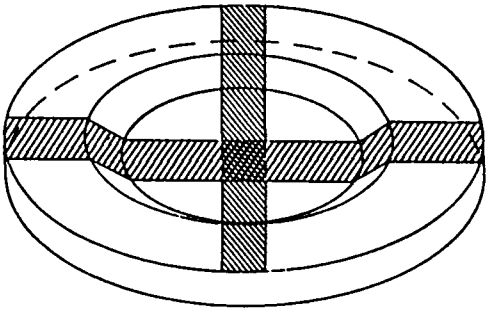


Figure 1. Ring-Supported Resonator Configuration, Showing Novel Electrode Geometry.

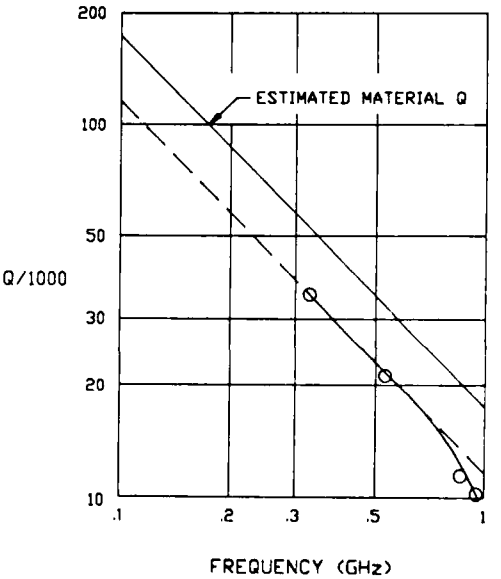



Figure 2. Best Measured Q's for Fundamental Mode SC-Cut Resonators. Upper line is the estimated intrinsic Q.

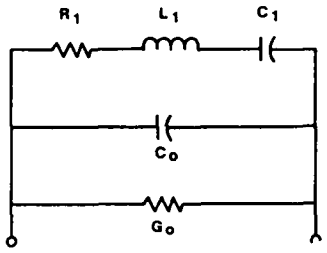
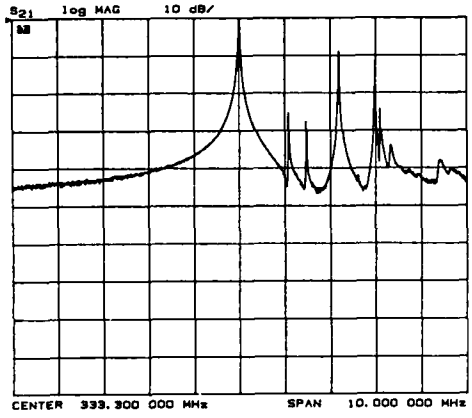
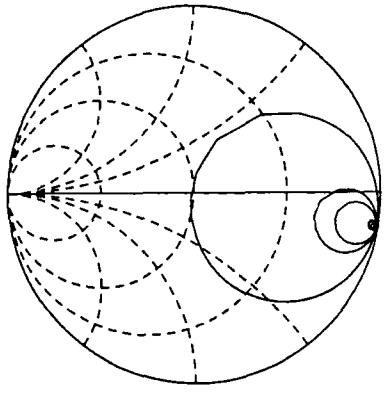


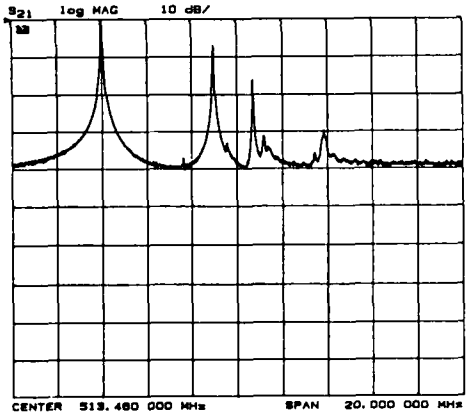

Figure 3. Two-Terminal Resonator Equivalent Circuit. Conductive loss is represented by G0.



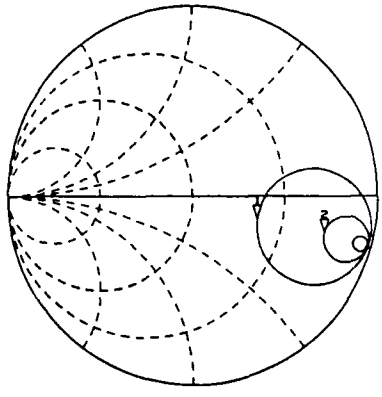

a)



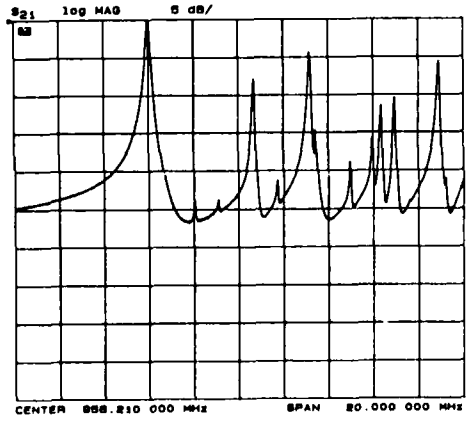
b)



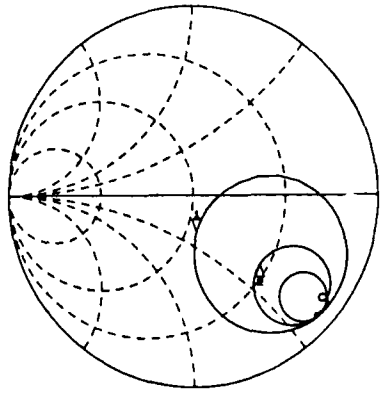

c)



d)




e)



f)

Figure 4. Resonator Frequency Characteristics: a), c), & e) are unwanted mode plots of 333 MHz, 500 MHz, and 1 GHz resonators No. 74, 31, & 25; b), d), & e) are Smith chart admittance plots of the same resonators.

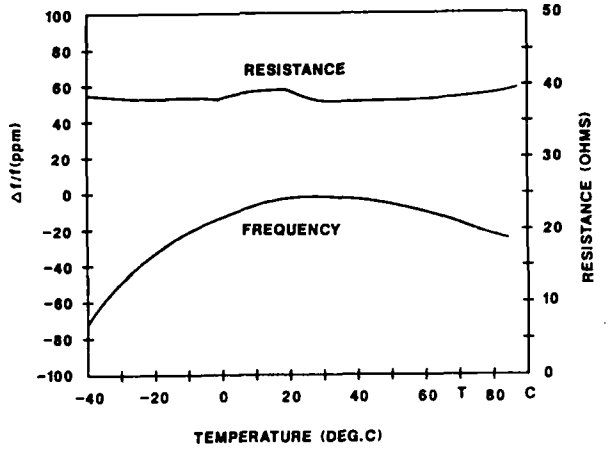



Figure 5. Frequency and Resistance vs. Temperature for 333 MHz Resonator No. 32.

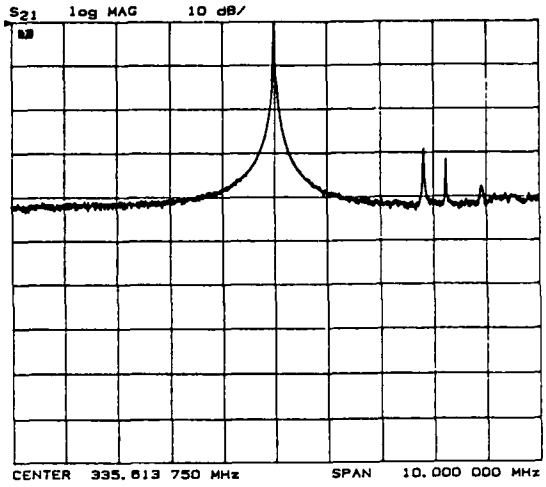



Figure 7. Unwanted Mode Plot, 333 MHz Resonator No. 207, Showing Improvement in Unwanted Mode Performance due to Reduced Electrode Width.

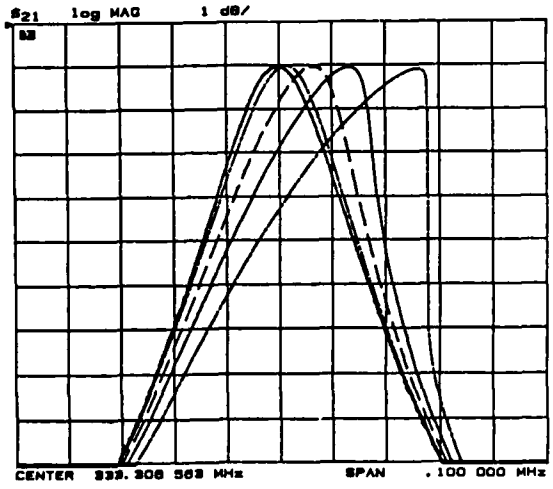



Figure 6. Non-Linear Resonance Curves, 333 MHz Resonator No. 3. Maximum motional arm currents are (left to right): 0.31, 1.2, 2.5, 3.6, and 5.0 mA.

42nd Annual Frequency Control Symposium - 1988

# A LOW COST FORCE SENSING CRYSTAL RESONATOR APPLIED TO WEIGHING

William C. Albert
Crystal Gage Inc.
P.O. Box 668
Franklin Lakes, NJ 07417

## SUMMARY

A general purpose, force sensing, crystal resonator has been developed that is low in cost. The very simple design employs the single vibrating beam approach that also includes a dynamic isolator system at each end to maintain high Q. The beam is a rotated X-cut which exhibits a high turnover temperature (up to 38°C) and performs very well in atmosphere as well as vacuum. Typical bias frequencies are from 20 to 26 kHz, depending on beam geometry, and they are capable of changing frequency typically + or - 10% in response to 700 gram axial tension or compression forces respectively. Although these resonators are suitable for many applications, the following describes how the resonators are used as sensors in a load cell. The load cell is designed to use the resonators as deflection sensors so that the same resonator can be used in load cells from 5 to 500 kilogram or higher capacities. Weighing to an accuracy of better than one part in 20,000 has already been demonstrated and indications are that one part in 40,000 or better is feasible. In this load cell application, the resonators are used in push-pull pairs and the frequency signals are differenced so that common mode rejection of thermal, nonlinearity, and other error sources is achieved.

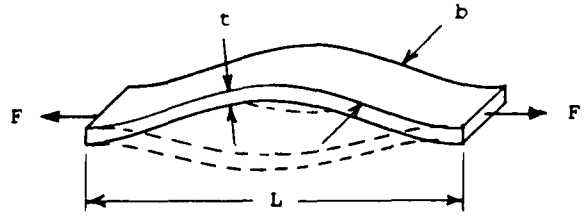



FIGURE 1
FIXED-FIXED VIBRATING BEAM
UNDER AXIAL FORCE

## BACKGROUND

A previously published Rayleigh method solution of the frequency-force effect of a vibrating flexure beam resulted in the following equations.[1,2]

$$f = f_o\sqrt{1 + SF} \tag{1}$$

Where:

$$f_o = a_o \frac{t}{L^2}\sqrt{\frac{E}{\rho}} \tag{2}$$

$$S = a_s \frac{L^2}{Ebt^3} \tag{3}$$

Where:

The beam geometry, vibration mode, and boundary conditions are described by Figure 1.

E and ρ are respectively the elastic modulus and density of the beam material.

The dimensionless constant $a_o$ and $a_s$ are respectively 1.026 and 0.294 for $f$ in Hz, the fundamental vibration mode, and the fixed-fixed boundary conditions.

For transducer applications, usually a form of the measurand F is to be determined. This is done by squaring Equation 1 and then rearranging to obtain the following.

$$F = \frac{1}{S}\left[\left(\frac{f}{f_o}\right)^2 - 1\right] \tag{4}$$

CH2588-2/88/0000-078 $1.00 C 1988 IEEE

Because of assumptions made in determining Equation 1 along with other sources of nonlinearity, additional empirical terms are sometimes added to achieve greater curve fitting accuracy.

$$F = A\left[\left(\frac{f}{f_o}\right)^2 - 1\right] + B\left[\quad\right]^2 + C\left[\quad\right]^3 \cdots \qquad (5)$$

Another useful relationship is obtained by applying the binomial expansion to Equation 1.

$$f = f_o\left(1 + \frac{S}{2}F - \frac{S^2}{8}F^2 + \frac{S^3}{16}F^3 \ldots\right) \qquad (6)$$

Equation 6 will prove useful later in describing the dual beam operation of a prototype load cell.

Another interesting relationship is obtained by combining Equations 1 and 3 and then differentiating with respect to F to obtain the approximation of Equation 7.

$$\frac{\Delta\,(f/f_o)}{\Delta F} \approx \frac{a_s}{2}\,\frac{L^2}{Ebt^3} \qquad (7)$$

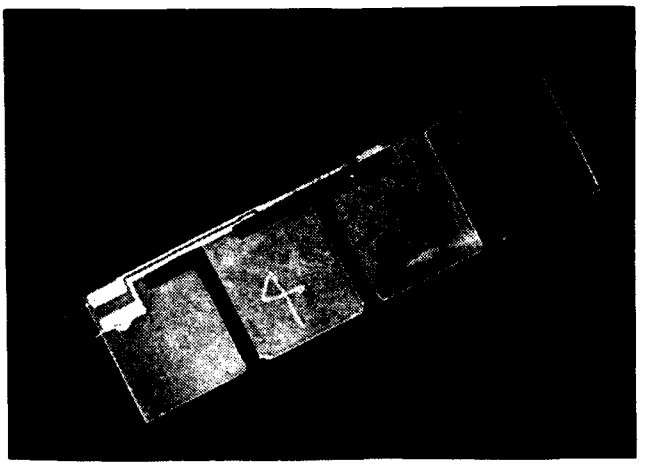
Recognizing that Δ F/bt is the applied tensile stress (s), Equation 7 can be modified to determine sensitivity in terms of applied stress as follows.

$$\frac{\Delta\,(f/f_o)}{\Delta s} \approx \frac{a_s}{2E}\left(\frac{L}{t}\right)^2 \qquad (8)$$

Equation 8 is included as background information to illustrate the strong influence of the length to thickness ratio (L/t) on the sensitivity of the basic frequency-force relationship.

## NEW RESONATOR DESIGN

Figure 2 illustrates the patented[3] crystal resonator design and describes how the very simple geometry provides the three basic necessary features of a vibrating beam, an isolation system, and mounting provisions with just a minimum of quartz machining operations. This is a key to low cost. The single vibrating beam approach (as opposed to the double ended tuning fork approach) requires an isolation system to prevent beam vibration energy from escaping and thereby causing a degradation of Q. With the aid of Figure 3, the operation of the isolation system is described as follows



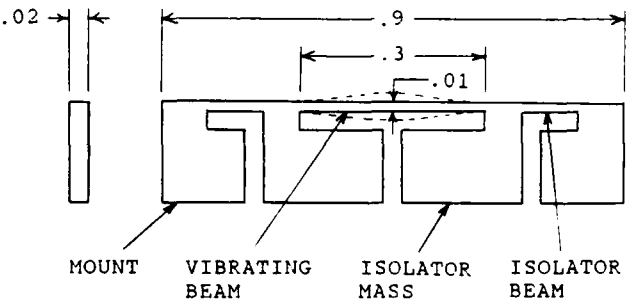



FIGURE 2
NEW CRYSTAL RESONATOR
(TYPICAL DIMENSIONS IN INCHES)

The vibrating beam is in reality a virtual fixed-fixed beam since there is slight movement at the beam root. Root movement ($y_R$) is typically on the order of 1% of maximum beam displacement ($y_{MAX}$). The isolator beam and isolator mass combine to become a resonant spring-mass system with a natural frequency ($f_I$) on the order of 10% of the vibrating beam frequency ($f_B$).

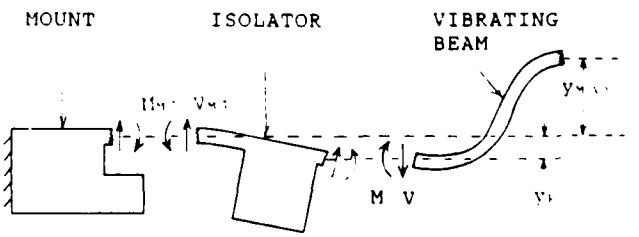



FIGURE 3
VIBRATION ISOLATION SYSTEM

The forcing functions on this system are the moment (M) and shear (V) reactions at the vibrating beam root. Because the natural frequency of the isolator system is much lower than the frequency of the forcing function, the reaction of the isolator system is primarily inertial and therefore, the relative phase of vibrating beam and isolator mass displacements are as indicated in Figure 3. By simple dynamic analysis, it can be shown that the ratio of mount reactions ($M_{MT}$ and $V_{MT}$ ) to vibrating beam root reactions can be described by the approximation of Equation 9.

$$\frac{V_{MT}}{V} \approx \frac{M_{MT}}{M} \approx -\left(\frac{f_I}{f_B}\right)^2 \qquad (9)$$

Other resonator characteristics are as follows:

- Typical dimensions are as illustrated in Figure 2.
- $f_o$ bias frequencies are from 20 to 26 kHz depending on beam geometry.
- Full scale frequency changes are from ± 7 to ± 10% (again depending on beam geometry) in response to 700 gram tension or compression forces respectively.
- The vibrating beam is a rotated X-cut exhibiting a turnover temperature as high as 38° C.
- Vibrations are sustained by a conventional X-cut flexure drive system. Because of the high coupling, the beam vibrates well in atmosphere as well as vacuum.

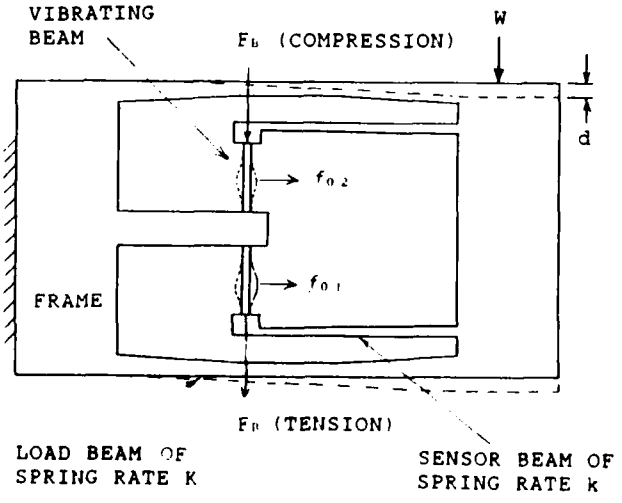



FIGURE 4
LOAD CELL SCHEMATIC

## LOAD CELL APPLICATION

Figure 4 describes schematically how the resonators are used in a load cell application which has the following characteristics.

- The applied load (W) is essentially all taken up by the load beams resulting in a deflection (d). Deflection of the resonators is negligible compared to d.
- The resonators are each attached to the load cell frame at one end and to a sensor beam at the other end. Note that the resonators are arranged so that the applied load places one vibrating beam in tension and the other in compression.
- The sensor beams are designed to achieve full scale resonator force ($F_B$ tension or compression) at full scale frame deflection d.The resonators are therefore used as deflection sensors so that the same resonator can be used in load cells from 5 to 500 killogram capcities. Different capacities are achieved by different load beam geometries.

### Frequency Differencing

When the load W is applied, the frequency of the tension beam will increase and that of the compression beam will decrease. By making the proper substitutions into Equation 6 (note that a compression force is $-F_B$ ) the frequency behavior of each resonator as well as the difference frequency can be analytically described by Equations 10, 11, and 12.

$$f_1 = f_{01} + k_{11}W + k_{21}W^2 + k_{31}W^3 \qquad (10)$$

$$f_2 = f_{02} - k_{12}W + k_{22}W^2 - k_{32}W^3 \qquad (11)$$

$$\Delta f = f_1 - f_2 = (f_{01} - f_{02}) + (k_{11} + k_{12})W + (k_{21} - k_{22})W^2 + (k_{31} + k_{32})W^3 \qquad (12)$$

Frequency differencing is a technique used in crystal resonator accelerometer applications[2] and also used here to achieve the following advantages:

- The bias frequency ($f_{01}$ - $f_{02}$) becomes small. More importantly, the benefit of common mode rejection (CMR) of bias frequency thermal drift is achieved since both essentially identical crystals will have similar temperature sensitivities. Other error sources are also reduced by CMR.

- The dominant first order scale factor ($k_{11}$ + $k_{12}$) becomes double that of an individual resonator.

- The most dominant nonlinearity term ($k_{21}$ - $k_{22}$) is greatly reduced by cancellation for improved linearity.

As test data will show, the best accuracy was obtained by determining W from a third order fit of the difference frequency as described by Equation 13 where $\Delta f_0$ is the no-load difference frequency.

$$W = J(\Delta f - \Delta f_o) + K(\quad)^2 + L(\quad)^3 \qquad (13)$$

Load Cell Design Inherent Temperature Compensation

In addition to the common mode rejection of resonator bias frequency temperature sensitivity as just described, the load cell design has an additional temperature compensating mechanism described next.

By design, load beam spring rate K is much greater than sensor beam spring rate k and the resonator axial deflection is negligible compared to d. Making those assumptions results in the following relationships.

$$d = \frac{W}{2K} \quad and \quad F_B = dk \qquad (14)\ (15)$$

Combining Equations 14 and 15 results in the following.

$$\frac{F_B}{W} = \frac{k}{2K} \qquad (16)$$

Equation 16 is next modified to include temperature sensitive effects.

$$\frac{F_B}{W} = \frac{k_o\left(1 + \frac{\Delta k}{\Delta T}T\right)}{2K_o\left(1 + \frac{\Delta K}{\Delta T}T\right)} \qquad (17)$$

Since both the sensor beams and load beams are flexure beams made of the same load cell frame material, $\Delta K/\Delta T = \Delta k/\Delta T$ and therefore, when Equation 17 is differentiated with respect to temperature, the following is obtained.

$$\frac{\Delta(F_B/W)}{\Delta T} = 0 \qquad (18)$$

Experimental data on the effectiveness of this inherent temperature compensation will be presented later.

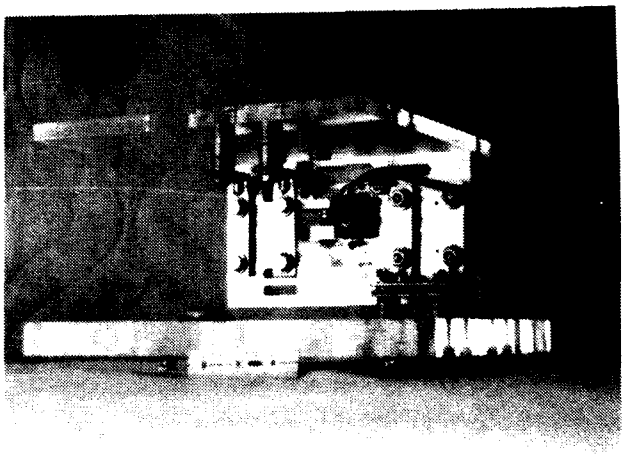


FIGURE 5
20 KILOGRAM SCALE

PROTOTYPE LOAD CELL

Several 20 kilogram FS experimental load cells were fabricated and then installed in the platform scale arrangement of Figure 5. In this weighing application, crystal resonator technology is expected to have the following advantages over conventional strain gage technology.

- The output will be inherently digital with no need for A/D conversion. This greatly simplifies the signal processing electronics. The simple oscillator circuits are similar to those found in wrist watches and other low cost crystal devices.

- Inherent accuracy will be high because of the nature of the quartz resonators and also because the signal frequencies can be accurately determined by comparing them to the system reference frequency which is invariably another crystal.

- For the same reasons, signal resolution will also be great. Incremental weight resolutions as low as 2ppm of full scale have been demonstrated.

- Because the crystal resonators are used as deflection sensors, the design stresses of the load beams and sensor beams of the load cell frame can be kept low. There is no need for highly concentrated stresses as with strain gage technology. Frame beam stresses can also be more evenly distributed as suggested by the varying beam cross sections illustrated in Figures 4 and 5. Lower stresses result in reduced hysteresis and creep effects.

- Using the resonators as deflection sensors insures that the resonators themselves will not experience excessive loads. By design, the deflection (d ) is relatively large so that overload stops can be easily installed and adjusted. In addition, the sensor beams are designed to have minimum mass to reduce shock susceptibility. The resonators themselves are mechanically fastened to the load cell frame. Avoiding the use of an epoxy for this purpose has greatly improved hysteresis and creep performance. This is unlike conventional strain gage technology where epoxy bonding is a limiting factor in obtaining good creep and hysteresis performance.

Performance Test Results

Several scales of the configurations of Figure 5 were tested with many subtle variations of the basic resonator design of Figure 2. The performance achieved is presented next.

Linearity The best linearity achieved to date on both individual resonators as well as the difference frequency was achieved by third order least square curve fits. For the individual resonator, the expression used was a third order form of Equation 5. For the difference frequency, a third order form of Equation 13 was used. The bias frequencies, full scale frequency changes, and full scale contributions of the nonlinear coefficients are given in Table 1. The residuals to a least square curve fit are plotted in Figures 6 and 7. This data was taken from two runs on the same unit and represents one of the better performing combinations of resonator and frame.

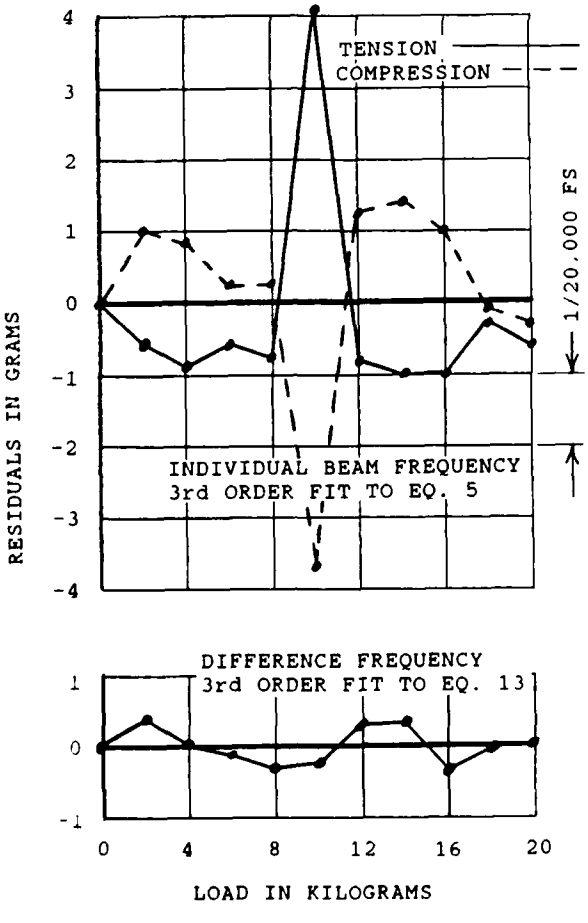



FIGURE 6
LINEARITY TEST (RUN 1)

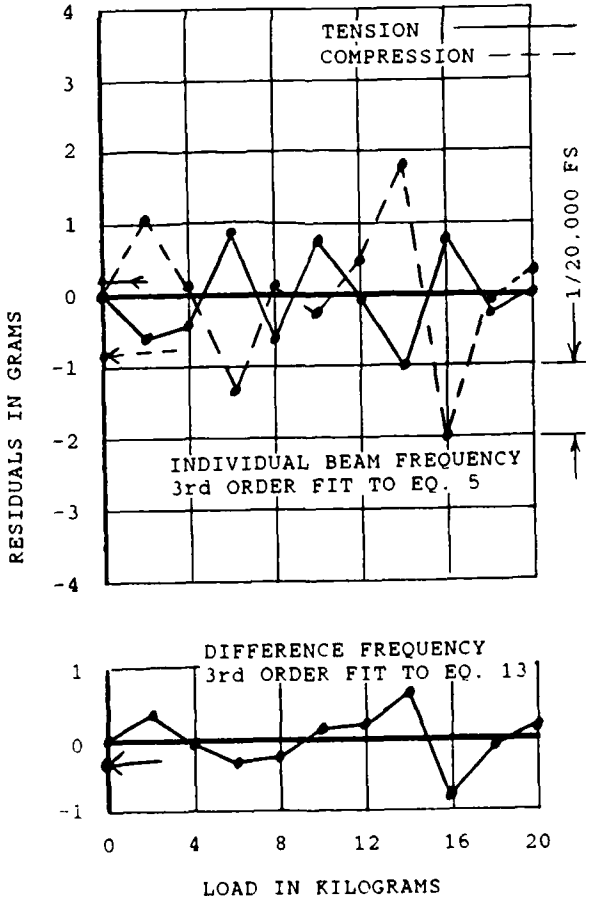



FIGURE 7
LINEARITY TEST (RUN 2)

TABLE 1
VARIOUS COEFFICIENTS

| COEFFICIENT | TENSION BEAM | COMPRESS BEAM | DIFF FREQ. |
|---|---|---|---|
| BIAS FREQUENCY | 22,32 k Hz | 22,02 k Hz | 306,6 HZ |
| FS FREQ CHANGE | + 1,30 k Hz | -1,44 k Hz | + 2740 HZ |
| CONTRIB OF 2ND ORDER EFFECT AT FS | -1,74% | -,18% | -1,66% |
| CONTRIB OF 3RD ORDER EFFECT AT FS | + 0,41% | + 0,49% | -,31% |

The following observations apply to the results presented in Figures 6 and 7 as well as Table 1.

- It is readily apparent that frequency differencing can improve the overall performance by an order of magnitude or more. This has also been the experience of accelerometer applications of this technology.[2] Common mode rejection of bias frequency error sources is the reason.

- A benefit of common mode rejection is dramatically illustrated in Figure 6 where both tension and compression resonators show an anomaly at the 10 kilogram load level. The reason for this anomaly is not known for sure but off-center load effects are suspected. Nevertheless, note that the difference frequency did not reflect this anomaly because of CMR.

- Based on data obtained to date, (exclusive of the anomaly) it is estimated that weighing to one part in 5,000 to 10,000 is presently possible for single resonator load cells. Weighing to one part in 20,000 to 50,000 is estimated for the dual resonator push-pull approach.

- Table 1 shov s an inconsistency in the contribution of the second order nonlinear effect between the tension and compression resonator. It is suspected that bending of the vibrating beam due to bending of the sensor beam during load application is the cause. Future designs will greatly reduce or eliminate this effect.

Hysteresis and creep For practical purposes, quartz crystal can be assumed purely elastic. Any hysteresis or creep effects are due to instabilities of the resonator-frame interface and plastic behavior of the frame material. By eliminating epoxy and limiting the frame stresses, hysteresis and creep magnitude has been kept within the residuals of Figures 6 and 7. Figure 7 includes zero repeat data that is typical and shows that these error sources have been kept within one part in 20,000.

Temperature Sensitivity The load cell frames were made of aluminum primarily because its use is traditional in the conventional load cell field. Aluminum is a very poor thermal expansion coefficient match to quartz crystal. Nevertheless, the CMR effects of frequency differencing along with the inherent temperature compensation of the design as previously described, greatly reduced thermal errors. A scale of the Figure 5 configuration was tested for temperature sensitivity. The bias (no load) temperature sensitivity for both individual resonators as well as the difference frequency is presented in Figure 8. Note that the frequency drift of the individual resonators was quite pronounced due to the thermal expansion rate of aluminum being much greater than quartz. However, the difference frequency drastically reduced this effect. The difference frequency bias thermal shift was about -0.0018% FS/°C).

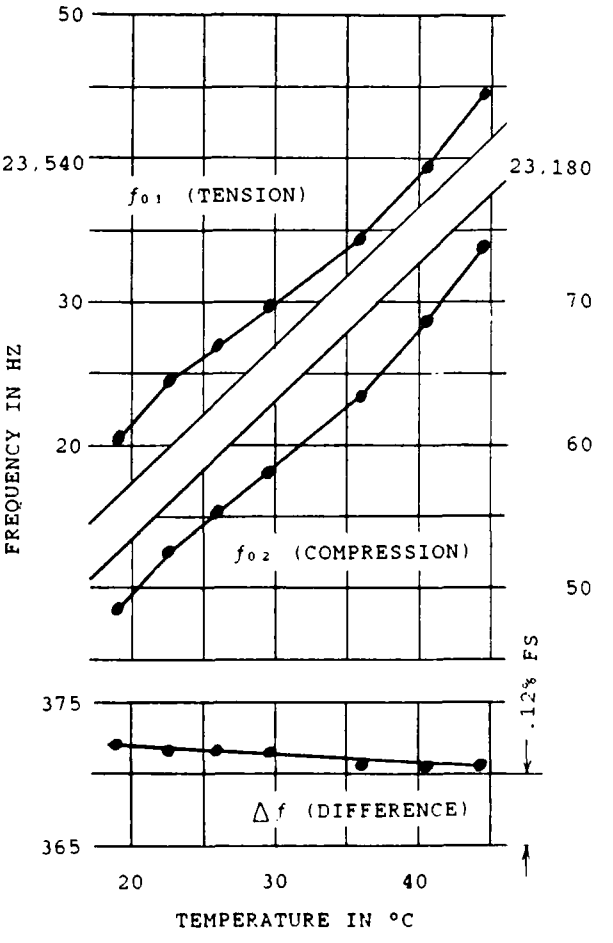



FIGURE 8
BIAS (NO LOAD) FREQUENCY
TEMPERATURE SENSITIVITY

The temperature sensitivity of scale factor was also determined by determining the half-scale frequency shift at various temperatures. For these tests, the scale factor data on the individual resonators was erratic due to off-center load sensitivity of this particular design. However, the off-center load effects were common mode rejected so that good difference frequency scale factor data was obtained that is presented in Figure 9. Note that the scale factor temperature sensitivity is about + 20 ppm/°C. This low scale factor temperature sensitivity is due to the inherent low temperature sensitivity of quartz crystal as well as the inherent temperature compensating behavior of the load cell frame design approach as previously explained.

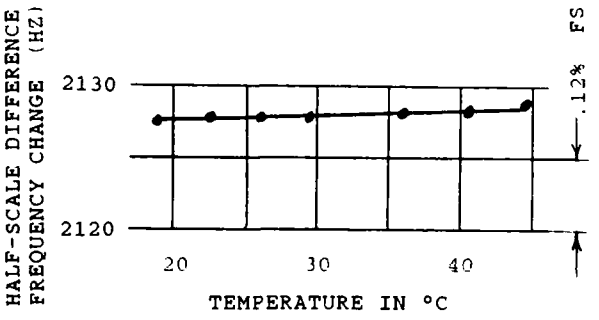



FIGURE 9
SCALE FACTOR
TEMPERATURE SENSITIVITY

## CONCLUSIONS AND FUTURE IMPROVEMENTS

The test results presented have demonstrated that force sensitive crystals have a place in the weighing field. The development of a low cost resonator which works well in air makes the application of this technology practical. This is particularly true for counting scales where accuracy and high resolution are needed. In this area, it is believed that crystal resonators will have an advantage over conventional strain gage technology. Much has been learned from designing, assembling, and testing the engineering prototype scales discussed. As part of the next design iteration, improvements will be made in the following areas.

- The load cell frame envelope will be redesigned so that it conforms to one of the standard load cell configurations preferred by scale manufacturers. In some cases it may be possible to substitute the redesigned frame for existing strain gage based load cells.
- Resonator improvements will continue to improve linearity, temperature sensitivity, and obtain better matching for improved common mode rejection of error sources.
- Resonator mounting techniques will be improved to provide simpler assembly and also further reduce hysteresis and creep effects.
- Tests on the engineering prototypes along with design analysis have revealed the causes of off-center load effects. The next design will greatly reduce this effect.
- Frame materials other than aluminum will be investigated. Material properties considered will be a better thermal expansion match to quartz, along with higher yield strengths for lower hysteresis and creep effects.

## ACKNOWLEDGEMENT

The author gratefully acknowledges the contribution of the Ohaus Scale Corp. in performing the linearity tests.

## REFERENCES

(1) Albert, W.C., "Force Sensing Using Quartz Crystal Flexure Resonators". Proceedings of the 38th. Frequency Control Symposium, 1984, pp 233-239.

(2) Albert, W.C., "Vibrating Quartz Crystal Beam Accelerometer". ISA 28th. International Symposium, Vol. 28, No. 1, 1982, pp 33-44.

(3) Albert, W C., U.S Patent No. 4,743,790, Force Sensing Vibrating Beam Resonator, May 10, 1988.

42nd Annual Frequency Control Symposium - 1988

# Stroboscopic X-ray topography of quartz resonators

**A.ZARKA, B.CAPELLE & Y.ZHENG**
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

Stroboscopic X-ray topography of 1ns time resolution with synchrotron radiation has been used to study the vibrations in quartz resonators. The study has shown that time-resolved images of the vibration reveal the existence of particular modes which cannot be observed on time-integrated images by conventional methods. A theoretical calculation is being carried out to characterize Pendellosung fringes on stroboscopic X-ray section topographs in weakly vibrating crystals. The agreement between the theoretical simulations and the experimental results demonstrates that for low vibration amplitude range ($u_0 \leq 7$Å), Pendellosung fringes can be related to vibration amplitudes. Stroboscopic X-ray traverse topographs are also discussed and it appears that equal-amplitude curves of the vibration can be imaged. All these examples show that the technique of stroboscopic X-ray topography is very useful to the study of the vibrations in resonators.

## Introduction

Stroboscopic X-ray topography has been performed to observe the acoustic vibrational states in quartz resonators with a time resolution of 1ns. Such observations have revealed interesting features concerning the existence of particular vibration modes and the vibration amplitudes, which cannot be provided by the other classical methods. The experimental results and the theoretical simulations show that the stroboscopic technique is a powerful tool to study the acoustic vibrations.

## Technique of stroboscopic X-ray topography

The synchrotron radiation from a Storage Ring is pulsed at a frequency of some MHz corresponding to the rotation frequency of particle bunch. For the D.C.I. Storage Ring (LURE, France) which we have used, the frequency of pulsed X-ray radiation is *3.169280 MHz. The stroboscopic topography appears* to be well adapted to study the resonators because the frequency of the vibrations in the crystals is of the same order (some MHz).

The experimental set up consists of two parts[1]:

- *The topographic setting*. It consists of the Laue setting, i.e. the quartz sample is placed in the white incident beam to obtain a Laue pattern of topographs[2]. On this pattern, the crystal is simultanously imaged with different diffraction vectors **g** whose sensitivity to the vibration displacement vector **u** is connected to the quantity **g.u**. In particular, the contrast due to the vibration vanishes on the topograph when **g.u**=0. This means that with one particular diffraction vector **g**, one component of **u** can be selected and studied. By limiting the incident beam with a very fine slit (about 15μm), section topographs are obtained. Traverse topographs are made with a wide incident beam and they can be understood as an addition of an infinite number of section topographs.

- *The electronic control*. To obtain an exact synchronization between the X-ray pulses and the piezoelectric vibrations of the resonators, a pulse signal obtained from the positron bunch (by a capacitive pick up) is used to generate the excitation signal of the resonator. This signal is shaped in sinusoidal form by filtering at the recurrence frequency of the synchrotron

CH2588-2/88/0000-085 $1.00 C 1988 IEEE

($f_o$=3.169280MHz) or at an harmonic of this frequency $n.f_o$ (with n = 2,3 ...). A phase shifter permits to vary the relative phase between the X-ray pulse and the sinusoidal signal used to excite the resonance of the crystal. To have an exact resonance condition, the resonators are adjusted by metallization to frequencies very close to $f_o$ (or $n.f_o$). A fine adjustment (a few Hz) is made with a large serial variable capacitance. The electrical parameters of the resonator are measured, in situ, with a vector voltmeter. Another vector voltmeter is used to measure the relative phase angle between the resonator current and the X-ray pulse. After calibration this relative phase is known with a precision of about 3°, most of the uncertainities resulting of the effects of the temperature fluctuations on the resonator. The general principle of the set-up is given in Figure 1.

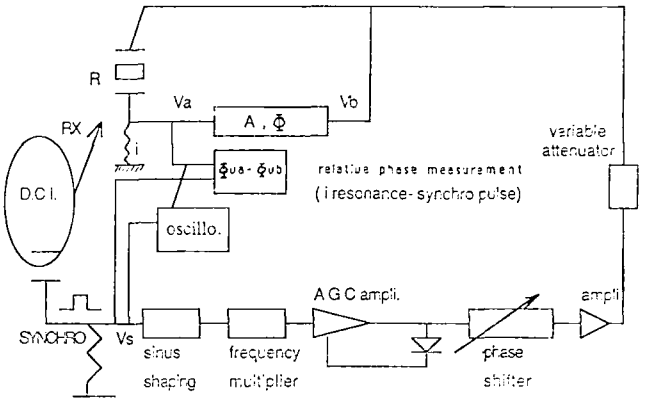



Fig.1 Experimental set up

## Interest of stroboscopic X-ray topography

The stroboscopic technique, because it provides a time-resolved image of the crystal, enables to reveal particular vibration features as it is shown in the following examples.

The first sample is a natural AT-cut plane resonator with the thickness $2h_o$=0.253mm and it operates at the fundamental thickness-shear mode (f= 38560MHz). As established by the theory[3,4], the dominating vibration component is the thickness-shear one $u_1$. The components $u_2$ and $u_3$ of the displacement are usually weak in regard to the $u_1$ component, when no coupling to flexure or plane shear occurs. Two selected diffraction vectors are $\mathbf{g_1}=2\bar{1}0$ and $\mathbf{g_2}=01\bar{3}$ where the first one $\mathbf{g_1}=2\bar{1}0$ is only sensitive to the component $u_1$ and the second one $\mathbf{g_2}=01\bar{3}$ is sensitive principally to $u_3$ and a little to $u_2$.

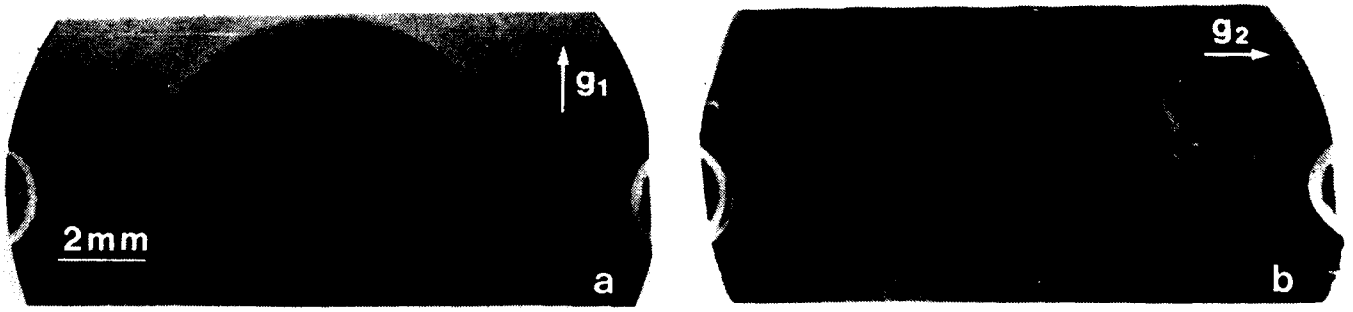



Fig.2 Classical X-ray topographs. a) Imaged with $\mathbf{g_1}=2\bar{1}0$; b) Imaged with $\mathbf{g_2}=01\bar{3}$

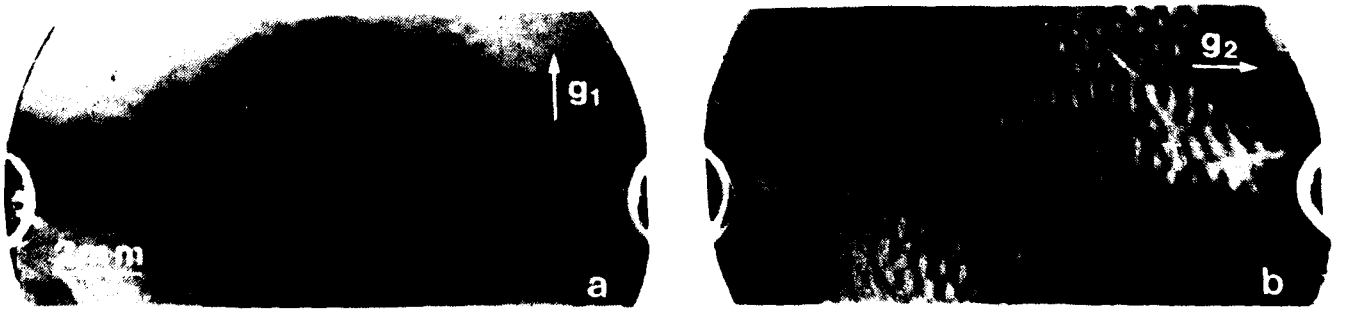



Fig.3 Stroboscopic X-ray topographs. a) Imaged with $\mathbf{g_1}=2\bar{1}0$; b) Imaged with $\mathbf{g_2}=01\bar{3}$

Two classical X-ray topographs obtained with $\mathbf{g}_1$ and $\mathbf{g}_2$ are shown in Figures 2a,b respectively and they are time-integrated images. In Figure 2a where only $u_1$ component is seen, a black contrast exists in the centre of the resonator and it corresponds to the strongly vibrating zone. In Figure 2b where $u_3$ and $u_2$ could be imaged, no contrast reveals their existence.

Similar X-ray topographs were recorded using the stroboscopic technique (Fig.3a,b). In Figure 3a obtained with $\mathbf{g}_1=2\bar{1}0$, a black contrast similar to that in Figure 2a can be seen, but some difference appears on the edge of the black contrast. This shows that the $u_1$ component over the whole resonator is not of the smooth form as described by the theory. This is probably due to the high excitation on the resonator so that the theory is not completely satisfactory. More interesting aspect is shown in Figure 3b which reveals the existence of the $u_3$ and/or $u_2$ components, which is absent on the time-integrated image (Fig.2b). Further, this particular component appears to be a vibration mode which extends up to the resonator edges. This aspect is important, because it suggests the existence of one coupled flexural mode and the origin of energy dissipation in the resonator can be due to the extention of this vibration mode. The identification of this mode and its modelling should be very interesting to understand the coupling between the vibration modes and the energy dissipation in the resonator.

## Stroboscopic X-ray section topography

Stroboscopic section topographs were made with the $\mathbf{g}=2\bar{1}0$ diffraction vector in order to study the $u_1$ component. The second sample studied was a synthetic AT-cut plano-convex quartz resonator (Fig.4a). The resonator has a radius of curvature 175mm and a maximal thickness 1.3196mm. It operates at the fifth overtone thickness-shear mode (f=6.338560MHz).

## 1. Basic understanding of the dynamic X-ray theory

The incident X-ray beam arriving at the crystal surface can be considered as a spherical wave. This spherical wave entering into the crystal is splitted into two types of wave fields because of the crystal periodic potential. Thus section topographs give the result of interferences on the crystal exit surface between the two wave fields. The interference fringes, called also Pendellosung fringes, are characterized by the parameter $\Lambda_o$, the extinction distance, which is related to the structure factor $F_g$ of the reflection.

For a vibrating crystal, two cases can be distinguished. When the deformation induced by the vibration is strong, nodal planes in thickness are imaged on section topographs without any other details[5]. When the vibration is weak, the crystal can be still considered perfect locally and section topographs represent images of interferences between two weakly modified wave fields in regard to the case of a perfect crystal. Thus for the last case, section topographs constitute an excellent method to characterize the vibration through the Pendellosung fringes. It is shown below that a relation between the fringe number and the vibration amplitude can be established.

## 2. Pendellosung fringes in a weakly vibrating crystal

To study Pendellosung fringes in a vibrating crystal, two approaches are available. The first one consists of resolving the Takagi's equations which are partial derivative equations correlating the direct and diffracted X-ray beams[6,7,8]. This

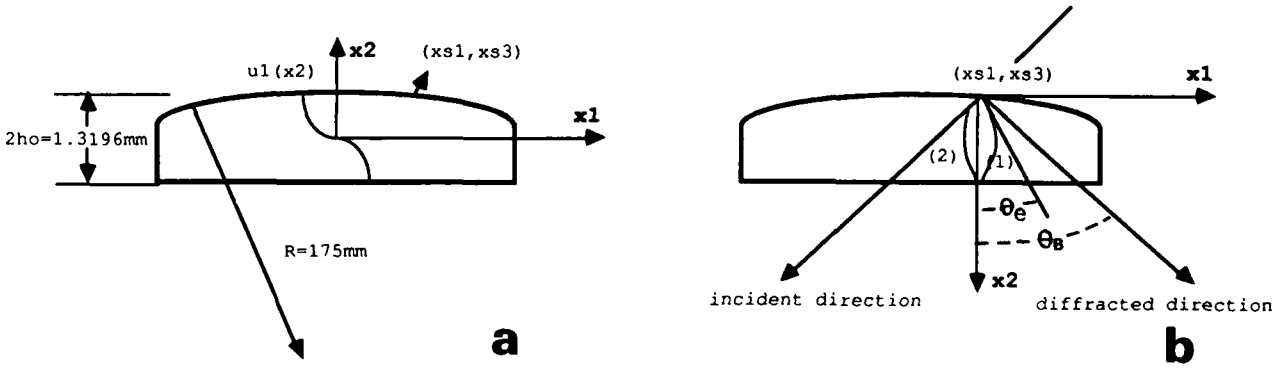



Fig.4 a) Vibration axis. ($xs_1$, $xs_3$) is a symbolic point of the X-ray beam arriving on the crystal surface; b) Diffraction axis. (1) and (2) are two schematic X-ray trajectories; $\theta_B$ is the Bragg angle and $\theta_e$ is the angle between a trajectory and the $x_2$ axis at the entrance point.

approach is rigorous, complicated and difficult to understand. The second one is based on the theory developed by Penning & Polder[9] and Kato[10,11] which keeps notions of the geometric optics, such as X-ray trajectories, intensity (or energy transported along each X-ray trajectory) and phases between different trajectories. This theory, called here the Kato's theory, is valid for a weakly deformed crystal. As it is concerned here, results deduced from the Kato's theory will be summarized in order to get an understanding of Pendellosung fringes[12]. In the next section, theoretical predictions from both theories will be compared with experimental results.

It will be considered here an ideal case of a pure thickness-shear mode vibration:

$$u_1(x_2) = u_o \cos(\frac{n\pi x_2}{2h}) \qquad (1)$$

where n is an odd number, 2h is the crystal thickness and the origin of the axis is taken at the X-ray entrance point on the crystal surface (Fig.4b). The Kato's theory is valid in the present case for the following range of vibration amplitude $u_o$:

$$\frac{u_o}{d_o} \leq \frac{1}{\pi^2}\left(\frac{L_o}{\Lambda_o}\right)^2 \qquad (2)$$

where $d_o$ is the spacing of reflecting planes, $L_o$ is the acoustic wavelength ($L_o=4h/n$) and $\Lambda_o$ is the extinction distance. For our case, $d_o=2.45$Å, $L_o=527.84\mu m$ and $\Lambda_o=101.6\mu m$, it means that the theory is applicable for the range of $u_o \leq 7$Å.

Two parameters characterizing the X-ray trajectories and the vibration amplitude need to be defined:

- The trajectory parameter a is related to the initial inclination of a trajectory at the entrance point. If $\theta_e$ is the angle between a trajectory and the $x_2$ axis at the entrance point and $\theta_B$ is the Bragg angle (Fig.4b), then $a = \dfrac{tg\theta_e/tg\theta_B}{\sqrt{1-(tg\theta_e/tg\theta_B)^2}}$.

- The vibration parameter b is a ratio between the acoustic characteristics $u_o/L_o$ and the diffraction ones $d_o/\Lambda_o$: $b = 2\pi \dfrac{u_o}{d_o}\dfrac{\Lambda_o}{L_o}$.

Thus all the X-ray characteristics in a vibrating crystal can be expressed in terms of these two parameters.

***Trajectories:***

The trajectories of the two X-ray wave fields are as following:

$$\frac{x_1}{2h.tg\theta_B} = \int_0^{\frac{x_2}{2h}} \frac{a \mp b\sin(n\pi\varphi)}{\sqrt{1+\{a \mp b\sin(n\pi\varphi)\}^2}} d\varphi \qquad (3)$$

where the upper and lower signs are attributed to the X-ray wave fields (1) and (2) respectively. It can be noted that the family of trajectories of the wave field (1) is symmetric with that of the wave field (2) with respect to the $x_2$ axis. In comparison with the straight trajectories in a perfect crystal, X-ray trajectories are of an oscillating form related to that of the acoustic wave. In the sense of the geometric optics, the X-ray optical paths are longer in a vibrating crystal than in a perfect crystal.

***Intensities:***

The diffracted intensity can be calculated for each trajectory depending on the boundary conditions of the trajectory on the entrance and exit surfaces:

$$I = \frac{\pi/(4\Lambda_o 2h.\sin^2\theta_B)}{(1+a^2)\displaystyle\int_0^1 \frac{d\varphi}{\sqrt{1+\{a \mp b\sin(n\pi\varphi)\}^2}^3}} \qquad (4)$$

***Phase and total intensity:***

For any point of the exit face, we have two trajectories ending at this point and belonging to wave fields (1) and (2) respectively. Their trajectory parameters $a_1$ and $a_2$ respectively can be determined from the equation (3). Thus the phase difference S between these two wave fields is established as following:

$$S = -\frac{\pi}{2} + \frac{\pi 2h}{\Lambda_o}\int_0^1 \sqrt{1+\{a_1 - b\sin(n\pi\varphi)\}^2}\, d\varphi + \frac{\pi 2h}{\Lambda_o}\int_0^1 \sqrt{1+\{a_2 + b\sin(n\pi\varphi)\}^2}\, d\varphi \qquad (5)$$

The total diffracted intensity at any point of the crystal exit suface is the sum of the intensities of the two wave fields and their interference:

$$I_{total} = I_1 + I_2 + 2\sqrt{I_1 I_2}\cos(S) \qquad (6)$$

The term of phase difference is important, because it determines the Pendellosung fringe contour. In particular, the fringe maximum positions are defined by $S=2m\pi$ (m=0,1,2...). It can be shown that in a vibrating crystal, the phase difference between two wave fields for any given point will be increased with the increasing vibration amplitude, i.e. some extra Pendellosung fringes, with respect to the case of the perfect crystal, appear with increasing vibration amplitude.

***Relation between extra fringes and vibration amplitude:***

In a section topograph, fringes are mostly spaced at the central part. To relate extra fringes to the vibration amplitude, the total diffracted intensity at the central point along the exit face is calculated as a function of $u_o$. From the trajectory

equation (3), it can be noted that the two trajectories of wave fields (1) and (2) respectively arriving at the central point along the exit face $x_1(x_2=2h)=0$ have the opposite trajectory parameters $a_1=-a_2$. $a_1$ is noted here as $a_0(b)$. Further, the two wave fields at this point have the same intensity $I(a_1)=I(a_2)=I(a_0)$. Thus the total intensity is related to the vibration amplitude through the function $a_0(b)$.

$$I_{total} = 2I(a_0)(1+\cos(S)) \qquad (7)$$

It is recalled that the phase difference for the present case is

$$S = -\frac{\pi}{2} + \frac{2\pi 2h}{\Lambda_0}\int_0^1 \sqrt{1+\{a_0-b\sin(n\pi\varphi)\}^2}\, d\varphi \qquad (8)$$

The total diffracted intensity at the central point along the exit face is plotted as a function of $u_0$ (Fig.5) where the crystal thickness is fixed to $2h_0$ and n=5. This curve shows that when the vibration amplitude $u_0$ increases from 0Å to 5Å, 12 extra Pendellosung fringes appear. So, the number of extra fringes appearing on section topographs is already a very sensitive indication of the vibration amplitude range.

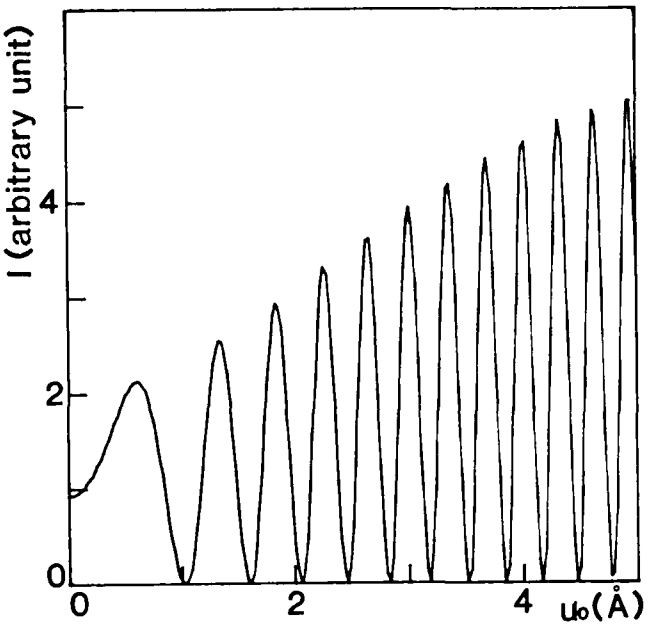



Fig.5 Total diffracted intensity at the central point along the exit face as a function of the vibration amplitude $u_0$.

## 3. Application to quartz resonator

### *Deformation model:*

The piezoelectric theory[4,13] leads to a simplified deformation model for the fifth overtone thickness-shear mode as following (in the vibration axis of Figure 4a):

$$u_1 = u_0 \exp\{-\alpha_5 x_1^2/2-\beta_5 x_3^2/2\}\ \sin(\frac{5\pi x_2}{2h}) \qquad (9)$$

where $u_0$ is related to the excitation voltage V and empirically to the quality factor Q, when the vibration is at/near the resonance:

$$u_0 = \text{constant} \cdot V \cdot Q \qquad (10)$$

In our case, the quality factor Q is about $1\times10^6$ in room environment. It means that for an applied voltage of about 1mV, the vibration amplitude $u_0$ is about 1Å. The value of $u_0$ is approximate, because Q is not known very precisely.

Further, under stroboscopic experiments, the displacement $u_1$ appears to be fixed and has an apparent maximal amplitude $u_0\cos\Phi$, where $\Phi$ is the relative phase angle between the vibration and the pulsed X-ray. The apparent displacement $u_i$ in the diffraction axis (Fig.4b), when an X-ray beam arrives at a point $(xs_1, xs_3)$ on the entrance surface, is thus

$$u_i = u_0\cos\Phi \exp\{-\alpha_5(x_1+xs_1)^2/2-\beta_5 xs_3^2/2\}$$

$$\cos(\frac{5\pi x_2}{2h(x_1+xs_1,xs_3)}) \qquad (11)$$

The crystal thickness 2h and the vibration amplitude depend here on the $x_1$ and are not strictly constant over the Borrmann triangle. But these variations are small and can be neglected. In putting the values at the entrance point $(xs_1,xs_3)$, an approximate expression of $u_i$ is:

$$u_i = u_0(xs_1,xs_3)\cos(\frac{5\pi x_2}{2h(xs_1,xs_3)}) \qquad (12)$$

where $u_0(xs_1,xs_3) = u_0\cos\Phi \exp\{-\alpha_5 xs_1^2/2-\beta_5 xs_3^2/2\}$

corresponds to the local vibration amplitude, and $2h(xs_1,xs_3)$ is the local crystal thickness. With this expression where the vibration is approximated to a pure thickness-shear mode locally, the total intensity at the central point along the exit face can be calculated for any value of $xs_3$ by using the formulae of the previous section.

### *Comparison with experimental images:*

The first experimental section topograph (Fig.6b) was taken in the middle of the electrode on the resonator: $xs_1=2.07$mm, and with a relative phase angle $\Phi=\pi/4$. The excitation voltage V was 6.046mV which implies a vibration amplitude for the top part of the section topograph, $u_0(xs_1,xs_3)=u_0(2.07,0)$, of about 1.61Å for $Q=1\times10^6$. By taking the extra fringe positions from the central line of the experimental image, the vibration amplitude $u_0(xs_1,xs_3)$ was determined by comparison with the intensity profile $I(xs_3)$ according to the Kato's theory (Fig.6a): $u_0(xs_1,xs_3)=u_0(2.07,0)=1.35$Å. Using this value of $u_0(xs_1,xs_3)$, the simulation was performed by resolving the Takagi's equations including the $x_1$ dependent term for the vibration amplitude into the deformation model (Fig.6c). It can be seen that the extra fringe positions in the experimental image are reproduced either by the simulated image or

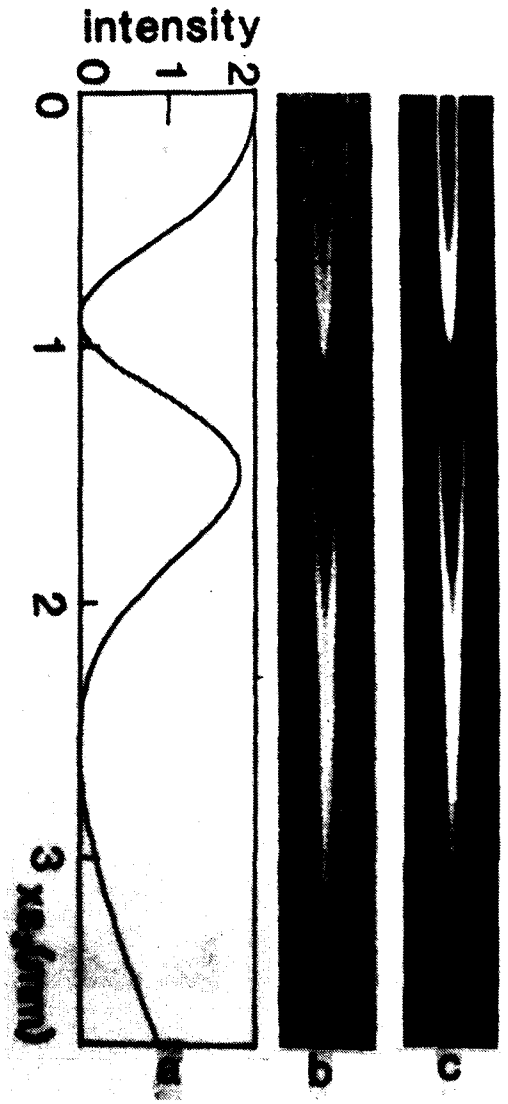



Fig.6 Comparison between intensity profile, experimental image and simulated image. a) intensity profile $I(xs_3)$ using the values: $xs_1=2.07$mm, $\Phi=\pi/4$ and $u_0(xs_1,xs_3)=u_0(2.07,0)=1.35$Å; b) Experimental image where $xs_1=2.07$mm and $\Phi=\pi/4$; c) simulated image performed with the same values as in (a).

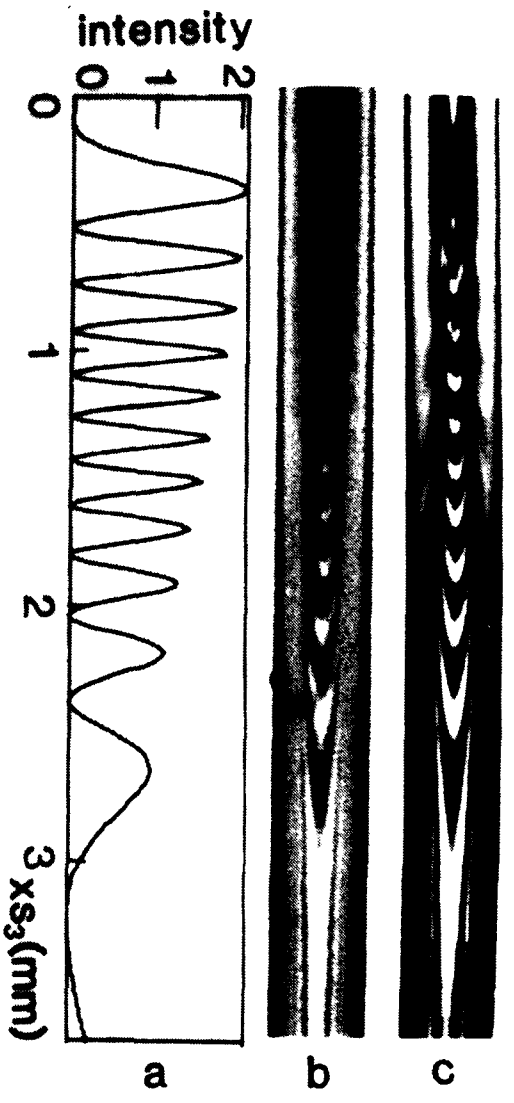



Fig.7 Comparison between intensity profile, experimental image and simulated image. a) intensity profile $I(xs_3)$ using the values: $xs_1=1.11$mm, $\Phi=0$ and $u_0(xs_1,xs_3)=u_0(1.11,0)=4.80$Å; b) Experimental image where $x_s=1.11$mm and $\Phi=0$; c) simulated image performed with the same values of $xs_1$ and $\Phi$ as in (a) and a corrected value $u_0(xs_1,xs_3)=u_0(1.11,0)=5.00$Å.

the intensity profile. This demonstrates that either Kato's theory or Takagi's theory can lead to a precise value of $u_0$ for the present case.

A second experimental image with a higher vibration amplitude is shown in Figure 7b where $xs_1=1.11$mm and $\Phi=0$. The estimated value of the vibration amplitude at the top part of the section topograph, $u_0(xs_1,xs_3)=u_0(1.11,0)$, was about 4.30Å. As previously, the determination by the intensity profile according to the Kato's theory leads to: $u_0(xs_1,xs_3)=u_0(1.11,0)=4.80$Å (Fig.7a). The simulation (Fig.7c) was performed by resolving the Takagi's equations and it was found that the image with a corrected value $u_0(xs_1,xs_3)=u_0(1.11,0)=5.00$Å reproduced the best experimental fringe positions. Small difference exists between the experimental image and the two calculations for the upper part of the section topography where the vibration amplitude is high (about 5Å). This is probably due to the experimental conditions not taken into account (saturation effect of the recording film, diffusion effect and polychromatism of the incident beam). Some asymmetry in fringe patterns with respect to the central line can also be observed in the experimental and simulated images. This fact is

mostly pronounced on the experimental image. This asymmetry may suggest the beginning of the vibration amplitude range where the Kato's theory becomes unavailable. It is recalled that the validity range previously estimated is $u_o \leq 7$Å.

These two examples show that extra Pendellosung fringes can be understood and analyzed by Kato's theory for a range of low vibration amplitudes. In particular, low vibration amplitudes can be determined from extra fringes in section topographs. The precision of such calculation must be discussed. It is found for the present case that an error less than 0.3Å concerning the value of $u_o$ may be reached and that corresponds, in terms of fringe positions, to the spacing between one black fringe and one white fringe on the images. The stroboscopic X-ray section topography appears to be a precise method to determine low vibration amplitudes. The knowledge of vibration amplitudes over the resonator is important, because it permits to obtain resonator parameters experimentally such as the amplitude attenuation factor $\beta_5$ and the quality factor Q.

## Stroboscopic X-ray traverse topography

Stroboscopic X-ray traverse topographs were performed for a vibration amplitude range similar to previously. A typical image is shown in Figure 8 where the estimated value of $u_o$ is 6.2Å. "Circular contrasts" can be seen and they are perturbed by two dislocations indicated by two arrows.

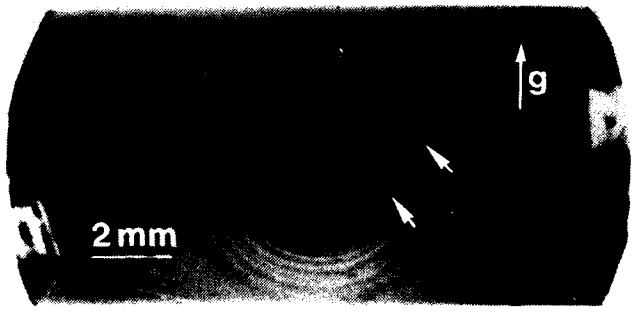



Fig.8 Stroboscopic X-ray traverse topograph made with $\mathbf{g}=2\bar{1}0$ where dislocations are indicated by arrows.

The understanding of these contrasts is very important, although a theoretical calculation seems to be very difficult for the present case because of the long calculation time required. Traverse topographs can be understood as an addition of section topographs over the whole crystal, so the "circular contrasts" probably result of the extra fringes observed in section topographs. As it was shown that the extra fringe positions are determined by the local vibration amplitudes, the "circular contrasts" represent most probably the equal-amplitude curves of the vibration. This point is supported by the two following remarks. Firstly, for a same excitation voltage, the number of "circular contrasts" on traverse topographs is found to be the same as the number of extra fringes on corresponding section topographs. Secondly, these "circular contrasts" have a slight ellipticity which corresponds to the elliptic equal-amplitude contour from the piezoelectric theory. It is recalled that the equal-amplitude curves are defined by

$$u_o \exp\{-\alpha_5 x_1^2 - \beta_5 x_3^2\} = \text{constant} \quad \text{or}$$

$$\alpha_5 x_1^2 + \beta_5 x_3^2 = \text{constant} \qquad (13)$$

and the ratio between the two axial lengths of the ellipses is then equal to $\sqrt{\alpha_5/\beta_5}$ (=0.935). This value was actually measured on the experimental images.

The fact that the "circular contrasts" can be related to the vibration amplitudes provides an another method to study the vibration in the resonator. It is noted that in comparison with section topographs, we have a bidimensional image of the vibration amplitudes. The characterisation of "circular contrasts" on traverse topographs in relation with vibration amplitudes will be carried out. It is expected that the determination of equal-amplitude curves on a resonator is very useful for the resonator study. The perturbation of such contrasts by dislocations is also a method to study the influence of dislocations on acoustic vibrations.

## Conclusion

The examples given here show that the technique of stroboscopic X-ray topography is a powerful method to study the vibration in resonators. It permits to reveal the existence of particular vibration modes which could be one origin of energy dissipation in the resonator. For low vibration amplitudes, section topography through extra Pendellosung fringes and traverse topography through "circular contrasts" can characterize precisely vibration amplitudes. Further studies will be carried out in order to characterize the vibrations on traverse topographs and to understand the influence of crystal defects such as dislocations on the performance of the resonators.

**Acknowledgement:** The authors wish to thank Dr. Epelboin whose assitance permitted the performance of the simulation images using his program DEFV. This work was supported by the DRET through the contract no.8734138004707501

## References


[1] Zarka,A., Capelle,B., Detaint,J. and Schwartzel,J. (1987). Proceeding of the 41th Annual Frequency Control Symposium. Philadelphia. U.S.A. p. 236.

[2] Tuomi,T., Naukkarinen,K. and Rabe,P. (1974). Phys. Stat. Sol.(a) 25, 93.

[3] Tiersten,H.F. (1976). J. Acoust. Soc. Am. 59, 879.

[4] Tiersten,H.F. and Smythe,R.C. (1979). J. Acoust. Soc. Am. 65, 1455.

[5] Gluer,C.C., Graeff,W. and Moller,H. (1983). Nucl. Inst. & Meth. 208. 701.

[6] Takagi,S. (1962). Acta. Cryst. 15, 1311.

[7] Takagi,S. (1969). J. Phys. Soc. Jap. 26, 1239.

[8] Taupin,D. (1964). Bull. Soc. Fr. Minéral. Crist. 87, 469.

[9] Penning,P. and Polder,D. (1961). Philips Res. Rep. 16, 419.

[10] Kato,N. (1964). J. Phys. Soc. Jap. 19, 971.

[11] Kato,N. (1974). Continuously deformed crystals. In: X-Ray Diffraction, pp.389-434, edited by L.V.Azaroff *et al.*, International series in pure and applied physics, McGraw-Hill Book Company.

[12] Zheng,Y., Zarka,A., Capelle,B., Detaint,J. and Schwartzel,J. Submitted for publication to Acta Cryst.

[13] Stevens,D.S. and Tiersten,H.F. (1986). J. Acoust. Soc. Am. 79, 1811.

# PROBE EXAMINATION OF THICKNESS-SHEAR VIBRATIONS OF AT-CUT NATURAL QUARTZ CRYSTALS: SOME NEW RESULTS

## Introduction

Travelling probe examination of acoustic vibrations on quartz crystal surfaces has been developed by Koga and Fukuyo[1] and van Dyke[2]. Apart from the scanty use of the technique in literature, its use by large, has been restricted to quartz crystals all covered with electrodes, a slot-being cut in the grounded electrode to contain the moving probe. In this paper, the probe technique of Koga and Fukuyo has been used to explore the nature of surface vibrations on unelectroded parts of AT-cut (natural) lapped and etched quartz crystals of dimensions 38 x 28 x 0.8 $mm^3$, (unless otherwise stated) the length of the crystals being along the X-axis (direction of thickness-shear vibrations), the crystals having electrodes of two shapes (rectangular and circular) and dimensions (13 x 10 $mm^2$ and 1.2 cm diam.). As such, the effectiveness of energy trapping apparently taking place in such crystals (vibrating in the fundamental mode) has been investigated and additionally the existence of an asymmetrically located dormant line on the quartz crystal resulting from change of phase of vibrations has been detected.

## Experimental

Since the unelectroded exposed surface of the crystal had to be examined, care was taken to fashion the probe such that the stray r.f. pick-up was negligible or small enough at almost all places of the crystal. About 1 mm of the exposed (unshielded) probe was found practical for the purpose, care being taken to have the probe tip almost touching the quartz crystal. The resulting stray pick-up was not more than 1 per cent of that at the central live electrode. By the simple expedient of use of a double-trace oscilloscope, the phase of surface vibrations could be compared to that of the vibrations below the central electrode. The results obtained by probe examination would be described below under the headings of "Energy Trapping" and "Phase-reversal".

## Energy Trapping

Energy trapping has been discussed in a number of publications[3,4,5] and so space would not be taken to review it here. To find out whether the usually employed technique for obtaining energy trapping (use of thin film patch electrodes) was really operative or not for obtaining the desired energy trapping, probe pick-up using vapor-deposited thin film patch electrodes of varying thicknesses and shapes was compared with the pick-up obtained when detatchable patch electrodes of corresponding shapes and dimensions were loosely placed on the quartz crystals.

Note. The probe-pick-up on any one half-length of the crystal was used for the investigation. The reason for this would be clear when discussing "Phase-reversal" below.

The effectiveness of energy trapping could be straightforwardly checked by determining the nature of fall of acoustic pick-up with distance on the bare surface of the crystal starting from the nearest edge of the central electrode. For energy trapping to hold, the fall of voltage for the thin film electrodes should be exponential or almost so, so that hardly any pick-up should remain after some short-distance from the central electrode. Figs. 1(a) and 1(b) depict the results for use of thin film and loose electrodes. It can be seen that both for use of thin film and loose electrodes the rate of fall of voltage (acoustic pick-up) is almost linear. Also, the pick-up remained almost finite uptil the edge of the length of the crystal.

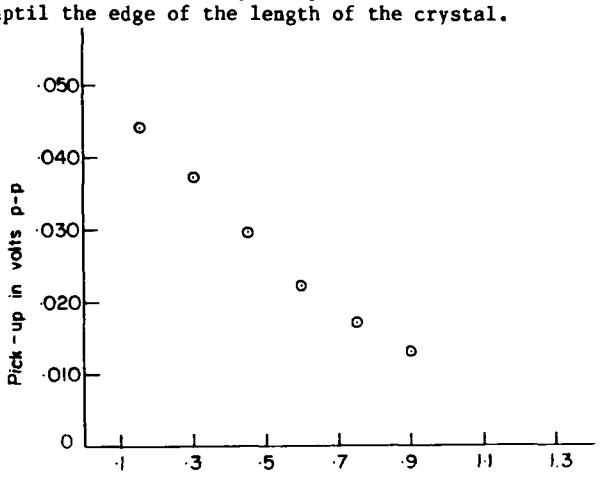



Fig 1(a). Change of acoustic pick-up on vibrating crystal surface with distance from nearest edge of rectangular film electrode of thickness 10,000 Å.

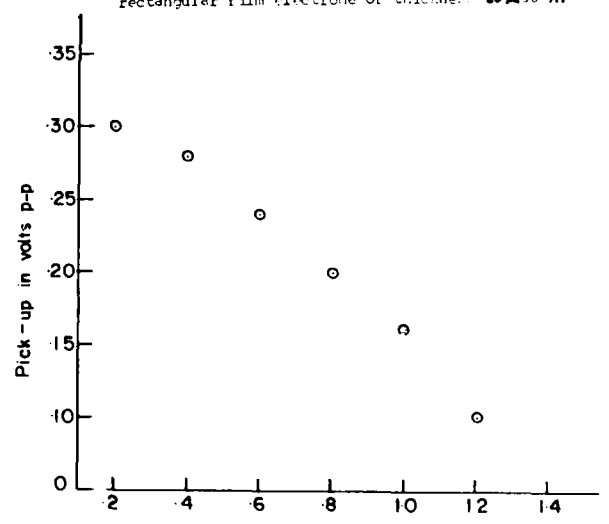



Fig 1(b). Change of acoustic pick-up on vibrating crystal surface with distance from nearest edge of detatchable rectangular electrode.

Another way of checking for energy trapping would be by comparing the ratio of voltages at the central electrode and at a point on the unelectroded surface just outside the electrode (2 mm in our observations) for the use of the two kinds of electrodes (thin film and detatchable). The ratios obtained were, using rectangular electrodes, 80, 130 and 225 for film thicknesses 2000 Å, 3500 Å and 10,000 Å respectively and 50 and 25 for loose electrodes, using rectangular and circular electrodes respectively.

Discussion. The linear slope of fig (1) denotes that energy trapping is not effectively taking place. On the other hand the higher voltage ratios for thin

CH2588-2/88/0000-093 $1.00 C 1988 IEEE

film electrodes indicate greater confinement of vibrational energy (by the square of the ratios) under these electrodes (thin film) than under the loose electrodes in the way demanded by the energy trapping theory. At this stage we are not in a position to state whether the greater energy concentration (about 2,5 and 16 times for the film electrodes of increasing thickness) is in quantitative accordance with energy trapping theory or only qualitatively so.

The present work suggests renewed investigation of energy trapping.

Before closing the topic, it should be stated that though greater film thickness desirably yields greater concentration of vibrational energy, yet the use of thick enough films is not an unmixed advantage since the vibrational damping also tends to increase with increasing thickness (as indicated by the observed decreasing output voltage for increasing thickness of the electrode for unchanged circuit settings). This fact also suggests further investigations in the matter.

## Phase-reversal

Fig. 2 represents the nature of pick-up obtained on the upper surface of a quartz crystal in certain directions having a centrally placed detatchable live electrode (diameter 1.2 cm, mass 5 gms) on it. The length of the arrows in the figure represent the relative magnitude of amplitudes and the direction of arrows the relative phases of vibrations. The existence of a line depicted by dots on one half-surface of the crystal indicates absence of vibrations and the site for change of phase. The location of the dormant line varied in different crystals but typically the line was about 2 mm. away from the nearest edge. The amplitude distribution on the two half-widths of the crystal was asymetrical, the lower amplitudes occurring on the half-width containing the dormant line.

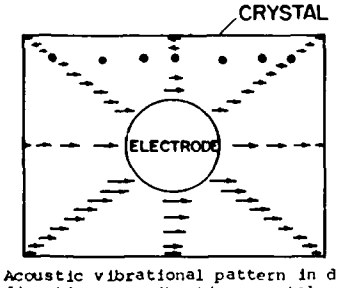



Fig 2. Acoustic vibrational pattern in different directions on vibrating crystal surface.

By reversing the crystal by the two alternative rotations of the crystal about the length and the width and examining different crystals for such operations, it was established that the dormant line occurred on that half-surface towards which the Y-axis was inclined by the acute angle 54° 75' (90° - 35° 15'). Pertinently, the dormant lines on the opposing surfaces were not located opposite.

Note. The particulars mentioned above existed for rectangular electrode shapes also and for both the kinds of electrodes, thin film and detatchable.

Commercial circular synthetic quartz frequency resonators having centrally placed circular thin film electrodes also exhibited regions of opposite phases of vibration, the phase-reversal taking place about a diametrical line.

The following additional observations were also made on rectangular crystals regarding the existence of dormant lines.

(1) With the movement of the patch electrode towards or away from the dormant line, the dormant line also followed in the direction of movement but by a much reduced displacement (in the ratio of about 6:1).

(2) As the electrode was put near the edge, on the dormant line, no phase-reversal was produced anywhere, the amplitude of vibration in the region containing the previous dormant line decreasing in the process.

(3) By shifting the patch electrodes together along the central length away from the center, regions of phase-reversal along one half-width still appeared.

(4) By completely covering one or both of the surfaces of the crystal with electrodes, lines of phase-reversal vanished. In case both surfaces were covered holes were spaced in the upper electrode for sampling the pick-up. The latter observation is in accord with that in literature[6].

(5) Using crystals with polished but subsequently unetched surfaces, new lines of phase-reversal appeared, the older intrinsic lines getting affected on some surfaces. The placing of the new lines on the same or different crystals had no correlation among them. By appropriate etching of the crystals, the new extra lines disappeared, the older intrinsic lines remaining as such or reappearing on different surfaces.

Discussion: At the present, the existence of the line of phase-reversal is obscure. Apparently, its asymmetry of placement can be associated with asymmetry of inclination of the Y-axis of the crystal. Observations with polished but unetched crystals suggest that surface effects are playing some yet undefined role in the phenomenon of phase-reversal.

## References

1. Koga, I and Fukuyo, H. (1953), J. Inst. Elec. comm. Eng. Japan, 36, 59.

2. van Dyke, K.S. (1957) Proc. Ann. Freq. Control Symp, vol 11.

3. H.Bahadur and R. Parshad, Phys. Acoustics, 16, 37, 1982 (Edit W.P. Mason and R.N. Thurston, Acad. Press, New York).

4. Shockley, W, Curran, D.R. and Konneval, D.J. (1963), Proc. Ann. Freq. Control Symp. 17, 88.

5. Shockley, W, Curran, D.R. and Koneval, D.J. (1963), Proc. Ann. Freq. Control Symp. 17, 88.

6. Koga, I. (1963), J. Appl. Phys. 34, 2357.

42nd Annual Frequency Control Symposium - 1988

# NEW MONOLITHIC CRYSTAL FILTER WITH WIDE TABS

Yukihiro OKAMOTO
Nihon Dempa Kogyo Co., Ltd., Tokyo, 151 Japan

Hitoshi SEKIMOTO
Faculty of Technology, Tokyo Metropolitan University, Tokyo, 158 Japan

## Abstract

It is very difficult to achieve both the low impedance level and the good spurious response characteristic with monolithic crystal filters (MCFs). The reason is that the large electrode area of resonators may cause strong anharmonic spurious modes to fall near the passband. In this paper, a new MCF with wide tabs is proposed. The wide tab electrodes are utilized for effectively waveguiding the energy of anharmonic modes to lossy mounting at plate edges. A important parameter of designing tab-waveguides of finite width is the cutoff frequency. The effect of Z'-tabs is examined on dual resonator MCFs operating in the third overtone thickness-twist mode. The theoretical and experimental results show that while the unwanted mode suppression achievable with the Z'-waveguides is limited by the occurrence of untrapped mode, it is possible to attenuate both the strong (3,1,3) and newly excited modes below - 20 dB. Experimental results for MCFs having wide Z'- and X-tabs are also presented.

## I. Introduction

Monolithic crystal filters (MCFs) are a class of bandpass filters using coupled mechanical resonators, which are fabricated on a single quartz plate. The available impedance level is limited by the acoustical properties of resonators. Any inductance is achievable, at least theoretically, by properly choosing the electrode area of resonators. That is not the case, however, when the very practical problem of achieving good spurious response characteristics is considered [1]. The reason is that the large electrode area may cause strong anharmonic spurious modes to fall near the passband. This limitation on the available impedance level still remains to be solved.

In this paper, a new MCF with wide tabs is proposed as a solution to the above problem. The tabs are utilized for effectively waveguiding the energy of trapped spurious modes, which otherwise has been confined in the resonator region, to lossy mounting at plate edges. Horton and Smythe [2] were the first to discuss such application of wide tabs to the suppression of spurious modes. However, the effects of tabs were only qualitatively understood, since the particular cutoff frequency associated with a tab of finite width was not recognized. Here, we introduce the more exact conception of "tab-waveguide" into the design of MCFs. The most important parameter of designing a tab-waveguide is the cutoff frequency, which is necessarily higher than that of an infinite plate fully electroded on one surface only. From the results for a resonator given by the authors [3], it is not difficult to expect that when an anharmonic trapped mode having phase reversals in the tab direction is at a frequency above the tab-cutoff, most of the energy will be guided along the tab and dissipated by the lossy mounting.

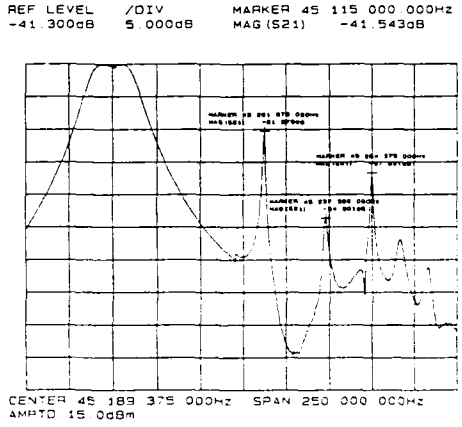



Fig.1 Measured spurious response in a 45 MHz dual resonator MCF. The electrode dimensions of the individual resonators are 0.8 mm (Z')*1.5 mm (X), the coupling gap between two resonators is 0.32 mm (Z'), and the width of Z'-tabs is 0.4 mm.

The effects of strip tabs extending along the Z' axis in AT-cut rectangular plates are examined on dual resonator MCFs, operating in the third overtone thickness-twist mode at about 45 MHz. Each of tabs is connected through a narrow lead, or directly, with a resonator. Both the theoretical and experimental results prove that the strongest first spurious (3,1,3) mode can be greatly suppressed by using the waveguiding effect of Z'-tabs, while the passband characteristics are unaffected. Furthermore, it is shown that although novel untrapped modes caused by the existence of wide Z'-tabs occur, both the (3,1,3) and newly excited responses can be adjusted to 20 dB less than the main ones. On the basis of these results, MCFs with wide Z'- and X-tabs are tentatively designed. The experimental results are also reported.

## II. Dual Resonator MCFs with Wide Z'-tabs

The spurious response of a 45 MHz dual resonator MCF is shown in Fig.1. This MCF was fabricated, according to the conventional design using narrow tabs, on a circular AT-cut quartz plate, and operated in the third overtone thickness-twist mode. The electrode dimensions of the individual resonators were 0.8 mm (Z')*1.5 mm (X), the coupling gap between two resonators was 0.32 mm (Z'), the width of Z'-tabs was 0.4 mm and the mass loading was set about to 0.0042. The relatively large electrode area of resonators, which yields an inductance of about 50 mH, resulted in a strong spurious resonance near the passband. The response is only 10 dB less than the main responses. We have examined the waveguiding effect of Z'-tabs on this mode, that is the anharmonic (3,1,3) mode having two phase reversals in the Z' direction.

CH2588-2/88/0000-095 $1.00 C 1988 IEEE

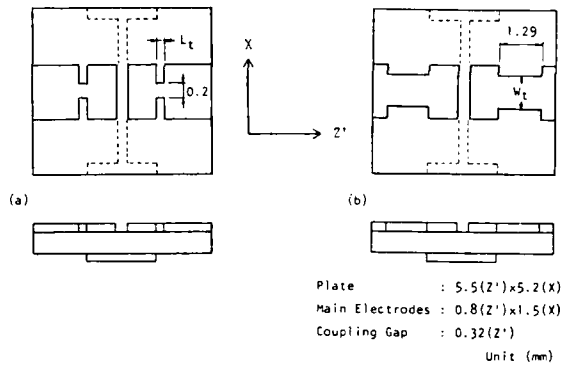



Fig.2 45 MHz dual resonator MCFs with two types of wide Z'-tabs. In the theoretical analysis, mechanical loads distributed along the Z' plate edges were included, and the narrow X-tabs were neglected.

Figs.2 (a) and (b) show the dual resonator MCFs with two types of wide Z'-tabs, on which the effects of tabs were investigated. These MCFs were fabricated on rectangular AT-cut plates of 5.5 mm (Z')* 5.2 mm(X), with the same resonator dimensions as in the MCF of Fig.1. The cutoff frequencies $f_0'$ and $f_0$ of infinite plates without and with full electrodes on both surfaces were about 45.25 MHz and 45.045 MHz, respectively, and the cutoff $f_t$ of an infinite plate fully electroded only on one surface was about 45.155 MHz. In type (a), where Z'-tabs of 1.5 mm width were connected with resonators through the lead electrodes of 0.2 mm width, the cutoff frequency $f_g$ of the tab-waveguides was set about to 45.173 MHz below the resonant frequency of the (3,1,3) mode. Accordingly, the wavequiding effects were examined as a function of lead-length $L_t$. In type (b), where Z'-tabs were connected directly with resonators, the variable parameter is the tab-width $W_t$ which determines the tab-cutoff $f_g$.

The measured spurious responses for types (a) and (b) are shown in Figs.3 and 4, respectively, with the calculated ones. The calculations have been done using a two-dimensional model with mechanical loads distributed along the Z' plate edges. This theoretical model is an extension of previous models [3],[4]. The mechanical loads were included to estimate a dissipation of lossy mounting. Since the acoustical properties of mounting were unknown, only a purely conductive value for loads was assumed in order to all in all fit the calculated characteristics to the measured ones. In addition, the narrow X-tabs were neglected. However, these calculated results have verified the measured results; the (3,1,3) mode can be suppressed by using the waveguiding effect of Z'-tabs, the passband characteristics are unaffected, and the novel untrapped modes caused by the existence of wide Z'-tabs occur.

From the results in Figs.3 and 4, it is apparent that as long as $f_g$ is below the frequency of the (3,1,3) mode designated as $S_1$, the Z'-waveguides can suppress greatly this anharmonic resonance. Furthermore, no influence of the waveguides on the passband characteristics can be found because $f_g$ is set above the passband. However, these results also show that the unexpected modes can be strongly excited by the wide tabs. The novel resonances designated as $A_1$ and $S_2$ become stronger with decreasing $L_t$ or increasing $W_t$, while the $S_1$ resonance becomes weaker. This is illustrated in Fig.5, where the measured levels for the $S_1$ mode and the stronger one of $A_1$ and $S_2$ modes are plotted with the calculated curves. Although the levels of $A_1$ and $S_2$ modes cannot be well approximated by the present theoretical model, the calculated curves show tendencies similar to the measured data. Accordingly, it should be noted that the unwanted mode suppression achievable with the Z'-waveguides is limited by the occurrence of $A_1$ and $S_2$ modes. Nevertheless, it is possible to attenuate both the (3,1,3) and newly excited modes below - 20 dB, This level corresponds to the additional 10 dB suppression of the (3,1,3) mode, when compared with the response shown in Fig.1.

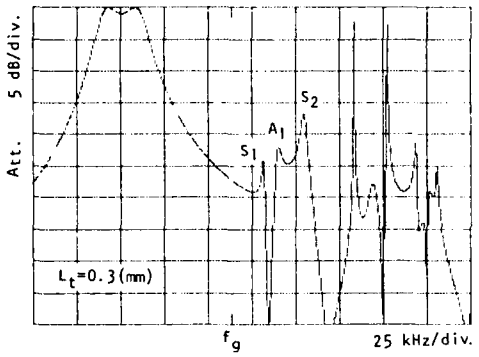
Three spurious resonances discussed above were identified by the calculation of vibrational fields, as symmetric $S_1$ and $S_2$ modes and an anti-symmetric $A_1$ mode. Fig.6 shows the patterns along the Z' axis in the MCF of type (a) with $L_t$=0.3 mm. The field of each mode parallel to the X axis was nearly the same as that of the main (3,1,1) mode. It is apparent that the $S_1$ mode belongs to the trapped (3,1,3) mode, and most of the energy is guided near the Z' edge. This verifies that the sufficient waveguiding effect can be obtained even from the wide tabs of finite length. On the other hand, the $A_1$ and $S_2$ modes obviously belong to the untrapped ones newly excited by the Z'-tabs because such virational patterns would be never present in MCFs without tabs. Therefore, we must take into account that the occurrence of untrapped modes is unavoidable in MCFs with wide tabs, fabricated on finite plates.

The discussion up to this point has been concerned with only the suppression of (3,3,1) mode. Since the tab-waveguides can affect the modes with phase reversals in the tab direction, Z'-tabs have little effect on the (3,3,1) and (3,3,2) modes. Hence these modes, which correspond to the higher frequency resonances in Figs.3 and 4, remain strong. Their suppression will require the X-tabs.

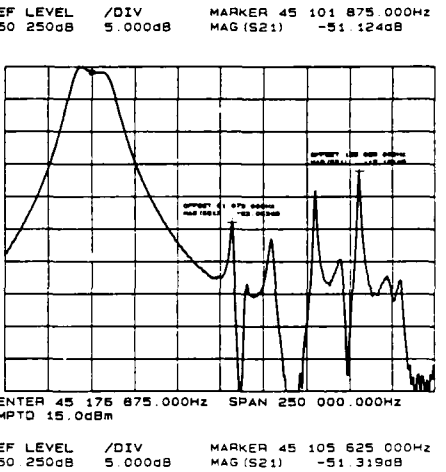
## III. MCFs with Wide Z'- and X-Tabs

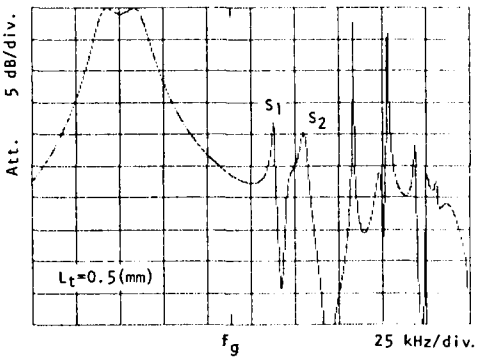
The results in the preceding section suggest that all of the spurious responses near the passband may be attenuated by using wide X-tabs as well as Z'-tabs. Fig.7 shows an attempt to improve the spurious characteristics in 82 MHz MCFs whose resonator inductances are about 20 mH. In this case, the spurious responses were all in all reduced to -19 dB. A comparison with the spurious responses of the current MCFs indicated that the utilization of wide Z'- and X-tabs provided an additional 10 dB suppression. However, we could not distinguish the effect of X-tabs in the spurious responses, since a theoretical model including both the Z'- and X-tabs has not been developed. We only presumed that the (3,3,1) and (3,3,2) modes would be also greatly attenuated. Another example of a 10.7 MHz MCF is shown in Fig.8, with a photograph of the electrode structure. This MCF was fabricated on a circular plate with 8 mm diameter, and operated in the fundamental mode. In spite of the large resonator area, the spurious responses could be attenuated below -23 dB.

Although these MCFs were only tentatively designed for an application in the cascaded form of two sections, they satisfied sufficiently the requirement of both the low terminal impedance level and the low spurious level. With further development, it should be possible to achieve the better spurious response characteristics.

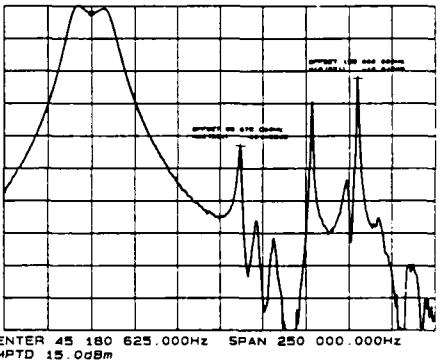


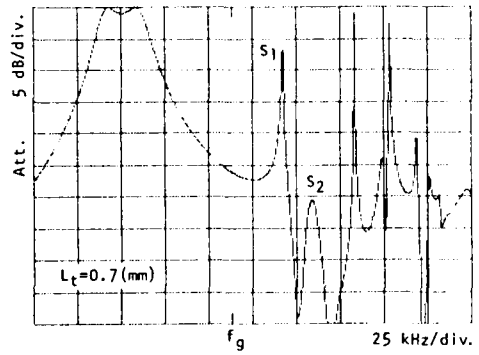

Lt=0.3(mm)



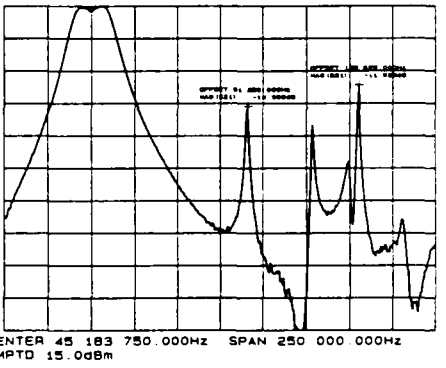



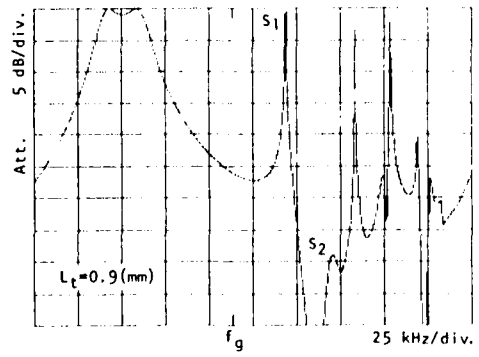

Lt=0.5(mm)



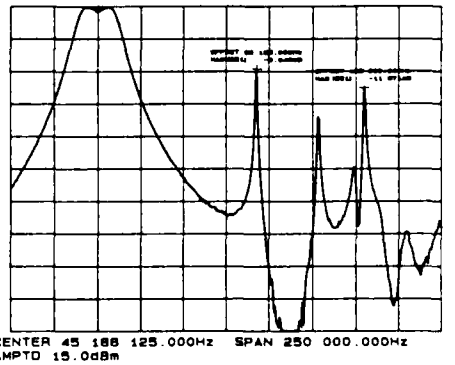




Lt=0.7(mm)



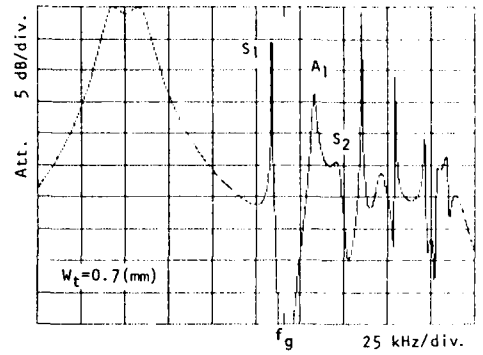

Lt=0.9(mm)




Calculated(passband center freq.=45.109Mhz, $f_g$=45.173Mhz)

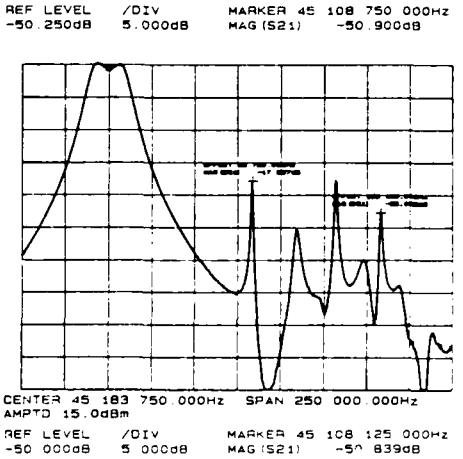
Measured

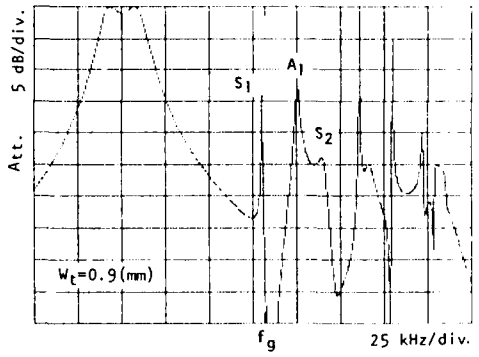
Fig.3 Calculated and measured spurious responses for 45 MHz MCFs with wide Z'-tabs of type (a).




$f_g$=45.199Mhz

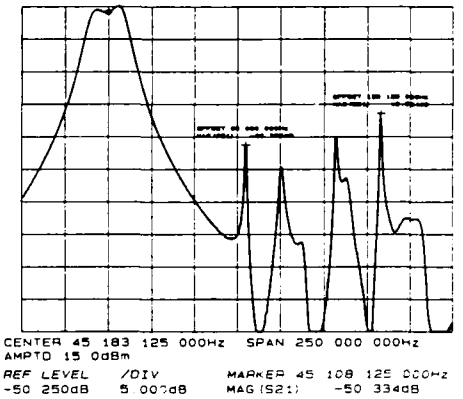
$W_t$=0.7(mm)

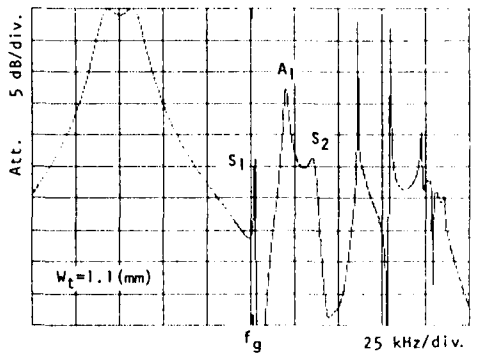



$f_g$=45.189Mhz

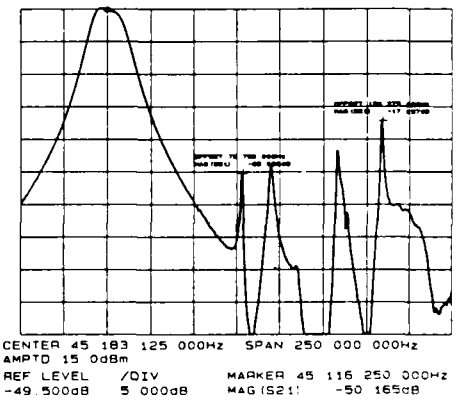
$W_t$=0.9(mm)

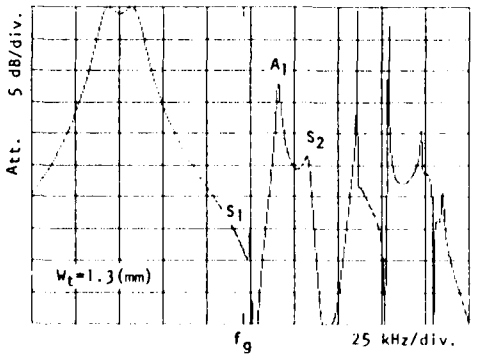



$f_g$=45.182Mhz

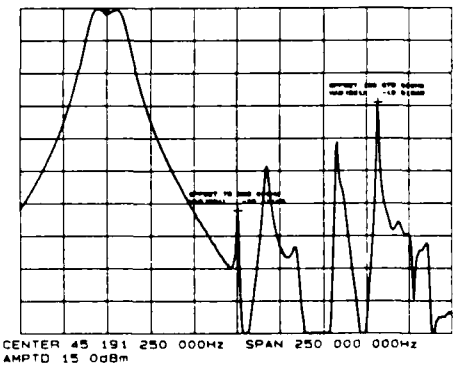
$W_t$=1.1(mm)




$f_g$=45.177Mhz

$W_t$=1.3(mm)




Calculated(passband center freq.=45.106Mhz)

Measured

Fig.4 Calculated and measured spurious responses for 45 MHz MCFs with wide Z'-tabs of type (b).

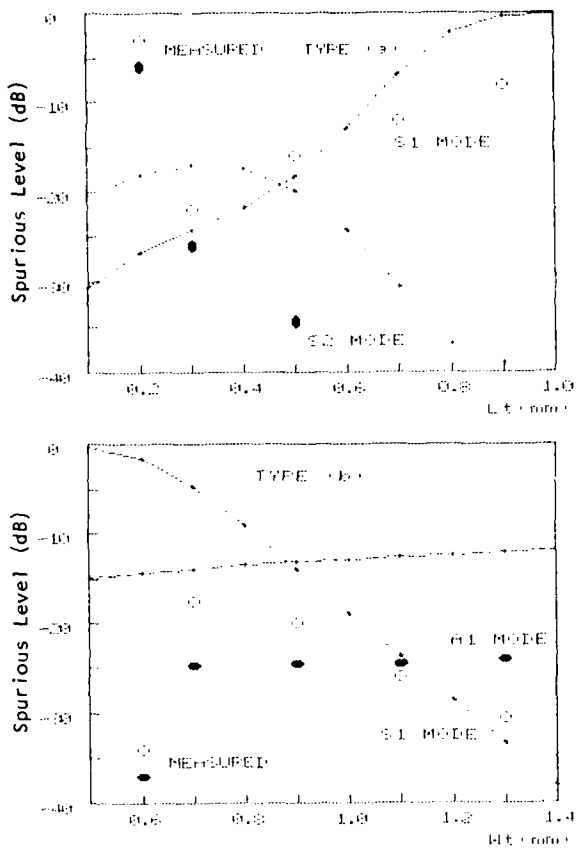



Fig.5 Calculated and measured levels for the $S_1$, $A_1$ and $S_2$ modes.

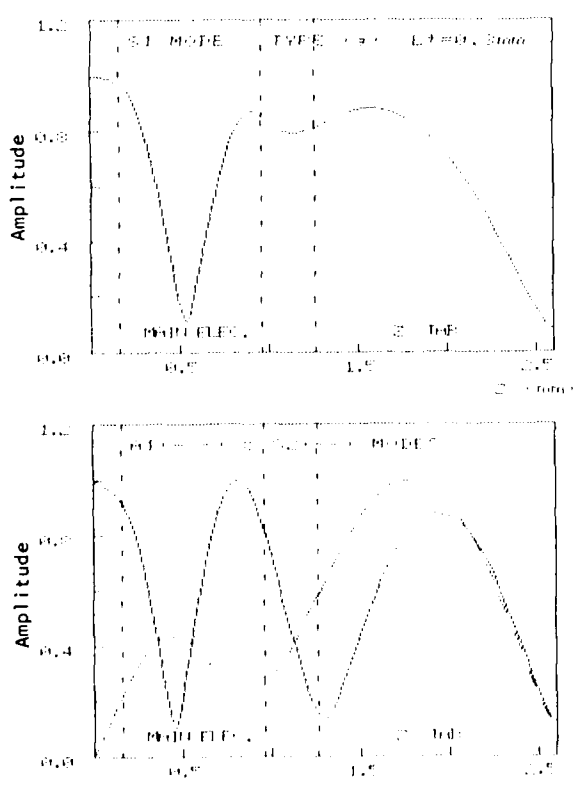



Fig.6 Calculated vibrational patterns along the Z'-axis in the MCF of type (a) with $L_t$=0.3 mm.

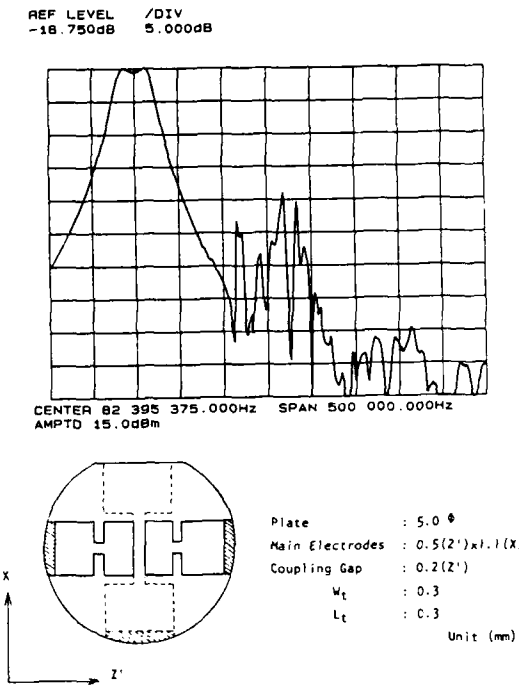



Fig.7 Measured spurious response for a 82 MHz dual resonator MCF with wide Z'- and X-tabs.

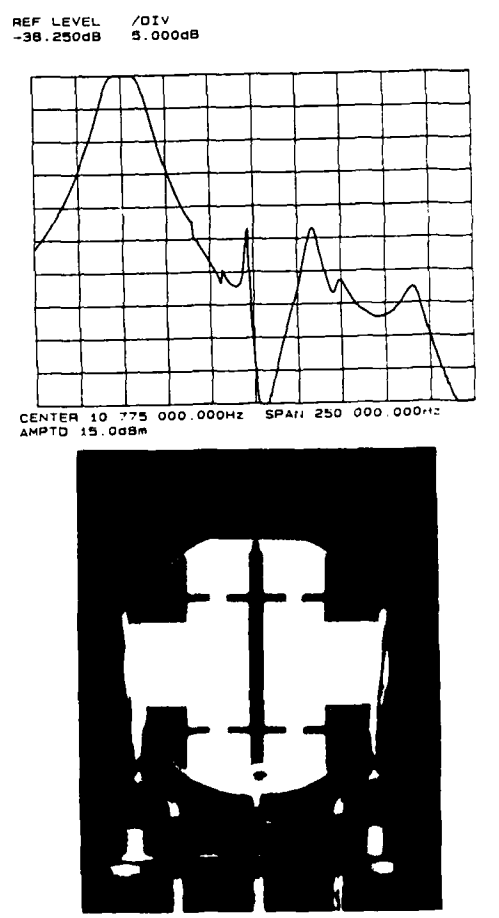



Fig.8 Measured spurious response for a 10.7 MHz dual resonator MCF with wide Z'- and X-tabs and the photograph of electrode structure. The MCF was fabricated on a circular plate with 8 mm diameter, and operated in the fundamental mode.

## IV. Conclusion

In this paper, new MCFs based on the conception of tab-waveguides were proposed. The wide tabs were utilized for effectively waveguiding the energy of trapped spurious modes to lossy mounting at plate edges. The effect of Z'-tabs was examined on dual resonator MCFs operated in the third overtone mode. The calculated and measured results showed that the most important parameter of designing the tab-waveguides of finite width was the cutoff frequency. The strongest (3,1,3) mode was greatly attenuated, as long as the tab-cutoff was set below the frequency of this mode. It was also shown that while novel untrapped modes were produced, the utilization of wide Z'- and X-tabs enabled us to suppress all of the spurious responses near the passband below -20 dB.

## References

[1] G.T.Pearman and R.C.Rennick, "Monolithic crystal filters," IEEE Trans. Sonics Ultrason., vol.SU-21, no.4, pp.238-243, Oct. 1974

[2] W.H.Horton and R.C.Smythe, "Theory of thickness-shear vibrations, with extensions and applications to VHF acoustically-coupled-resonator filters," in Proc. 21st Annu. Symp. Frequency Control, pp.160-178, 1967

[3] H.Sekimoto, H.Nakata and M.Miura, "Analysis of trapped energy resonators with tabs," IEEE Trans. Ultrason. Ferroelec. Freq. Contr., vol.UFFC-34, no.6, pp.674-680, Nov. 1987

[4] H.Sekimoto, "Two-dimensional analysis of monolithic crystal filters using mode-matching technique," Electron. Commun. Japan, vol.65, no.12, pp.10-19, 1982

42nd Annual Frequency Control Symposium - 1988

# SEMI-ANALYTICAL ANALYSIS OF COUPLED THICKNESS-SHEAR AND THICKNESS-TWIST IN MONOLITHIC CRYSTAL FILTERS

**Henri Carru and René Lefèvre**

**Centre National d'études des Télécommunications**
**196 av. H. Ravera 92220 Bagneux France**

**Jean Pierre Aubry and Serge Lechopier**

**Compagnie d'Electronique et de Piezoelectricité**
**95100 Argenteuil France**

## Summary

The wave equation, describing the displacement of rotated Y-cut quartz trapped energy resonators and two poles monolithic filters with rectangular electrodes and operating in fundamental and overtones of coupled thickness-shear and thickness-twist vibrations, due to H.F. Tiersten [1,2] is treated by a finite element method. Rectangular elements are used for element division of the surface of the plate both in the electroded and unelectroded regions and displacement at the corners of elements are taken as unknown. Natural boundary conditions imply that, in several cases, solution obtained differs from the analytical solution.

This method can be applied to the treatment of resonators and complex monolithic structures having N resonators ( with N >= 2 ) including the effect of size and position of resonator tabs. This analysis allows the design of monolithic filters both on AT-cut quartz and rotated Y-cut berlinite. Optimum designs for unwanted modes attenuation may be found.

## Introduction

Design of monolithic filters has to be more and more accurate to get low-cost devices especially in the following cases:

-High-frequency monolithic filters when electrode size is small compared with that of tabs.

-Bidimensional monolithic filters to introduce attenuation poles near the bandpass of the filter.

-Use of non rectangular electrodes to achieve high Q-factor or low spurious responses.

-Case of finite plates.

For these particular cases, purely analytical method established by H.F. Tiersten [1] is difficult to use because in structure with no constant electrode width, there are many corner-type regions for which only the $X_2$ dependence in the partial differential equation and boundary conditions on the major surfaces of the plate are satisfied [1]. In a same manner, finite element method applied to piezoelectrical equations set, gives a more complete description of phenomena but takes a large computer-time. Semi-analytical analysis that combines the two methods appears to be an efficient solution for the above mentioned cases.

## Principle of the semi-analytical method

For AT-cut quartz thickness-shear and thickness-twist modes devices, H.F. Tiersten has shown that the wave equation describing the displacement $U_1$ is :

$$M_n \frac{\partial^2 U_1}{\partial X_1^2} + C_{55} \frac{\partial^2 U_1}{\partial X_3^2} + A \times U_1 = 0 \qquad (1)$$

where

$$A = \rho\omega^2 - \left(\frac{n\pi}{2h}\right)^2 \times \hat{C}_{66} \qquad (2)$$

with

$$\hat{C}_{66} = \overline{C}_{66}\left(1 - \frac{8K^2}{n^2\pi^2} - 2R\right)$$

for a resonator

$$\hat{C}_{66} = \overline{C}_{66}(1 - 2R')$$

for an electroded region that is not a resonator.

$$\hat{C}_{66} = \overline{C}_{66}$$

for an unelectroded region.

R and R' are the plateback of these areas.

Considering a monolithic structure as shown on fig 1 and dividing a choosen domain (D) on it into rectangular elements, we get a mesh. Then we set displacement $U_1$ as a sum whose terms are composed of displacement values at each node of the mesh i.e. :

$$U_1 = \sum_j^n N_j(X_1, X_3) \times U_{1j} \qquad (3)$$

CH2588-2/88/0000-101 $1.00 C 1988 IEEE

where Nj is a product of two linear independent functions.

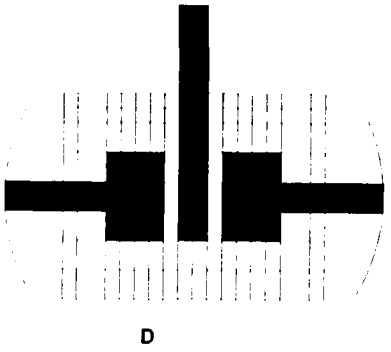



Fig 1: Monolithic cell divided into rectangular elements

To satisfy equation (1) a necessary and sufficient condition is:

$$I = \int\int_D N_i \left( M_n \frac{\partial^2 U_1}{\partial X_1^2} + C_{55} \frac{\partial^2 U_1}{\partial X_3^2} + \Lambda \times U_1 \right) dX_1 dX_3 = 0 \qquad (4)$$

Integral I may be split into two integrals:

$$I = I_1 + I_2$$

where

$$I_1 = \int_\Gamma N_i \left( M_n \frac{\partial U_1}{\partial X_1} \alpha_1 + C_{55} \frac{\partial U_1}{\partial X_3} \alpha_3 \right) d\Gamma$$

$$I_2 = \int \int_D \left( N_i A U_1 - M_n \frac{\partial N_i}{\partial X_1} \frac{\partial U_1}{\partial X_1} - C_{55} \frac{\partial N_i}{\partial X_3} \frac{\partial U_1}{\partial X_3} \right) dS = 0 \qquad (5)$$

where $\alpha_1$ and $\alpha_3$ are cosine directors of the normal to the contour $\Gamma$ of the considered domain (D).

The first integral $I_1$ is neglected, which is legitimate when the gradient of $U_1$ is low on a large $\Gamma$ ; when $\Gamma$ is a rectangle with sides parallel to $X_1$ and $X_3$ , it is equivalent to assume traction-free edges.

Substituting (2) and (3) into $I_2$ gives :

$$\sum_j U_{1j} \int\int_D \left( M_n \frac{\partial N_i}{\partial X_1} \frac{\partial N_j}{\partial X_1} + C_{55} \frac{\partial N_i}{\partial X_3} \frac{\partial N_j}{\partial X_3} + \left( \frac{n \cdot \pi}{2h} \right)^2 \hat{C}_{66} N_j N_i \right) dX_1 dX_3 =$$

$$\rho \omega^2 \sum_j U_{1j} \int\int_D N_j N_i dX_1 dX_3 \qquad (6)$$

Equation (6) has the following form:

$$[B] U_1 = \rho \omega^2 [D] U_1 \qquad (7)$$

where [B] and [D] are matrices.

Eigenvalues give the resonant frequencies of the structure, values of $U_{1j}$ are computed at each node to give $U_1$ over all the domain (D). Electrical equivalent circuit is then computed from Tiersten's equations.

The size of the domain has to be kept to the minimum necessary to give a fixed accuracy (unless it is enough subdivided,which may lead to prohibitive computing time); otherwise spurious oscillations and wrong frequencies appear[3].

Infinite plates or finite rectangular plates with rectangular electrodes and tabs imply a rectangular mesh easy to generate ( Input data are only dimensions and numbers of elements in each direction ).Infinite or circular plates with electrodes of arbitrary shape and slant tabs imply triangular mesh which needs a mesh generator.

## Computed and experimental results

We have applied this method to compute the two first main modes in some typical cases of monolithic cells both on AT-cut quartz and rotated Y-cut berlinite. Agreement is good between actual experimental device data and computer simulation results with tabs both for frequency modes and motional inductance of the first mode. In this example, see table 1, we have computed the variation of these frequencies when the plate is finite ( edges of the plate being at 1 mm from those of resonators ).

**TABLE 1**

Material: Quartz
Device: Two poles monolithic cell (from C.N.E.T.)
Unelectroded thickness frequency: 8192 KHz
Coupling: Thickness-shear
Electrode length(along the coupling axis): 2.24 mm
Electrode width: 4 mm
Inter-resonators spacing: 1.65 mm
Tab width(normal to the coupling axis): 1 mm
R(plateback of the resonators): 0.0107
Trimming strip: none

| | Exp. results | Sim. without tabs | Sim. with tabs |
|---|---|---|---|
| L(mH) | 53.9 | 52.5 | 52.6 |
| Fs | 8130.275 KHz | 8130.900 KHz | 8130.300 KHz |
| Fas | 8139.205 KHz | 8140.100 KHz | 8139.200 KHz |
| Fs | Finite plate | 8127.800 KHz | 8126.800 KHz |
| Fas | Finite plate | 8135.900 KHz | 8134.700 KHz |

As shown in this table, influence of tabs on frequency modes is not negligible while motional inductance remains practically constant. The contour curves of displacement $U_1$ of the first mode (symmetric mode) plotted on fig 2-A show distorsion brought by tabs.

Fig 2-B shows the same curves without tabs.

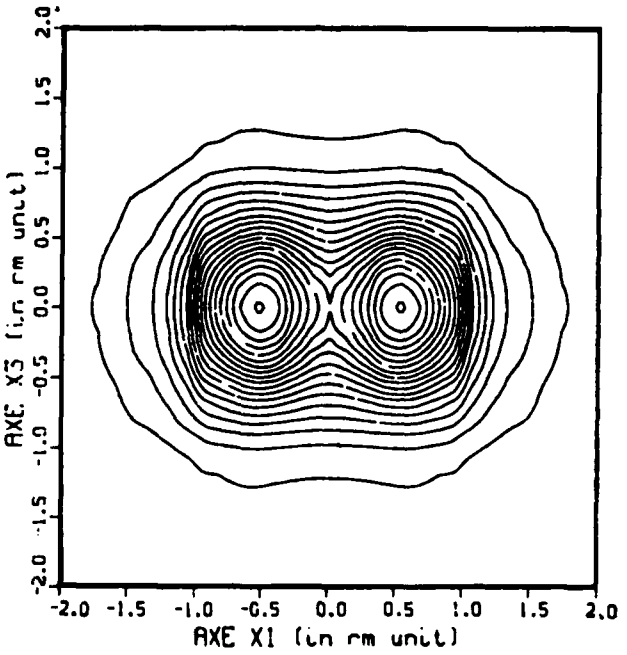



Fig 2-A:contour curves of displacement $U_1$ (with tabs)

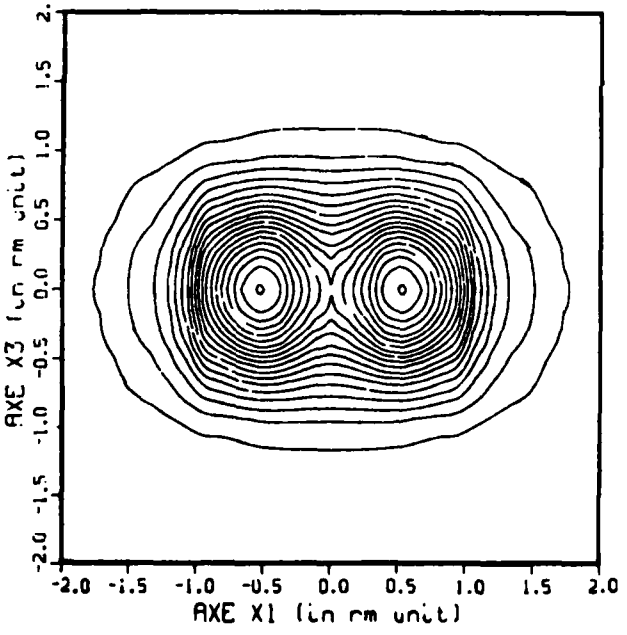



Fig 2-B:contour curves of displacement $U_1$ (without tabs)

In the table 2, we have computed frequency modes and $U_1$ for a quartz two-poles monolithic cell including a trimming strip in the inter-resonator spacing and tabs along the coupling axis. Practically the part of trimming strip out of inter-resonator spacing has a negligible effect du to its very low plateback.

**TABLE 2**

Material: Quartz
Device: Two poles monolithic cell (from C.E.P.E.) see fig 1
Unelectroded thickness frequency: 21791 KHz
Coupling: Thickness-twist
Electrode length(along the coupling axis): 0.9 mm
Electrode width: 1.45 mm
Inter-resonators spacing: 0.24 mm
Tab width(along the coupling axis): 0.2 mm
R(plateback of the resonators): 0.0103
R'(plateback of the trimming strip): 0.0051
Trimming strip length(along the coupling axis):0.08 mm

| | Exp. results | Sim. without tabs | Sim. with tabs |
|---|---|---|---|
| Fs | 21610.15 KHz | 21611.10 KHz | 21610.10 KHz |
| Fas | 21648.42 KHz | 21650.20 KHz | 21648.40 KHz |

In the fig 3 and fig 4, we have drawn contour curves and amplitude $U_1$ of the first two main modes. We can see that there is no distorsion of these curves by the trimming strip.

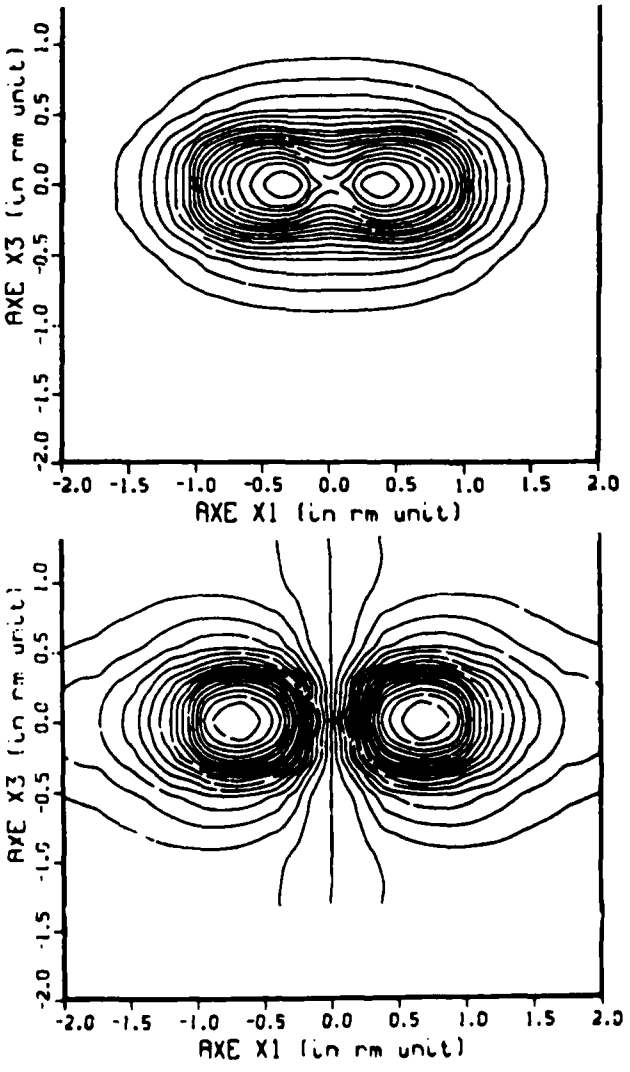



Fig 3: contour curves of displacement U1 of the two first main modes of a monolithic cell (with tabs)

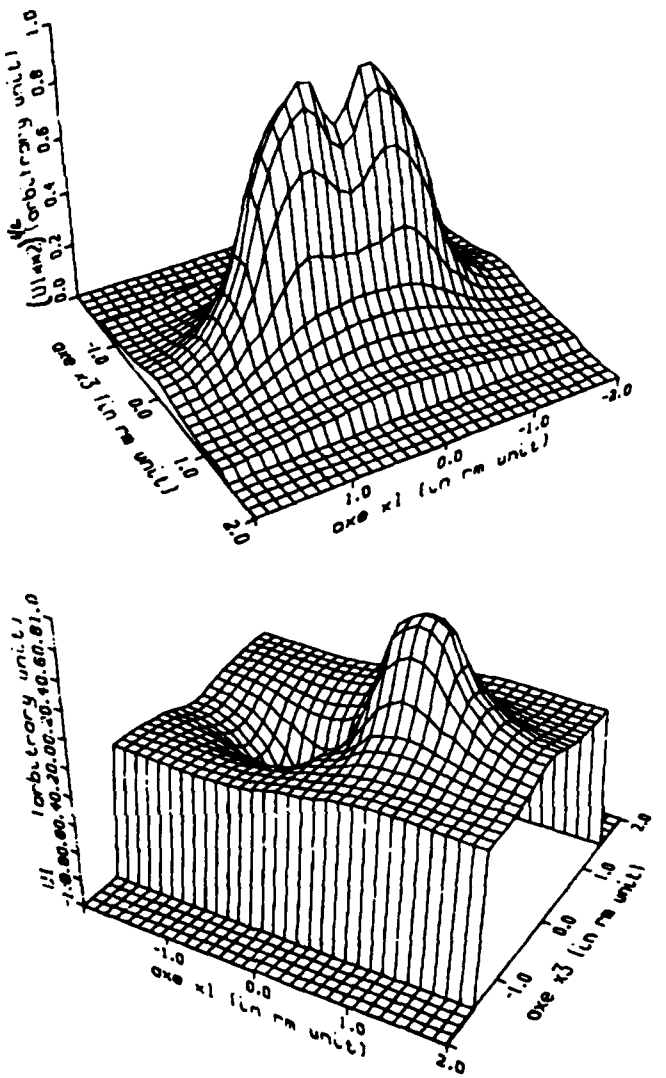


Fig 4: Displacement U1 of the two first main modes of a monolithic cell (with tabs)

Table 3 show analogue results on berlinite. In this table frequency differences between computed results and experimental results are due to a bad evaluation of the air-gap frequency.

| TABLE 3 | | |
|---|---|---|
| Material: Berlinite<br>Device: Two poles monolithic cell (from C.N.E.T.)<br>Unelectroded thickness frequency: 14846.86 KHz<br>Coupling: Thickness-shear<br>Electrode length(along the coupling axis): 0.8 mm<br>Electrode width: 1.5 mm<br>Inter-resonators spacing: 0.5 mm<br>R(plateback of the resonators): 0.0144<br>Trimming strip : none | | |
| | Exp. results | Sim. without tabs |
| L(mH) | 39.2 | 39.7 |
| Fs | 14712.19 KHz | 14714.80 KHz |
| Fas | 14756.21 KHz | 14759.20 KHz |

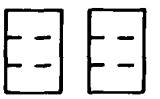
## Spurious modes considerations

In a previous paper [4,5], it has been established that, for a monolithic filter, unwanted anharmonic modes may be classified in two groups from resonator point of view : Quasi-symmetric modes (Q.S.) that have an even number of node lines in the resonators and Quasi-antisymmetric modes (Q.A.S.)that have an odd number of node lines in the resonators. When there is only one resonator, the Q.A.S. modes are not piezoelectrically active because the current, proportional to the integral of $U_1$ over area of resonator is equal to zero. In a monolithic filter, the other resonator causes an asymmetry then the amplitude of Q.A.S. modes is strongly dependent upon the coupling between the two resonators while that of Q.S. modes remains practically constant. In the same paper, we have introduced the so-called trapping factor:

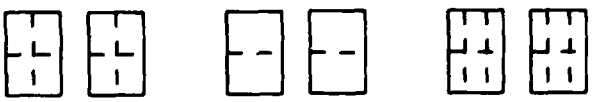
$$\Omega = \frac{\pi . L_i}{2h} \sqrt{\frac{\overline{C}_{66}}{C_{ii}}} \sqrt{\frac{F_0^2}{F_1} - 1}$$

where $L_i$ and $C_{ii}$ are respectively the length and elastic constant along the coupling axis while $F_0$ and $F_1$ are cut-off frequencies of unelectroded and electroded regions respectively. Q.A.S. modes activity is decreasing with an increasing trapping factor; then appear Q.S. modes.

In the present analysis (tri-dimensional analysis), we take into account the effect of electrode width and unwanted modes that have node lines in that direction(along to the coupling axis)fig 5. As in that direction the resonators are symmetric, Q.A.S. modes are purely A.S. modes then they have no activity (except if there is another cause of asymmetry).



purely Symmetric mode piezoelectrically active



purely Anti-symmetric modes piezoelectrically inactive

Quasi-symmetric mode piezoelectrically active

Quasi anti-symmetric mode piezoelectrically active

fig 5: Pattern of some unwanted anharmonic modes

As can be seen from this figure, Q.A.S. modes having only node lines normal to the coupling axis are piezoelectrically active. On fig 6 unwanted mode activity is plotted versus trapping factor and length to width electrode ratios. It appears that a length to width electrode ratio (equal to $\sqrt{M_n}/\sqrt{C_{55}}$ ) gives a good compromise between a lowest Q.A.S activity and the highest value of $\Omega$ for which Q.S modes appear. It results from that some optimum designs for monolithic filters.

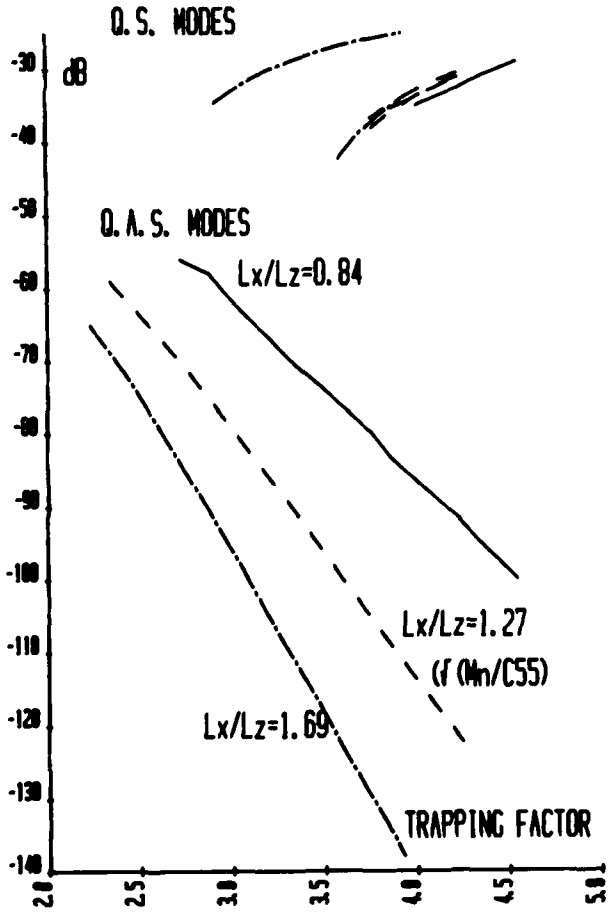



Fig 6: Activity of unwanted anharmonic modes versus trapping factor.

## *Conclusion*

Semi-analytical analysis allows the treatment of complex cases of resonators and monolithic filters. When electrode shape and tabs are rectangular, it leads to a computer program that takes low computer time (few seconds). From this study some optimum designs for a monolithic filter have been found.

## Aknowledgments

The authors thank J. Detaint to have pointed out the importance of Tiersten's equation and helped for contour and perspective views and C. Joly for its experimental work.

## References

1. H.F. TIERSTEN, "Analysis of overtone modes in monolithic crystal filters", J. Acoust. Soc. Am., Vol 62, No. 6, December 1977.

2. H.F. TIERSTEN, R.C. SMYTHE, "An analysis of contoured crystal resonators operating in overtones of coupled thickness-*shear and thickness-twist*", J. Acoust. Soc. Am., Vol. 65, No 6, June 1979

3. J. DETAINT, H. CARRU, J. SCHWARTZEL, C. JOLY, "Analysis of partially or totally contoured resonators with finite lateral dimensions", this Symposium.

4. A. GLOWINSKI, R. LANCON, R. LEFEVRE, "Effects of asymmetry in trapped energy piezoelectric resonators", Proc. of 27th Ann. Symp. on Freq. Control pp. 233-242 1973

5. R. LEFEVRE, "Monolithic crystal filters with high Q-factor and low spurious level", Proc. of 33th Ann. Symp. on Freq. Control pp 148-158 1979.

42nd Annual Frequency Control Symposium - 1988

# PERFORMANCE OF TFR FILTERS UNDER ELEVATED POWER CONDITIONS

R.S. Ketcham, G.R. Kline, and K.M. Lakin

Microelectronics Research Center
Applied Sciences Center
1925 Scholl Road
Iowa State University
Ames, Iowa 50011

## Introduction

Research into the application of the thin-film piezoelectric resonator (TFR) technology in novel monolithic bandpass filtering devices has been conducted at the Microelectronics Research Center (MRC).[1,2,3] TFR based filter devices appear uniquely suited for MIMIC receiver applications requiring ultraminiaturization, such as filter banks, and monolithic integration with active devices, such as channelized communication systems. Research into the performance of TFR based filters under elevated power levels was initiated to assess the suitability of this technology in systems requiring ultraminiature transmit and transceiver functions. Characteristics of particular interest include both the 1 dB compression point and the two-tone third order intercept point. This paper presents the results of research conducted on silicon based TFR bandpass filter devices under elevated power conditions.

## Background

### Filter Structures

TFR technology currently embodies a range of devices that utilize thin film piezoelectric materials to provide ultraminiature electromechanical resonators. Typically fabricated on silicon or GaAs substrates, single TFR structures can be used as frequency control elements in oscillators. In addition, complex multi-resonator structures can be used to synthesize any one of a number of filter networks.

The application of the TFR technology in filter devices was initially pursued on silicon with investigations into monolithic filter structures, illustrated in Figure 1a. These results were significant in demonstrating the use of thin-film piezoelectric membranes in ultraminiature filter structures. Additional thin film monolithic filter (TFM) results were recently reported by Lakin et al.[3] citing a center frequency of 1.4 GHz with an insertion loss of less than 3 dB.

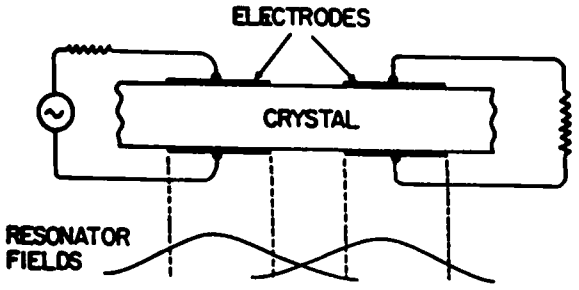



Fig. 1a) Thin film monolithic filter (TFM).

Of recent significance is the realization of the monolithic stacked crystal filter (SCF). Ballato et al.,[4] conducted initial research into quartz "hybrid" realizations of the SCF while Kline et al.[5] recently reported on the performance of monolithic SCF devices fabricated on silicon. Illustrated in Figure 1b, the SCF is a five layer structure consisting of two piezoelectric thin-films separated by a conducting ground-plane electrode. Low insertion loss filters can be realized due to direct thickness mode resonator coupling, while symmetrical bandpass characteristics are achieved due to the central ground-plane electrode. Using this structure, monolithic multi-section SCF devices, exhibiting higher selectivity, have been achieved on silicon by cascading single SCF sections, without interstate matching networks, and have demonstrated center frequencies above 1 GHz with insertion losses of .75 dB per section.

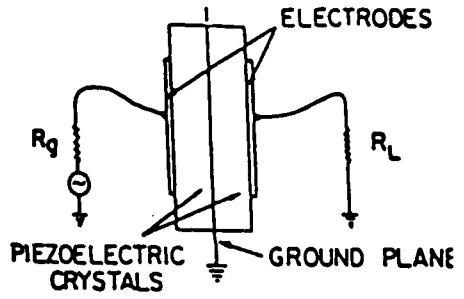



Fig. 1b) Single section stacked crystal filter (SCF).

Of primary interest for MIMIC applications are the SCF and the TFM structures because of their efficient coupling and cascading features. Classic ladder filter networks can be synthesized using TFR structures, but have not been rigorously pursued due to their inefficient utilization of die area and fabrication complexity.

### Materials

Because these structures are synthesized using thin-film deposition techniques, operation at microwave frequencies is possible. SCF and TFM networks, illustrated in Figure 1, with center frequencies of 500 MHz to 1 GHz, were utilized in this investigation. Both the SCF and TFM networks are multi-layer structures synthesized from two major types of materials: metallization layers for electrode structures and piezoelectric layers that provide for the electromechanical transduction. Optionally, these networks can incorporate portions of the supporting silicon membrane into their structures.

The SCF and TFM devices utilized in this paper are five layer structures. Filter electrodes were formed using evaporate Al with typical metallization thicknesses ranging from 1200 Å to 2400 Å. C-axis normal AlN was used as the piezoelectric material in these filter structures. Depending on the filter center frequency, AlN thicknesses ranged form 3 to 5.5 um.

AlN is a refractory material with exceptional thermal conductivity characteristics.[6] Its characteristics should enhance device performance at higher power

CH2588-2/88/0000-106 $1.00 C 1988 IEEE

levels by minimizing elastic nonlinearities and thermal gradients within the device structure itself. In addition, AlN has a thermal expansion coefficient that closely matches that of silicon. This should minimize finite strain effects induced by thermal mismatches within the composite structure.

## Intermodulation and Power Handling Characteristics

External Characteristics such as intermodulation and compression have been attributed to a number of phenomena that occur within the resonator structure.[7] Primary among these are non-linearities associated with the motional parameters, thermal effects, and non-linear coupling to unwanted vibrational modes. These effects have received extensive study in the context of specific quartz devices[8 9 10 11] and in a general context.[12]

Due to their recent development, TFR structures have not yet received extensive study, and precludes an exacting treatment of these phenomena in this paper. Primary limitations include: data concerning higher order elastic, dielectric, thermal, and piezoelectric material constants for piezoelectric thin films, and the derivation of electroelastic equations for specific composite TFR structures. Tsubouchi et al.[13] have established preliminary information concerning low order material constants of thin film AlN but provided no information concerning higher order material constants.

Elementary treatment of the thermal characteristics of TFR structures can be performed and used as point of comparison with other device technologies. Figure 2 illustrates a resonator cavity based on uniform piezoelectric material, such as quartz, with metal electrodes and the crystal holders. Valentin et al.,[14] demonstrated that dissipation of heat from low frequency quartz resonators is achieved primarily via lateral conduction through the piezoelectric material and the crystal holders. Thermal gradients exist across the crystal due to a finite thermal resistance and are also influenced by the finite number of holders in contact with the crystal. This differs from high frequency quartz units where the membrane device is symmetrically supported by the unmilled crystal blank.

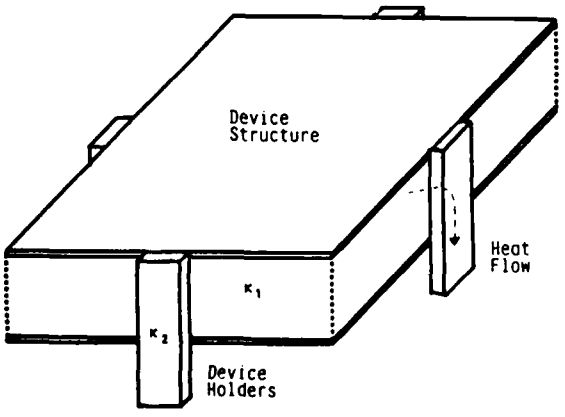



Fig. 2) Thermal environment of uniform resonator structure.

Figure 3 illustrates a simplified composite resonator structure consisting of multiple materials, each with differing values of thermal conductivity. As previously discussed, AlN has a theoretical thermal conductivity of 320 W/mK, silicon has a value of 153 W/mK, while Al has a value of 236 W/mK. For comparable structures, these materials will yield devices with approximately 50 times the thermal conductivity of quartz and should minimize the thermal gradients across

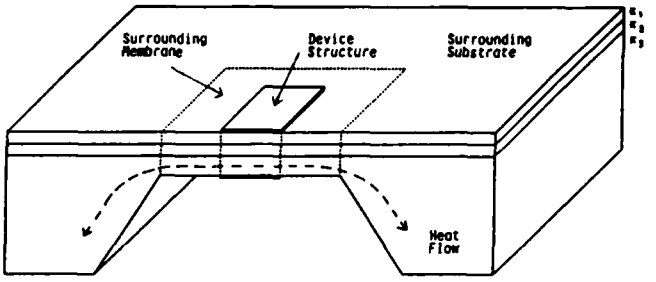



Fig. 3) Thermal environment of composite TFR structure.

the resonator structure. Thermal gradients are further reduced because the filter device is symmetrically held by the membrane and the surrounding high conductivity substrate material.

Transient thermal effects occur quite rapidly due to the high thermal conductivity and small heat capacity of the filter structures. Filters used in this study have volumes ranging from $10^{-4}$ to $10^{-3}$ $mm^3$, and have approximately $10^{-4}$ the heat capacity of lower frequency devices. Because of the reduced heat capacity and high thermal conductivity materials, TFR devices rapidly achieve thermal equilibrium.

## Results

Compression characterization was performed using the system diagramed in Figure 4. This system consists of an HP-8753 network analyzer interconnected to an external ENI-603L power amplifier and directional couplers to allow testing at elevated power levels. An HP-8569 spectrum analyzer was used to calibrate the testing power levels and to record the transmitted power levels. Due to limitations of the power amplifier, power levels of +35 dBm could be achieved only through 800 MHz. Testing above 800 MHz could be performed, but at lower power levels. This system was used to perform both normalized and unnormalized swept measurements.

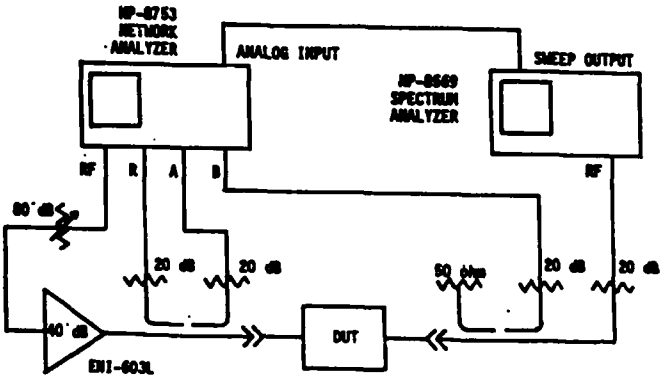



Fig. 4) Compression measuring system.

## Stacked Crystal Filters

Stacked crystal filters exhibit a series of bandpass responses that correspond to the resonant modes of the composite structure. Figure 5 illustrates the three primary modes that can be excited with mode n=2 representing the optimum bandpass responses. The samples that were evaluated consisted of cascaded single section SCF that have been monolithically fabricated onto a single membrane. These filters are designed to operate directly into 50 ohm loads and can be cascaded without interstage matching networks, illustrated in Figure 6. The volume of these filters are approximately $1.2 \times 10^{-3}$ mm$^3$.

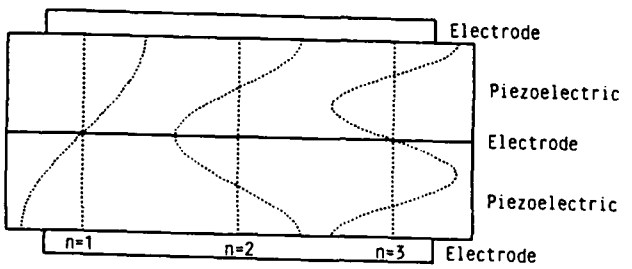



Fig. 5) Primary SCF mode structures.

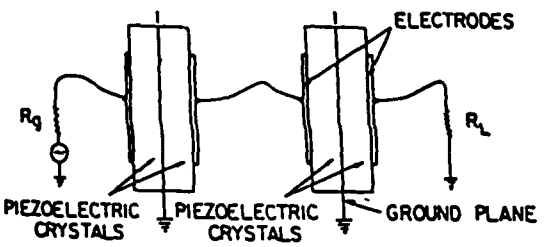



Fig. 6) Cascaded double section SCF. Note lack of interstage matching network.

In the samples that were evaluated, the optimum bandpass response exhibits a low insertion loss and typically occurs above 1 GHz. Figure 7 illustrates an n=2 response obtained from a cascaded SCF. The unwanted fine structure present on the bandpass is the result of a low level spurious mode induced by the electrode structure.

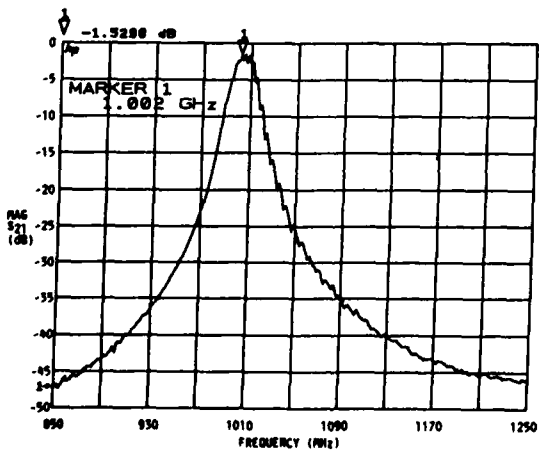



Fig. 7) Response of monolithic double section SCF (mode n=2).

Figure 8 illustrates a series of normalized responses taken with incident power levels adjusted from +5 dBm to +25 dBm in 5 dBm increments. This particular sample has a minimum insertion loss of 1.84 dB and was tested using on-wafer microwave probes. The return loss is typically -25 dB. Figure 8a is an 801 point high resolution swept response using a sweep time of 80 seconds. There is negligible passband compression at +25 dBm and only a slight downshift of the high side of the passband due to filter heating. Figure 8b is a 26 point low resolution swept response using a sweep time of 12.5 ms. This figure illustrates the ability of this structure to rapidly reach thermal equilibrium.

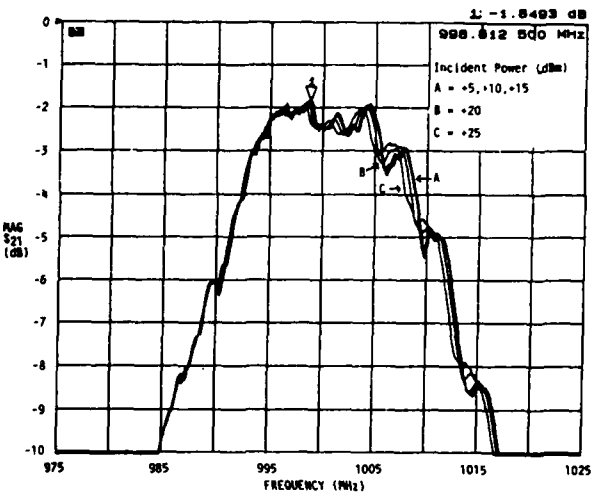



Fig. 8a) High resolution swept responses of double section SCF (mode n=2, N=801 points, $t_S$=80s).

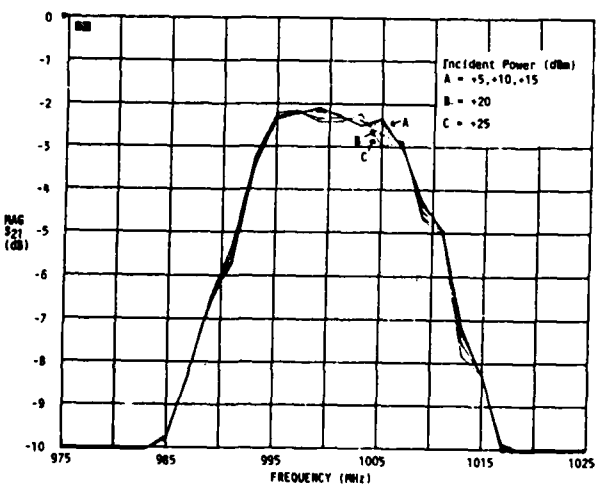



Fig. 8b) Low resolution swept responses of double section SCF (mode n=2, N=26 points, $t_S$=12.5ms).

Due to the limited power capability at 1 GHz, SCF testing at higher power levels was conducted at a lower frequency by operating the filter in the n=1 mode. Figure 9 illustrates an unnormalized series of swept measurements that were made on an SCF die mounted in a power transistor package. High resolution, 801 point, swept measurements were taken from 0 dBm to +30 dBm in 1 dBm increments using an 80 second sweep time. This figure clearly illustrates a downward shift in the passband, a result of the temperature uncompensated filter being subjected to internal heating. The apparent increase of the filter bandwidth is a measurement artifact that results from the filter thermally tracking the swept signal.

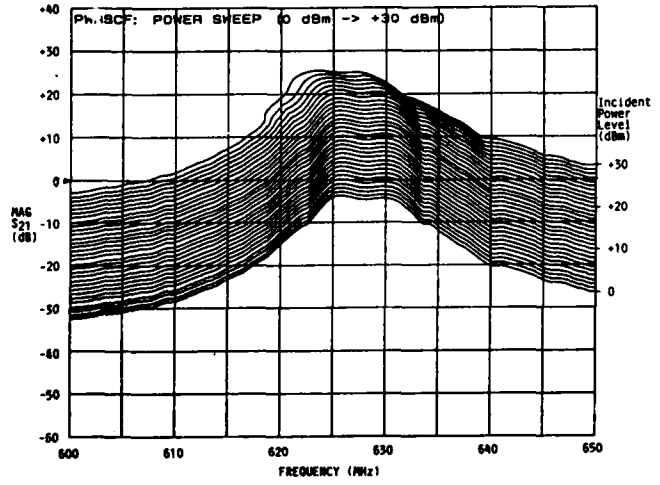



Fig. 9) Unnormalized high resolution swept responses of double section SCF (mode n=1. N=801 points. $t_S$=80s).

Monolithic Filters

Figure 10 illustrates a series of normalized swept responses that were collected from a TFM device with a center frequency of 975 MHz. Incident power levels were adjusted from +5 dBm to +20 dBm in 5 dBm increments. Figure 10a is a high resolution, 801 point, swept measurement using a sweep time of 80 seconds. This filter experiences thermally related problems at +20 dBm incident power level and is probably attributable to the fact that the volume of this filter is $1.5x10^{-4}$ $mm^3$. This is an order of magnitude smaller that the SCF units previously discussed and impacts the filter by causing power densities to increase and the heat capacity of the filter structure to be reduced.

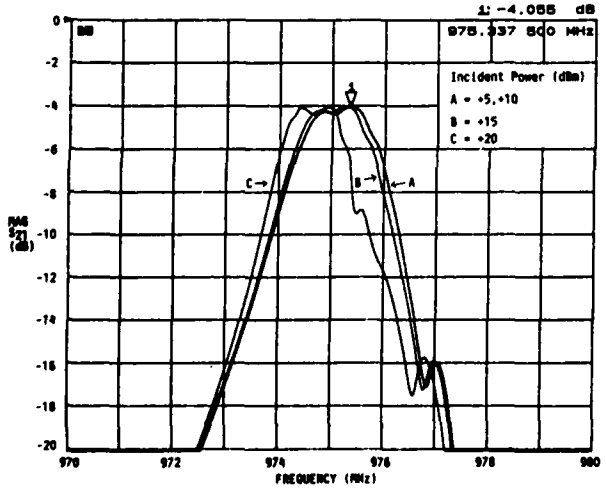



Fig. 10a) High resolution swept responses of TFM (N=801 points. $t_S$=80s).

Figure 10b is a low resolution, 26 point, swept measurement using a sweep time of 21.5 ms. This figure clearly illustrates the ability of these structures to rapidly attain thermal equilibrium.

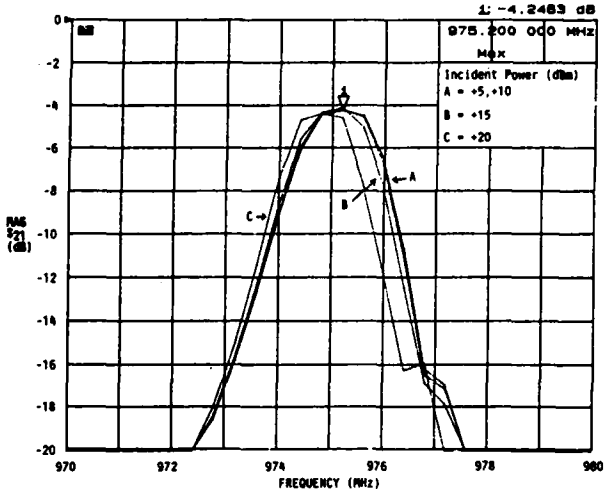



Fig. 10b) Low resolution swept responses of TFM (N=26 points. $t_S$=12.5ms).

Figure 11 illustrates a series of unnormalized swept responses that were collected from a TFM device with a center frequency of 585 MHz. Incident power levels were adjusted from 0 dBm to +35 dBm in 1 dBm increments. As was evident with the SCF sample, this figure clearly illustrates a downward shift in the passband, a result of the temperature uncompensated filter being subjected to internal heating. The apparent increase of the filter bandwidth is a measurement artifact that results from the filter thermally tracking the swept signal. This particular sample has a volume of $1.8x10^{-3}$ $mm^3$, which is 12 times larger than the previously described TFM sample.

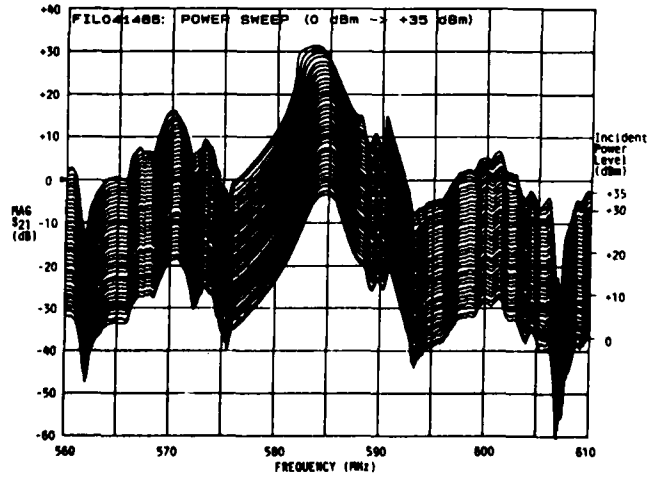



Fig. 11) Unnormalized high resolution swept responses of TFM (N=801 points. $t_S$=80s).

Figure 12 illustrates a series of normalized swept responses that were collected from this same TFM device. Incident power levels were adjusted from +10 dBm to +35 dBm in 5 dBm increments. Figure 12a is a high resolution, 801 point, swept measurement using a sweep time of 80 seconds. Figure 12b is a low resolution, 26 point, swept measurement using a sweep time of 12.5 ms. Of significance in Figure 12 are the swept responses obtained at the +35 dBm power level. At the higher sweep speed, the passband does not exhibit near the distortion as that displayed with the slower sweep time. This strongly suggests that under isothermal conditions, the filter is extremely linear with respect to its elastic and piezoelectric constants, and that the majority of the observed nonlinear related phenomena are thermally induced. This has implications for filters that are used in pulsed power applications.

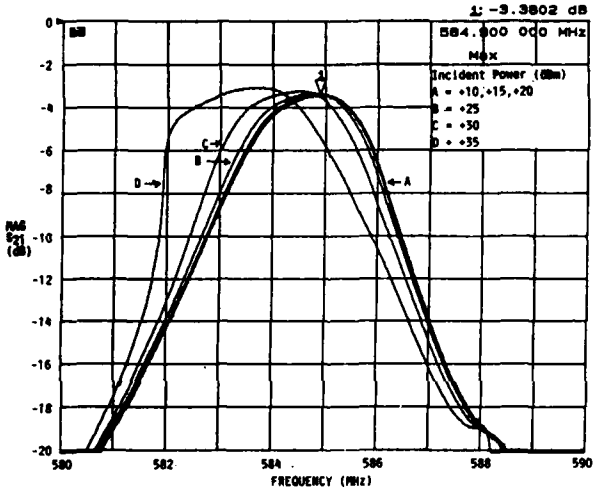



Fig. 12a) High resolution swept responses of TFM (N=801 points, $t_S$=80s).

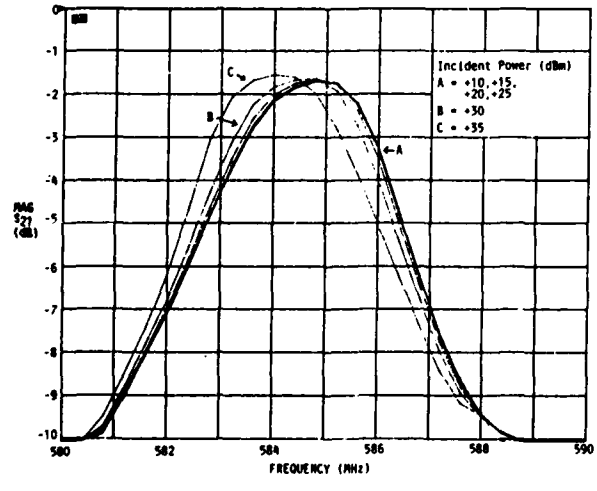



Fig. 12b) Low resolution swept responses of TFM (N=26 points, $t_S$=21.5ms).

Intermodulation testing was performed on this filter sample using the system diagramed in Figure 13. The power/frequency capabilities of this system are identical to the compression test system. The HP-8569 spectrum analyzer was used to record the intermodulation signal levels. Direct filtering of the fundamental signals to improve the dynamic range was not performed causing tests to be conducted at high incident power levels.

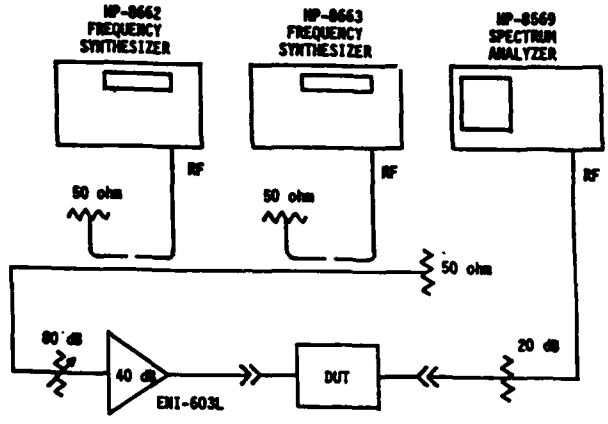



Fig. 13) Intermodulation measurement system.

For this test, two test tones were situated symmetrically in the filter passband so that the intermodulation products would occur outside of the filter passband. Figure 14 illustrates the results of this test using an incident power level of +30 dBm. The top trace was obtained without the filter and indicates a third order intercept point of +42 dBm, which represents the third order intercept point of the power amplifier. The lower trace was obtained with the filter and shows an inband attenuation corresponding to the filter's insertion loss and attenuation of the IM products which correspond with the skirt attenuation of the filter. This data suggests that the third order intercept point of the filter exceeds +42 dBm.

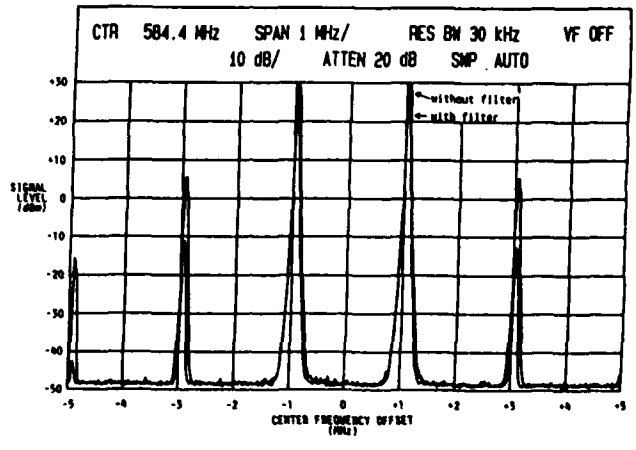



Fig. 14) Intermodulation test results for TFM pictured in Figure 11.

## Conclusion

Results obtained from this preliminary research indicate that TFR based filters are capable of operating under large power density conditions. The use of AlN in a composite structure appears to yield devices with exceptional linearity and excellent thermal characteristics.

As indicated earlier in the paper, additional work needs to be conducted in the areas of transient thermal modeling, composite structure modeling, and characterization of the constitutive materials. This information will be necessary for incorporating these devices with active devices in such applications as monolithic transmit and transceiver systems. This work will also assist in the transition of this technology onto different substrate materials such as GaAs.[15]

Additional work needs to be conducted in pulse power characterization of these filters under conditions commensurate with their bandwidth. The swept measurement methods used in this paper approximate this condition, but introduce quasi-static thermal effects which must be considered when evaluating data.

The authors would like to thank the Collins Defense Communications Division of Rockwell International and the Electrical Engineering Department of Iowa State University for making available additional RF characterization equipment.

## REFERENCES

1. K.M. Lakin and J.S. Wang, "Acoustic Bulk Wave Composite Resonators," Applied Physics Letters, Vol. 39(3), February 1, 1981, pp. 125-127.

2. G.R. Kline and K.M. Lakin, "1.0 GHz Thin Film Bulk Acoustic Wave Resonator on GaAs," Applied Physics Letters, Vol. 43(8), October 15, 1983, pp. 750-751.

3. K.M. Lakin, G.R. Kline, R.S. Ketcham, A.R. Landin, W.A. Burkland, K.T. McCarron, S.D. Braymen, and S.G. Burns, "Thin Film Resonator Technology," 41st Annual Symposium on Frequency Control, 1987.

4. A. Ballato and T. Lukaszek, "Stacked Crystal Filters," Proceedings IEEE, Vol. 51, October 1973, pp. 1495-1496.

5. G.R. Kline, R.S. Ketcham, and K.M. Lakin, "Low Insertion Loss Filters Synthesized with Thin Film Resonators," 1987 Ultrasonics Symposium Proceedings.

6. Y. Kurokawa, K. Utsumi, H. Takamizawa, T. Kamata, and S. Noguchi, "AlN Substrates with High Thermal Conductivity," IEEE Transactions on Components, Hybrids, and Manufacturing Technology, Vol. CHMT-8, No. 2, 1985.

7. M. Howard, R. Smythe, and P. Morley, "Monolithic Crystal Filters Having Improved Intermodulation and Power Handling Capability," 39th Annual Symposium on Frequency Control, 1985.

8. H. Tiersten, "Analysis of Intermodulation in Rotated Y-Cut Quartz Thickness-Shear Resonators," 28th Annual Symposium on Frequency Control, 1974.

9. R. Smythe, "Intermodulation in Thickness-Shear Resonator," 28th Annual Symposium on Frequency Control, 1974.

10. J. Gagnepain, "Nonlinear Properties of Quartz Crystal and Quartz Resonators: A Review," 35th Annual Symposium on Frequency Control, 1981.

11. R. Smythe and P. Morley, "Experimental Evaluation of the Effective Non-Linear Elastic Constant for Trapped Energy and Contoured Resonators," 39th Annual Symposium on Frequency Control, 1985.

12. J. Gagnepain, "Non-Linear Constants and Their Significance," 41st Annual Symposium on Frequency Control, 1987.

13. K. Tsubouchi, K. Sugai, and N. Mikoshiba, "Zero Temperature Coefficient Surface-Acoustic-Wave Devices Using Epitaxial AlN Films," 1982 Ultrasonics Symposium Proceedings.

14. J. Valentin, G. Theobald, and J. Gagnepain, "Temperature-Induced Frequency Shifts in Quartz Resonators," Journal of Applied Physics, Vol. 58(3), August, 1985.

15. G.R. Kline, R.S. Ketcham, and K.M. Lakin, "Thin Film Microwave Acoustic Filter on GaAs," accepted for presentation at the 1988 IEEE GaAs IC Symposium.

42nd Annual Frequency Control Symposium - 1988

# FREQUENCY SORTING BASED ON PLANAR YIG RESONATORS*

S.M. Hanna and S. Zeroug
Polytechnic University
Department of Electrical Engineering and Computer Science
Weber Research Institute
Route 110, Farmingdale, NY 11735 USA

## Abstract

A new approach for an integrated microwave channelizer is introduced. This planar channelizer is based on magnetostatic wave (MSW) resonators. It is basically a contiguous bank of filters fabricated on one substrate. The filters are realized using MSW resonators. Due to their low magnetic losses, epitaxially grown single-crystal yttrium iron garnet (YIG) films are used for the MSW propagation. The planar YIG resonators are rectangular YIG films with the microwave signals coupled to and from these resonators using microstrip structures. Experimental results describing the characteristics of the resonators are presented. An important feature of these resonators is their tunability with a linear frequency-bias field characteristic. Our theoretical analysis of the resonators gave good agreement with the experimental results. To increase the bandwidth and improve the filter response of each channel, magnetically coupled resonators have been investigated. We are using coupled MSW resonators of different dimensions to realize different channels. Our approach for frequency sorting requires no gradient in the bias magnetic field. A three-channel device based on this new approach is demonstrated. The simple structure of our channelizer is an important feature, especially for applications which normally require a large number of channels such as electronic warfare (EW) receivers.

## Introduction

Many military applications require sorting of signals according to their frequencies. Desirable features in devices used in such applications are the ability to tune them and operating directly at signal frequencies in the microwave range. As a result, the need for signal-frequency downshifting with the attendant complexity of local oscillators, filters, and multiplexers is eliminated. Although surface acoustic wave (SAW) channelizers with excellent pass-band shapes and rejection of out-of-band signals have been demonstrated[1-3], their operating frequencies are limited to about 2GHz.

Magnetostatic waves (MSW's) can operate up to and beyond 20 GHz. MSW channelizers have been recently developed by Westinghouse's researchers[4,5]. These filterbanks used narrow delay lines arranged along an input manifold. Each delay line consisted of a strip of yttrium iron garnet (YIG) film. The delay lines were identical except that there was a gradient in the bias magnetic field, so that each delay line experienced a different bias field and hence, operated at a different center frequency.

A channelizer used in an application such as electronic warfare (EW) receivers, normally is required to have a large number of channels. Thus, a simple approach for realizing the band-pass filters needed for each channel is important. In this paper we are reporting on a new approach for MSW channelizing. We are using single and coupled MSW resonator to realize different channels. These resonators are rectangular YIG films with different dimensions. The microwave signals are coupled to and from the resonators using microstrip structures. Our approach for frequency sorting requires no gradient in the applied magnetic field. The simple structure of the resonators lends itself to straightforward analysis and modeling. This approach can be readily extended to the realization of multipole multichannel filters with more rectangular passbands than are possible using the delay line approach.

## Characteristics of Single and Coupled MSW Resonators

The band-pass filters used in our channelizer are based on an MSW resonator introduced by the researchers at the Hewlett-Packard Laboratories to be used in tunable oscillator applications [6,7]. The basic configuration of this planar resonator is shown in Fig. 1.

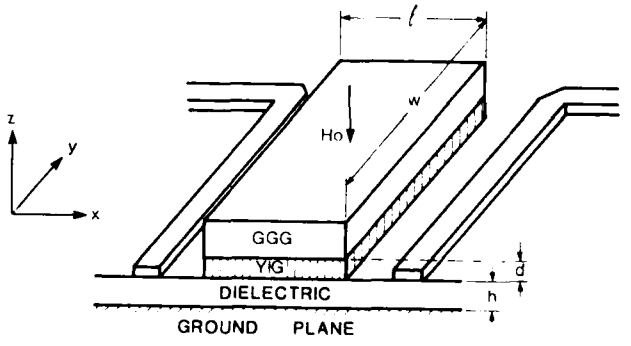



Fig. 1. Configuration of a magnetostatic wave single resonator

The resonant cavity is a rectangular sample of YIG film epitaxially grown on a nonmagnetic substrate of gadolinium gallium garnet (GGG). A magnetic bias field is applied normal to the YIG film surface, so that the magnetostatic forward volume waves (MSFVW's) are excited. These magnetostatic waves get reflected by the straight edges of the YIG film. Two microstrip lines are used to couple energy in and out of the resonator. We have designed, fabricated, and tested single YIG resonators with relatively low insertion loss (less than 6 dB). Shown in Fig. 2 are the results of the transmission and input impedance measurements of a single MSW resonator

*Supported in part by the NSF (Grant No. ECS-8722155) and New York State Center for Advanced Technology in Telecommunications.

CH2588-2/88/0000-112 $1.00 C 1988 IEEE

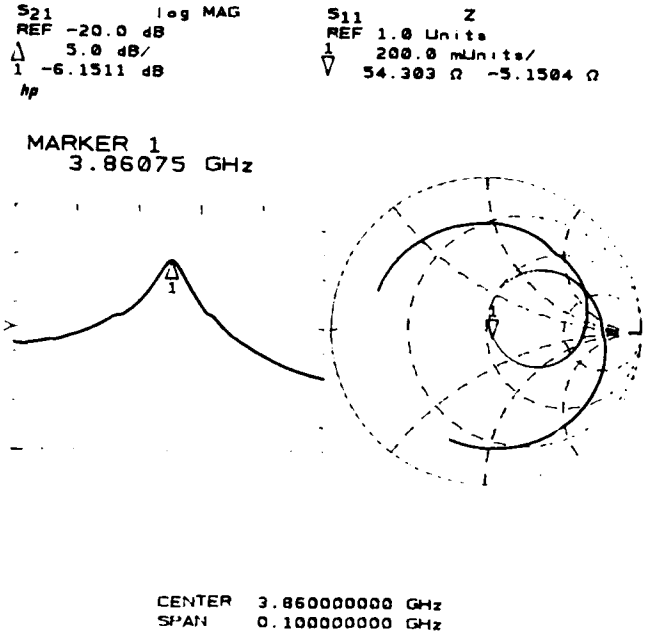



Fig. 2 a) Transmission response, and
b) Input impedance of a single MSW resonator.

In common with other waveguiding structures, YIG films can support higher order MSW modes which produce spurious resonances and hence out-of-band signals when the resonator is implemented as a filter. Since MSFVW's have nearly isotropic propagation characteristics in the YIG film plane, the frequencies of the resonant modes depend on both the length ($l$) and the width (w) of the rectangular resonator. Spurious responses from higher order width modes can seriously degrade the out-of-band rejection of the device. We have carefully analyzed the spurious modes in single MSW resonators[8]. The results of our theoretical calculations of mode separations are in good agreement with our experimental measurements.

An important feature in these MSW resonators is their tunability. By scanning the bias magnetic field, the resonator center frequency can be easily swept. In Fig. 3(a,b,c) we show the frequency response of a single resonator tuned to different frequencies by varying the applied magnetic field. The YIG film resonator (sample #2) has the dimensions $l$=2.53 mm, w=9.83 mm, and d=83$\mu$m. The microstrip lines used were fabricated on a Duroid substrate ($\epsilon_r$=2.2). The separation between the two lines is 4.5 mm.

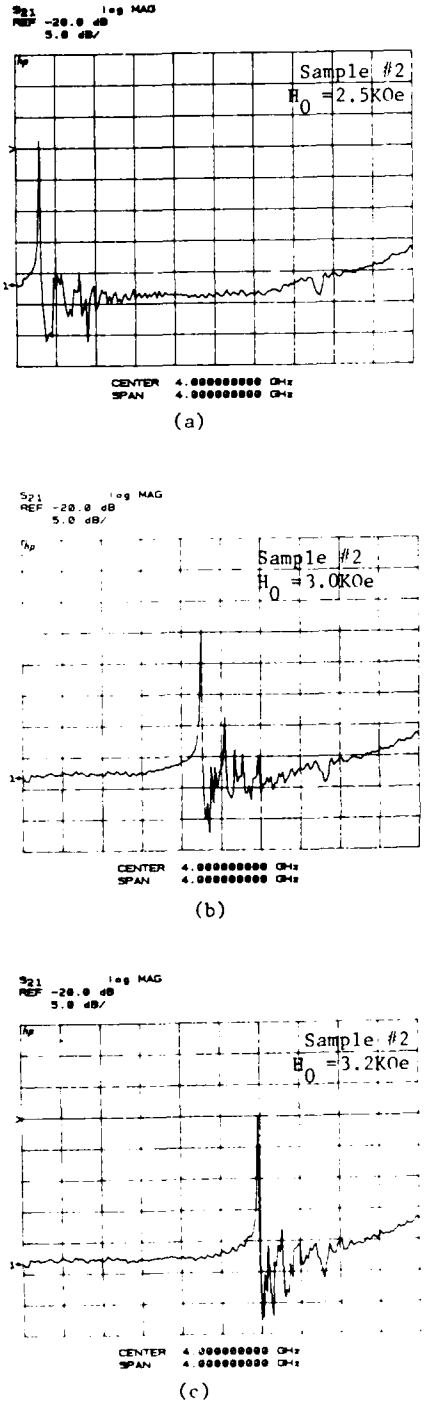



Fig. 3. Frequency response of a single MSW resonator (Sample #2) for different values of the applied magnetic field.

The tunability is clearly demonstrated by the above figures 3(a,b,c). It is worth noting that the insertion loss at resonance does not change appreciably (less than 2 dB) over the tuning range. For the sample #2 used for the experimental data shown in Fig. 3(a,b,c), we have calculated the resonance frequency as a function of the applied magnetic field. This can be done using the dispersion relation describing MSFVW's propagating in a YIG sample in the configuration shown in Fig. 1. This dispersion relation can be written as:

$$2\alpha \ \mathrm{Cot}(\alpha kd)=(\alpha^2+1)e^{-2kh} + (\alpha^2-1) \qquad (1)$$

where k is the constant of the propagation along the x-direction, and

$$\alpha = \sqrt{\frac{\omega^2-\omega_s^2}{\omega_0^2-\omega^2}}\ , \ \omega = 2\pi f,$$

$$\omega_s = \sqrt{\omega_0(\omega_0+\omega_m)}\ , \ \omega_0 =\gamma H_i\ , \ H_i=H_0-4\pi M_s$$

$\omega_m=\gamma(4\pi M_s)$ , $\gamma$=2.8MHz/Oe, and $4\pi M_s$ = 1750G.

In Fig. 4 we are plotting the results of both our experimental measurements and our theoretical calculations using the dispersion relation given by Eq. (1). The graph shows the resonance frequency as a function of the magnetic field for sample #2.

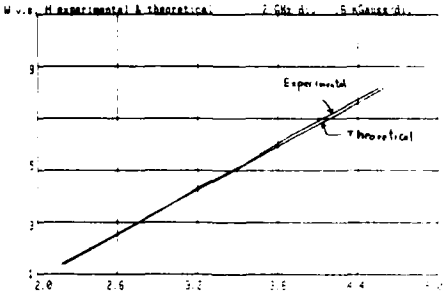



Fig. 4. Tuning characteristics of an MSW resonator.

As it is well known the use of coupled-resonators can improve the frequency response of a filter as compared with a single-resonator filter. Hanna and Zeroug have introduced a simple configuration for coupled MSW resonators.[8]. We have implemented coupled-MSW resonators in realizing the individual filters constituting our channelizer. Shown in Fig. 5 is the configuration we used for two-coupled resonators. The gap between the two YIG resonators was simply introduced using a wafer saw as a part of the process of dicing the garnet sample into the specified rectangular shape. The two rectangular resonators are magnetically coupled through the rf magnetic field linkage across the gap shown in Fig. 5. In liquid phase epitaxial growth of YIG films, two layers are normally grown on both sides of the GGG substrate. One of these YIG films was routinely lapped off. This was done to eliminate the coupling and interfering resonances introduced by a second magnetic film on the other side of the GGG substrate.

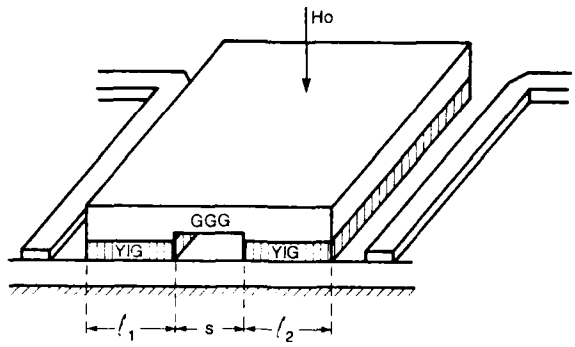



Fig. 5. Configuration for two-coupled MSW resonators.

Shown in Fig. 6 is the transmission response of a two-resonator filter. As expected, we were able to increase the 3 dB bandwidth of the filter and improve its out-of-band rejection by the use of coupled resonators. We expect further improvement in the out-of-band rejection by implementing more than two coupled resonators.

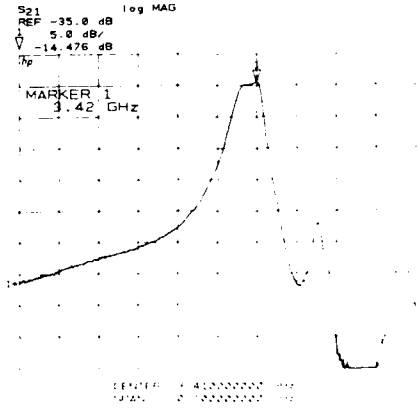



Fig. 6. Transmission response of a two-coupled MSW resonator filter.

## MSW Resonator Based Channelizer

We have implemented the MSW coupled-resonators[8] in fabricating a planar channelizer. The schematic for a three-channel filterbank is shown in Fig. 7.

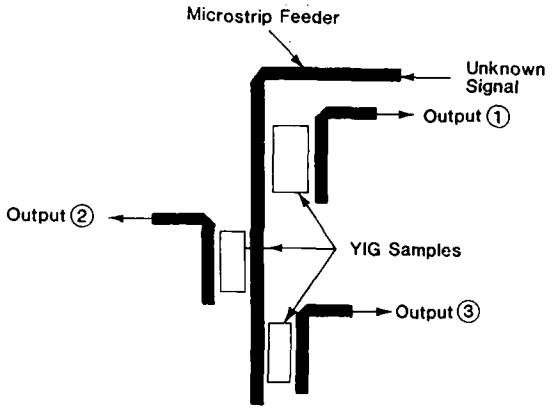



Fig. 7. Schematic for a three-channel resonator-based channelizer.

In this channelizer the filters are coupled MSW resonators of different dimensions. Thus, each filter is tuned to a slightly different center frequency. Our channelizer requires only a uniform bias magnetic field, which is easier to provide compared with the spatially varying magnetic field needed in the case of MSW filterbanks based on delay lines. We are implementing the magnetostatic forward volume waves. Thus, the dc field is applied normal to the YIG film, which is a convenient arrangement reducing the gap between the poles of the magnet supplying the bias field.

In Fig.8, we show the frequency response for the 3-channel filterbank. The channel center frequency insertion loss is 18±2dB. Obviously the development of this channelizer is far from complete. Examples of the improvements needed are: 1) the passbands of the channels need to be spaced more uniformly in frequency. 2) the flatness and the uniformity in shape from one channel passband to the other, and 3) feed through from different channels have to be minimized.

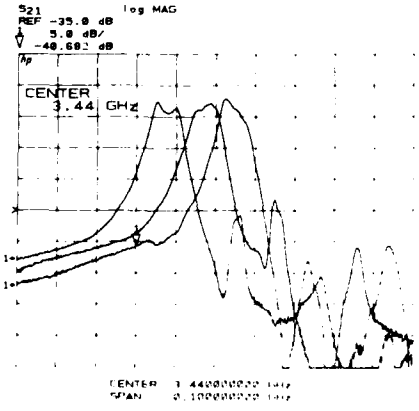



Fig. 8. Frequency response of a 3-channel filterbank.

It should be pointed out that we have not yet investigated the temperature stability of our MSW channelizer. However, a promising technique for temperature compensation of similar MSW devices has been recently proposed.[9] In this technique a bias magnet with a temperature coefficient for the coercive force of opposite sign to that of the YIG can be used. Good temperature stability can be achieved with a combination of Neodymium-Iron-Boron magnets and a part of the magnet yoke structure made from a temperature dependent permeability material.

## Summary and Conclusions

We have demonstrated a new approach for MSW filterbanks. The use of coupled resonators offers a simple means for realizing multipole filters for channelizers. The filter passband shapes using this approach should, in principle, be superior to those based on delay lines.

In our channelizer only a uniform bias field is needed. This makes the magnet design much simpler compared with the delay line-based MSW filterbank in which the frequency alignment of individual channels depends on the bias field spatial profile. The simplicity of the structure we introduced lends itself for good potential in EW applications requiring a large number of channels.

## Acknowledgement

The equipment grant (HP 8510 Network Analyzer) awarded by the Hewlett-Packard Corporation, made the measurements reported in this paper possible. We also thank W.Ishak of HP laboratories for the valuable discussions and for supplying some of the garnet samples used in this investigation.

## References

1. D.E. Allen et al., "Surface Acoustic Wave Components for Frequency Sorting Receivers", IEEE 1979 Ultrasonics Symposium Proceedings, pp. 555-557 (1979).

2. J.B. Lagrange et al., "SAW Devices: The Answer for Channelized Receivers?", Microwave Systems News 7, pp.87-96 (1977).

3. D.C. Webb, "SAW Filters Simplify Signal Sorting", Microwave Systems News, 8, pp. 75-84 (1978).

4. J.D. Adam, M.R. Daniel and S.H. Talisa, "MSW Filterbanks", Microwave Journal, 31, pp. 107-124 (1987).

5. J.D. Adam, "Magnetostatic Wave Multichannel Filters", Proc. of the 36th annual Frequency Control Symposium, pp. 419-427 (1982).

6. K.W. Chang and W.S. Ishak, "Magnetostatic Forward Volume Straight-Edge Resonators", Proc. of the 1986 IEEE-MTT Symp., pp. 473-475 (1986).

7. W. Ishak and K.W. Chang, "Tunable Microwave Resonators Using Magnetostatic Waves in YIG Films", IEEE Trans. on Microwave Theory and Tech., MTT-34, pp. 1383-1393 (1986).

8. S.M. Hanna and S. Zeroug, "Single and Coupled MSW Resonators for Microwave Channelizers", to be presented at the 4th Joint MMM-Intermag Conference, Vancouver, British Columbia, Canada, July 1988.

9. J.D. Adam, M.R. Daniel and S.H. Talisa", A 13-Channel Magnetostatic Wave Filterbank", 1988 IEEE MTT-S Symposium Digest, pp. 879-882 (1988).

# HIGH PERFORMANCE QUARTZ

R. A. [illegible]DISE and R. L. BARNS, AT&T Bell Laboratories, Murray Hill, NJ 07974.
J. J. STEVENS, AT&T Bell Laboratories North Andover, Massachusetts 01845.
E. E. SIMPSON, AT&T Network Systems, North Andover, Massachusetts 01845.
H. BROWN, AT&T Bell Laboratories, Crawfords Corner Road, Holmdel, New Jersey 07733.

Recent experiments in the laboratory and under factory conditions and a re-assessment of the literature aimed at developing techniques for preparing high performance quartz especially r-bar quartz for SAWS, radiation hard quartz and low channel and dislocation density material are reported. We find r-bar material to be preferrable for SAW wafers because the wafer is cut parallel to the growing face resulting in greater homogeneity and increased yield. For the wafer sizes required 5"X growth (growth on a face rotated 5° from (0001)) requires 6 months while wafers large enough for devices can be grown in 6 weeks on r-bar seeds (growth on $(10\bar{1}1)$ face). Possible procedures for Al impurity reduction so as to improve radiation resistance are described. Recrystallization of natural quartz nutrient and crushing the grown crystals for nutrient does not result in significant purification unless several recrystallizations are employed. A single recrystallization is shown to be equivalent to single pass zone refining and would not be expected to result in significant purification. Procedures for using a very low Al content $SiO_2$-glass as nutrient are described. Etch Channels are shown to be decorated dislocations and measures which reduce dislocations: using dislocation free (DLF) seeds, reducing particulate inclusions and minimizing seed-grown material lattice mismatch are effective. Additional means of channel reduction may depend upon avoiding decoration of dislocations by Na.

CH2588-2/88/0000-116 $1.00 C 1988 IEEE

## 1.0 Introduction

Since the discovery of efficient techniques for synthesizing quartz in the 1940's and 50's and the transfer of these techniques to the factory in the 1950's and 60's, there has been a continual need for improved quality material. In response to this need a continuing series of investigations has produced almost continuously improved quartz. In a recent paper[1] we have reviewed the perfection of cultured quartz and its connection to growth processes. In that paper we describe the physical chemical understanding needed and the growth process changes required to eliminate dendritic growth and obtain commercially useful rates, reduce OH content, obtain high acoustic Q at commercial rates and grow dislocation free material in the laboratory. In addition, we outlined the current status of understanding regarding the genesis and reduction of etch channels and the sources and amelioration of radiation damage in quartz. The reader is referred to that paper and to the review by Brice[2] for background information, definitions and a summary of the status of these subjects up to late 1986. In this present paper, we wish to report our recent studies aimed at producing high performance quartz, especially r-bar grown material (growth on $(10\bar{1}1)$ face seeds) for surface acoustic wave devices, radiation hardened quartz, and low channel density material. Finally, we have re-examined the literature concerning dislocation-free quartz and wish to emphasize an important, often overlooked mechanism for defect generation, i.e., seed-grown material lattice mismatch.

## 2.0 Experimental

The quartz examined was all hydrothermally grown in steel autoclaves using Brazilian lascas nutrient, milky Arkansas nutrient, nutrient cut from previously hydrothermally grown material or high purity $SiO_2$ glass. (Dynasil Corp., Berlin, NJ). The vessel size used at Bell Laboratories was 1 inch inside diameter by 10 inches inside length. In some experiments, production sized autoclaves (10'' ID × 10' IL) were used at AT&T Network Systems Merrimack Valley quartz facility. Typical growth conditions are given in Table 1. Growth rate depended strongly on the nature of the nutrient and was higher and not very reproducible when glass was used.

Crystals for SAW (surface acoustic wave) applications were evaluated by processing them through our Merimick Valley device-line including orientation, cutting, metallization, etc. and observing their SAW behavior.

## 3.0 Results and Discussion

### 3.1 Quartz for SAW Devices

Surface acoustic wave devices are finding increasing application because they are usable at higher frequencies than bulk devices. The importance of SAW wafer homogeniety has been emphasized by Bulst and Willibald.[3] They observed that the reproducibility of the resonant frequency of SAW devices was poor when using material which showed banding in X-ray topographs. Wafers cut from cultured material were more uniform than those from natural quartz and wafers cut from Z-sectors were more uniform than wafers from X-sector material. Bye and Cosier[4] reported that the presence of growth striations caused property deterioration such as increases in equivalent series resistance even in BAW (bulk acoustic wave) resonators.

We have compared the utility of 5°X (growth on seeds rotated 5° about the X axis from the (0001) surface) and r-bar grown quartz for SAWs and find r-bar more effective. The usual AT orientation used for SAWs leads to plates nearly parallel to the growing face for r-bar material but highly inclined for 5°X material. As is well known, impurity concentration striations and gradients are always much more severe perpendicular to the growth face so that, in the plane of a SAW wafer, r-bar material will be more homogeneous. Further evidence of the role of wafer homogeniety in device behavior and the advantage of r-bar growth can be gleaned from our experience with large-scale production of A-6 monolithic crystal filter units (where the device wafer is nearly parallel to the r-face and where very stringent device requirements had to be met). r-bar material gave increased yields compared to 5°X grown quartz because of the greater uniformity of material within a wafer. The combined evidence from these studies convinced us that for high performance SAW devices, homogeneity was an important advantage and that r-bar growth was to be preferred.

For the wafer sizes required and bearing in mind the requirement that one not cut devices through the seed, to get an appropriate size wafer from 5°X growth requires 6 months while r-bar quartz large enough for wafers can be growth in 6 weeks. In addition, because of the edge effects and other wastage, the yield for r-bar is ~10% greater than 5°X. Thus, factory costs are lower and quality, especially homogeneity, is better for r-bar. Finally, Table 2 compares insertion loss for r-bar and 5°X material. As can be seen, there is no significant difference.

### 3.2 Radiation Hardening

The effects of ionizing radiation (x-rays and gamma rays) are important for quartz devices for use in space and defense applications. Radiation changes the elastic constants and hence the resonant frequency and other parameters of a device. Transient effects during irradiation are very complicated and are functions, for example, of the total dose, the dose rate, the crystal cut, the material of the device enclosure, the ambient gas or vacuum, the electrode material, and the impurity content in the quartz and gradients in impurity content.[5–9] Permanent frequency changes after irradiation also depend on many of these factors.[5–9] There is fairly general agreement that Al is the most important impurity with respect to radiation effects. This has been discussed in many reports, for example ref.[9–11]

Lascas from Brazil is the most common nutrient for quartz growth. Iwasaki et al.[12] and Suzuki et al.[13] show that the Al content from various mines in Brazil varies from 15 to 214 ppm. Milky Arkansas nutrient[14] has about 30 ppm Al as shown in Table. 3. However, the intrinsic inhomogeniety of natural specimens will often result in high Al in some parts of a natural quartz nutrient batch.

It is clear that quartz grown with reduced Al should have improved radiation hardness. We tested two types of nutrient for Al reduction:

a. cultured quartz
b. high purity glass.

First, we used cultured quartz as nutrient employing large autoclaves in our Network Systems growth facility. Analysis was by atomic absorption spectroscopy. Table 3 summarizes the results. The nomenclature of Table 3 is explained as follows: Milky Arkansas quartz (32 ppm Al) was used as nutrient for growth on an r-bar seed and produced crystal r-bar-1 (with 65 ppm Al. The increase in Al content over the nutrient probably was caused by an Al rich region in the nutrient. Al rejection in X growth and preferential segregation in r-growth is also partly responsible. The r-bar material was reused as

nutrient. Recrystallization produced crystal r-bar-2 where the Al content is 58 ppm. The results of subsequent recrystallizations using r-bar nutrient are shown in the top part of the table. In another series of experiments, Arkansas nutrient was used for growth on a 5°X seed to produce 5°X-1. This crystal was crushed and reused as nutrient for growth on r-bar seeds to produce r-bar-a. 5°X-1 nutrient was also used to produce 5°X–2 which was used as nubrent for r-bar-b etc. These experiments are shown in the bottom part of the table. Clearly, recrystallization using 5°X grown nutrient is more effective, but as Table 3 shows purification by recrystallization using conventional hydrothermal procedures is difficult.

The reasons for the inefficiency of such recrystallizations are relatively straightforward. Consider a typical recrystallization sequence. Quartz dissolves in the hydrothermal solution phase and crystallizes from the solution continuously throughout the process. The distribution of impurities is analogous to zone refining. Figure 1[15] is a schematic of impurity concentration vs. fraction crystallized for zone refining or hydrothermal crystallization for a case where the distribution constant for the impurity, $k_{eff}$ (the effective distribution constant for growth at a particular rate on a given crystallographic face), is less than 1 (which is true for Al in quartz) and $C_o$ is the initial concentration of impurity in the ingot to be zone refined or in the hydrothermal nutrient to be recrystallized.

Considering the quartz-Al purification case, as can be seen, the first quartz to recrystallize has an Al concentration

$$C_s = k_{eff}\, C_l \qquad (1)$$

where $C_s$ is the Al concentration in the initial growth and $C_l$ is the Al concentration in the hydrothermal solution. The concentration in the liquid is given by

$$C_l = C_o s \qquad (2)$$

where $C_o$ is the concentration of Al in the nutrient and s is the solubility of quartz. For the first material to recrystallize,

$$C_s = k_{eff} C_o s \qquad (3)$$

In general,

$$k_{eff} \neq k_o\ , \qquad (4)$$

the equilibrium distribution constant at very low growth rates so that the impurity profile of Fig. 1 is, in practice, a family of curves for various $k_{eff}$'s. Consider a case where essentially all growth is on a single crystallographic face. Growth using seeds of several crystallographic faces and growth on seeds which produce a habit with significant growth in several directions are not such cases. The deviations brought about by these effects will be covered below. $C_{Al}$ in solution increases over the interval a-b of Fig. 1 until

$$C_l = C_o s / k_{eff} \qquad (5)$$

at b so that

$$C_s = C_o\ . \qquad (6)$$

Since at b, Al dissolves from the nutrient in an amount equal to its incorporation, during the better part of the crystallization, i.e. the interval b-c, crystals grow with Al equal to $C_o$. When growth ends at d, Al is left behind in solution. The process is like zone refining with a single pass with the slight advantage that some Al is left behind in the solution. However, not much purification results unless one chooses to crop off material and use only the first part of the crystal to grow (in interval a-b).

The foregoing is for the case where all growth is in one crystallographic direction. In the present case with 5°X and r-bar seeds, some of the growth will be in the +X and −X direction, but the qualitative features described above still obtain. At the time growth ceases the $SiO_2$ left behind in solution is sV where s is the solubility of silica (under the usual growth conditions, this is 2-5%) and V is the volume of solution. Given the usual solution volumes, $SiO_2$ solubility and nutrient weight, the result is that the fraction of quartz crystallized is always >0.9, and the amount of Al left behind in solution is not very large. Thus, purification is not expected to be large. Table 3 shows that this is the case.

Preliminary experiments indicated that $k_{eff}$ for 5°X is smaller than $k_{eff}$ for r-bar so that as indicated in Table 3 purification is more efficient with 5°X growth. If growth surfaces of several different orientations are present in the same run, as is the case in real experiments, the surface with the larger $k_{eff}$ will "getter" Al from the solution relative to the surface with the smaller $k_{eff}$ and end up with more Al than the nutrient. This was probably a signficant contribution to the Al enrichment seen in the case of the first crystallization on r-bar seeds using Arkansas nutrient as shown in Table 3 (r-bar-1). As Table 3 shows, r-bar growth from r-bar nutrient is not as effective a purification strategy as r-bar growth from 5°X nutrient. It should be pointed out that, in the experiments of Table 3, the +X and −X growth regions were not cropped from the grown crystals. Further purification would, of course, result if that were done. However, clearly for the lowest Al content and highest radiation hardness, 5°X nutrient is to be preferred. If one wants the advantages for SAW's, recrystallization on r-bar seeds using recrystallized 5°X nutrient is appropriate. However even under the best conditions, recrystallization is not a very effective purification strategy as many steps are required to produce large reductions in Al. Previous work along these lines was reported in refs. 16-22 and the results are in good agreement.

From the foregoing, we conclude that a better strategy is to start with the highest purity nutrient available. Silica glass from Dynasil (and probably other suppliers) is one of the purest forms of silica and contains <0.02 ppm of Al. Clearly, if it could be used as a nutrient, low Al quartz should result. However, techniques for such use need development.

It is well known that amorphous nutrients almost always produce poor quality growth. This is because amorphous material has a higher solubility than crystalline material, leading to very high initial supersaturation and poorly controlled growth. Under hydrothermal conditions, the surface of amorphous material tends to convert *in situ* to thermodynamically stable crystalline quartz. Thus, the solubility will fluctuate as various regions of the amorphous nutrient surface are converted to crystalline quartz, are then dissolved away exposing more amorphous material which is reconverted to quartz, etc. The resulting concentration of $SiO_2$ in solution and hence the supersaturation will fluctuate violently with time, causing spontaneous nucleation on the walls and poor quality growth on the seeds. It might be expected that if the difference between the dissolving and growth temperatures, $\Delta T$, were made small during the early stages of growth (until the nutrient converted to quartz) the supersaturation, $\Delta s$, might be small enough to produce good growth. Later in the growth cycle, $\Delta T$ could be increased so as to grow at an economic rate. To test these ideas we carried out glass nutrient experiments in laboratory size autoclaves at Murray Hill. We tried a programmed $\Delta T$

procedure with some improvement in quality but great variation in results. Lack of reproducibility was no doubt due to the essentially random fluctuation in $SiO_2$ concentrations which were different from run to run. Furthermore, if $\Delta T$ was made too small the seeds were lost. An obvious procedure is to convert glass to quartz in a separate process and then use the converted quartz for a "normal" run. We found that even with $\Delta T = 0$, the product of such procedures was a flocculent powdery quartz too finely divided to be useful plus a cemented-together mixture of glass and crystals which was difficult to remove from the autoclave.

A. A. Armington told us of his success in "seeding" the glass-crystal conversion by adding a small fraction of crystalline quartz to the glass nutrient. We believe "seeding" results in the ready availability of crystalline quartz nuclei in the growth zone and promotes growth on these centers locally during high $\Delta S$ "pulses" when amorphous $SiO_2$ is dissolving. The result is that these high $\Delta S$ pulses do not propagate throughout the autoclave to produce wall nucleation or poor growth on the seeds. Using seeding we carried out the series of experiments summarized in Table 4 and illustrated in Fig. 2. The growth conditions were as described in Table 1 except that when glass was present the $\Delta T$ was 5-10° for the first day and then increased to 30-45°. Where possible, seeds were somewhat thicker, 0.1" vs 0.05", to assure that they did not dissolve away during the low $\Delta T$ warm up. In addition, as indicated, various amounts of quartz were mixed with the silica glass including at least 2 gm placed in the growth region to help saturation during warm up and protect the seeds from dissolving.

Figure 2 shows the quality obtained. As can be seen (Run Q-2) quite good quality at a very good rate occurs with ~30% quartz in the nutrient. Obviously, crystalline quartz can add Al to the grown material, so that it would be expected that the samples of Table 4 would contain more Al as the crystalline quartz content of the nutrient is increased. If further reduction should be required to produce the necessary radiation hardness, amorphous $SiO_2$ can be used as part of the nutrient for subsequent recrystallizations.

Using low Al nutrient may also have a beneficial effect by reducing particulate inclusions (see below).

### 3.3 Etch Channels and Dislocations

Etch channels (sometimes called etch pipes or etch tunnels) in quartz were discovered[23] and characterized[24] long ago and are reviewed in ref. (1). With the increasing application of lithographic techniques for fabricating quartz devices, the reduction or elimination of etch channels has become a matter for intensive study.[4,25–34]

It is clear that all dislocations which intersect a surface will produce an etch pit but only a few of these dislocations will form an etch channel. The cause of this difference is not clear, but a recent exploratory experiment that we have performed suggests that decoration of a dislocation with Na is required. A laser microprobe emission spectrograph was used to analyze a quartz surface having a known density of latent etch channels. A latent etch channel is a microscopic region in a pristine sample which, if properly etched, would produce an etch channel. The presence of latent etch channels is inferred, for instance, by etching adjacent slices from a crystal and assuming that latent channels are present in an unetched slice when its adjacent neighbors have etch channels which have geometries indicating that they would thread through the pristine sample. Using very high speed film with the spectrograph, Na was detected in a fraction of the spots excited by the laser. This fraction correlated with the expected fraction based on the etch channel density (ECD) and the fraction of the surface sampled by the laser. Al and Fe were not detected at these locations but the sensitivity for these elements was much lower than for Na.

It is known that sweeping removes Na from quartz and that sweeping reduces the ECD. The use of Au electrodes for sweeping is more effective in reducing the density than Pt, and the treatment of the Au-swept material with aqua-regia restores at least some of the etch channels.[33,34] We interpret this as evidence that Au plugs the latent channels and that the partial effectiveness of Pt is somehow connected to its not being as effective a "plug". Sweeping will sufficiently plug latent channels so that channels are not produced during photolithographic processing and etching.

Another strategy for reducing channels is suggested by the work of Yamashita, Shinomiya and Kumasuki.[35] They found that by replacing part of the NaOH mineralizer with KOH, the *total* Na + K in the crystal was about 0.4 ppm compared to 2 ppm Na when pure NaOH was used. If alkali decoration is necessary to form a latent etch channel, Na + K mineralizer might be effective in reducing latent ECD even when dislocations are not controlled.

Thus, although there are ways of reducing the ECD below the latent ECD, methods of minimizing latent etch channels are desirable. Minimizing the dislocation density would surely minimize the latent ECD. From a survey of the literature and from our own experiments we concluded that there are only three sources of dislocations in quartz: (1) dislocations in the seed, which propagate threading dislocations in new growth, (2) particulate inclusions (PI) grown into the quartz which may generate new dislocations and (3) lattice mismatch between the seed and the growth which introduces misfit dislocations.

To reduce dislocations due to (1) carefully selected low dislocation or dislocation free (DLF) natural quartz can be used as seed material. For commercial growth this is expensive (because of the necessity of not only finding DLF regions in natural quartz but also because of the need of avoiding twinning which is very common in natural crystals) but has been used to produce dislocation free growth.[36,37] Since the only dislocations which can propagate from the seed during growth are those which intersect the growth surface, Zarka, Liu and Buisson[38] were able to show that cutting basal seeds from the +X region is an effective method of simulating DLF seeds since the dislocations in the +X region lie nearly parallel to the basal plane. This method was subsequently confirmed by others[16,18,21,39–40] and could provide a route to effective DLF seeds without resort to the tedious selection of natural quartz.

Using swept seed plates to reduce dislocation density in growth on seeds has been tried in several studies[5,21,41–42] and by us without success. This is understandable since there is no reason to expect that sweeping would heal dislocations. However, in disagreement with others we note that Dowsett, et al.[43] claim to have produced low dislocation density growth by using swept seeds. Details were not given.

A method of producing r-bar seeds with low dislocation densities would be especially useful. We would like to suggest a possible procedure. The stereographic projection in Fig. 3 shows how r-bar seed plates with favorably oriented dislocations can be obtained from an r-bar stone. The center pole of the projection represents a minor-rhomb face (near an AT-cut) which is the large surface of the r-bar stone. The other minor-rhomb planes are marked $z_1$ and $z_2$ and lie at an angle of 85.76° from the seed surface. This small deviation from the ideal 90° should still allow eliminating the propagation of a large fraction of the dislocations present. New seeds would be cut parallel to $z_1$. To our knowledge, this strategy has not been employed.

It might be pointed out that both Ashahara and Taki[44] and Brice[2,45] found a good correlation between dislocation density

and IR Q. Thus, strategies which improve Q (such as slow growth) may also help in reducing latent etch channels.

To reduce dislocations due to particulate inclusions (2) the following are important considerations:

We have studied dozens of X-ray topographs and photomicrographs in the literature and are led to the following conclusions: PI's rarely occur in growth except very near or on the seed surface. This suggests that if particles are continuously liberated from the nutrient or from the vessel walls, they are rarely grown into the quartz. There is some evidence that a particle must exceed a certain size to cause new dislocations.[44] If a particle generates new dislocations, it generates many rather than just one.

It is likely that many more particles are liberated from the vessel walls during warm-up when the solution is undersaturated and the mineralizer is a relatively powerful etchant. In particular, the disturbances (drying, reaction with $CO_2$ in air, scraping, etc.) of removing the previous run and loading the new run are likely to create conditions in the wall deposit which facilitate the liberation of particles during warm up.

This suggests that saturation of the solution with $SiO_2$ before the run is started, e.g. using sodium silicate solution instead of NaOH solution, might be beneficial. We tried a few runs of this sort and found that growth was rapid and of poor quality. Clearly an exact $SiO_2$ balance must be preserved. Excess $SiO_2$ provides an initial supersaturation pulse and poor quality. Too little $SiO_2$ might encourage micro-flaking of the wall. It is important to bear in mind that under production conditions where multigeneration seeds must be used, only a small fraction of the dislocations are newly generated from nutrients i.e., the great majority propagate from the seed. Thus the most important attribute of PI's lies in the cumulative effect on quality degradation of seed plates over many generations. A good description of many kinds of PI was reported by Johnson and Irvine.[46]

Suzuki et al.[13] reported a correlation between PI and the amount of Al impurity in the nutrient. Thus, study of PI's in growth in noble metal lined autoclaves and under low Al conditions would be informative.

The third possible source of dislocations is the difference in the lattice parameters between the seed and new growth resulting in misfit dislocations.[3] Misfit dislocations must be a significant source of dislocations since when dislocations begin at the seed surface, only a fraction of these are associated with a PI or are the continuation of threading dislocations from the seed. The remainder must be misfit dislocations. The simplest calculation shows that a lattice parameter difference of only 1 ppm would result in about 400 misfit dislocations/cm$^2$. This assumes that there is a perfectly sharp transition between the seed and the new growth. For solution growth, a sharp transition is unlikely.

Cohen and Sumner[47] and Brice[48] have summarized the data on the lattice parameters of natural and cultured quartz. All impurities increase the lattice parameters but the effect of $OH^-$ is negligible. The effect of other elements is roughly proportional to their ionic radii.[49] The variation of the c parameter is smaller than for the a parameter and can be ignored for the purposes of this discussion. When results from samples having extreme values of impurities are ignored, the a-parameters for cultured and natural quartz are about 4.9136Å and 4.9132Å respectively. The difference is 80 ppm and the two populations overlap. There is a rather large uncertainty in this difference when the uncertainty of measuring lattice parameters by the powder method[50] are considered. The variability of natural quartz is much larger than cultured. Nevertheless, it is clear that the average natural quartz seed would be an effective source of misfit dislocations.

There are quantitative measurements in the literature of the lattice parameter differences between seed and growth. Yoshimura, Miyazaki et al.[52] gave results which are summarized in Table 5. An increase in impurities in the seed and increase in $\frac{\Delta a}{a}$ seems to correlate with an increase in dislocations. However, it is not clear how many of the dislocations were initiated at the seed surface and how many may be threading dislocations propagating from already existing dislocations in the seed. Zarka, Lin and Sauvage[53] showed that a lattice parameter difference between seed and growth of 30 ppm generated new dislocations.

Considering the case where quartz was grown on a natural seed, topographs in the literature show a wide range of dislocation densities. Spencer and Haruta[51] show a topograph of Z growth on a DLF natural seed and state that all of the dislocations in the growth are due to particles on the seed surface. Barns et al.[36] got DLF growth on DLF seeds. Croxall et al.[5] found that a small number of dislocations originated at the seed surface in Z growth on a DLF natural seed. These dislocations were said not to have been caused by particles. Contrast at the seed surface showed that there was a lattice parameter mismatch between seed and growth. Dowsett et al[43] showed a topograph with a fairly large number of dislocations originating at the seed surface in Z growth on a natural seed. The presence of particles was not mentioned. This variability of experience is consistent with the range of lattice parameters in natural and cultured, i.e., chance in many otherwise well controlled experiments determines whether DLF growth is obtained and depends on the accidental difference between the lattice parameters of the seed and the growth.

From the foregoing we can see that measures which reduce inclusions and decrease lattice mismatch and techniques which employ DLF seeds or at least the best seeds possible are required for DLF growth or for significant reduction in dislocation density. Low dislocation density is an assured route to low channel density crystals.

## 4.0 Conclusions

r-bar growth has been shown to be preferable to 5°X for SAW devices because of homogeniety and yield.

Al reduction is the key to radiation hardness. Nutrient recrystallization and the use of high purity glass nutrient are strategies for reducing Al. For recrystallization purification, many recrystallizations are required because of the intrinsic inefficiency of hydrothermal recrystallization which has a separation efficiency like single pass zone refining. High purity glass nutrient shows promise for high purity growth but special procedures for controlling nucleation and growth rate are necessary in using glass. "Seeding" with small quantities of crystalline quartz added to the glass is particularly effective in regulating growth.

Reducing the latent etch channel density of cultured quartz requires reduction of the dislocation density although some reduction may be possible by reducing the decoration of dislocations and sweeping may be used to plug channels making them resistant to the etchants used in photolithographic processing. Low dislocation quartz can only be grown on low dislocation seeds. High quality seeds can be obtained from natural quartz but this requires expensive screening. Seeds in which the existing dislocations lie nearly parallel to the intended growth surface can be cut from cultured quartz and this, coupled with careful seed selection, is the strategy most likely to allow the maintenance of seed quality over many generations. Other factors which are important to the control of dislocations (and

hence etch channels) are the reduction of inclusions and minimizing the lattice mismatch between seed and growth.

Quartz suitable for the most demanding applications has been grown in the laboratory by using high purity nutrient and nearly perfect seeds. The challenge of routine production of large crystals remains but already r-bar quartz for demanding SAW application shows great promise and can be produced under factory conditions.

**Acknowledgements**

We wish to thank Alton Armington (Rome Air Dev. Center, Hanscom AFB, MA) for discussions concerning the use of $SiO_2$ glass as nutrient, T. Y. Kometani for atomic absorption analyses and D. L. Nash for emission spectrography. A. C. Longton and K. A. Jackson gave great encouragement and much advice.

**Table 1 - Growth Conditions**

| | |
|---|---|
| Growth Solution | 1M NaOH |
| % Fill | 83% |
| Dissolving Temperature | 400°C |
| Growth Temperature | 350°C |
| Pressure | 22,000 psi |
| Baffle | 5% open |
| Seed Plates | 5°X or r-bar |
| Growth Rate* | |
| (100% crystalline nutrient) | 5°X, ~40 mils/day |
| (total increase in thickness) | r-bar, ~20 mils/day |

* increase in thickness of seed/da, 40 mil = 1 mm.

**Table 2**

Insertion Loss in 1.7 GHz SAW devices:

| Source | Insertion Loss (dB) | Number of Devices Measured |
|---|---|---|
| Supplier #1 34° cut (basal) | 14.4 ± 1.1 | 17 |
| Supplier #2 34° cut (basal) | 15.3 ± 1.8 | 11 |
| AT&T 34° cut (r-bar) | 14.1 ± 1.0 | 50 |

**Table 3. Al Content of Quartz Grown Using Various Nutrients.**

| Nutrient | Grown Crystal | ppm Al |
|---|---|---|
| Milky Arkansas | | 32 |
| Milky Arkansas | r-bar-1 | 65* |
| r-bar-1 | r-bar-2 | 58* |
| r-bar-2 | r-bar-3 | 51* |
| r-bar-3 | r-bar-4 | 47* |
| 5°X-1 | r-bar-a | 53* |
| 5°X-2 | r-bar-b | 32* |
| 5°X-3 | r-bar-c | 24* |

* (Al analysis in r-bar growth).

**Table 4. Experiments with Glass Nutrient.**

| Run No. | Nutrient Glass | + Quartz | Rate on 5°X* | Rate on r-bar* |
|---|---|---|---|---|
| Q-1 | 0 | + 40 gm | 40 mil/da. | ~20 mil/da. |
| Q-2 | 24 gm | + 12 gm | 20 mil/da. | 7 mil/da. |
| Q-3 | 34 gm | + 10 gm | 40 mil/da. | 15 mil/da. |
| Q-4 | 40 gm | + 5 gm | 32 mil/da.** | 12 mi/da.** |
| Q-5 | 40 gm | + 0 | *** | |

* Rate is total increase in thickness/da., 40 mil = 1 mm.
** ΔT reduced to aid in control.
*** Poor control, flocculent product + glass which could not be easily removed.

Table 5

**Lattice Parameter Differences (Δa/a) and (Al + Na) Impurity Concentrations (Ref. 52)**

| | Δa/a ppm for Z-growth rel. to seed | (Al + Na) Conc. in Various Regions | | | | Disloc. density per $cm^2$ |
|---|---|---|---|---|---|---|
| | | seed | Z-growth | +X-growth | −X-growth | |
| Xtal 1 | −14 | 27 ppm | 3 ppm | 16 ppm | 49 ppm | ~50 |
| Xtal 2 | −70 | 200 | 2 | 22 ppm | 59 | ~500 |

## References


[1] R. A. Laudise and R. L. Barns, "Perfection of Quartz and Its Connection to Crystal Growth", IEEE J. on Ultrasonics, vol 35, pp. 277-88, 1988.

[2] J. C. Brice, "Crystals for Quartz Resonators", Rev. Mod. Phys., vol 57, pp 105-146, 1985.

[3] W. E. Bulst and E. Willibald, "Ultrareproducible SAW Resonator Production", in Proc. of the 36th Annual Freq. Cont Symp., pp 442-452, 1982.

[4] K. L. Bye and R. S. Cosier, "An X-Ray Double Crystal Topogrphic Assessment of Defects in Quartz Resonators", J. Matl. Sci., vol 14, pp 800-810, 1979.

[5] D. F. Croxall, I. R. A. Christie, J. M. Holt, R. J. Isherwood and A. G. Todd, "Growth and Characterization of High Purity Quartz", in Proc. of the 36th Annual Symp. of Freq. Cont., pp 62-68, 1982.

[6] "Quartz for the National Defense Stockpile", Natl. Materials Advisory Board of NRC, NMAB-424, Natl. Acad. Press, Wash. DC, 1985.

[7] D. L. Griscom, "Point Defects and Radiation Damage Processes in a-Quartz", in Proc. of the 33rd Annual Freq. Cont. Symp., pp 98-109, 1979.

[8] J. C. King, "Analasticity of Synthetic Crystalline Quartz at Low Temperature", Phys. Rev., vol 109, no. 5, pp 1552-1553, 1958.

[9] B. R. Capone, A. Kahan, R. N. Brown and J. R. Buckmeister, "Quartz Crystal Radiation Effects", IEEE Trans. on Nuc. Sci., vol NS-17, pp 217-220, 1970.

[10] D. Chakraborty, "On the Incorporation of Metallic Impurities in Synthetic Quartz Single Crystals", J. Crvst. Growth, vol 44, pp 599-603, 1978.

[11] T. J. Young, D. R. Koehler and R. A. Adams, "Radiation Induced Frequency and Resistance Changes in Electrolyzed High Purity Quartz Resonators", in Proc. of the 32nd Annual Freq. Control Symp., pp 34-42, 1978.

[12] H. Iwasaki, F. Iwasaki, C. K. Suzuki, V. A. R. Oliveira, D. C. A. Hummel and A. H. Shinohara, "Characterization of Brazillian Lascas from Various Regions and Their Use For Synthetic Quartz Growth, Part 1: Lascas Study", in Proc. of the 40th Annual Freq. Cont. Symp., pp 39-46, 1986.

[13] C. K. Suzuki, H. Shinohara, A. R. Oliveira, S. Takiya, and J. Kiss, "Characterization of Brazillian Lascas from Various Regions and Their Use for Synthetic Quartz Growth, Part II, Properties Correlation". in Proc. of the 40th Ann. Freq. Cont. Symp. pp 47-53, 1986.

[14] E. D. Kolb, K. Nassau, R. A. Laudise, E. E. Simpson and K. M. Kroupa, "New Sources of Quartz Nutrient for the Hydrothermal Growth of Quartz", J. Cryst. Growth, vol 36, pp 93-100, 1976.

[15] W. G. Pfann, "Zone Melting", John Wiley and Sons, Inc., NY, p. 25, 1958.

[16] J. F. Balascio and A. F. Armingon, "Development Results for the Production of High Quality Quartz", in Proc. of the 40th Annual Freq. Cont. Symp., pp 70-75, 1986.

[17] A. F. Armington and J. F. Balascio, "The Growth of High Quality Quartz in Commercial Autoclaves", in Proc. of the 39th Annual Freq. Cont. Symp., pp 230-233, 1985.

[18] A. F. Armington and J. F. Balascio, "The Growth of High Purity, Low Dislocation Quartz", in Proc. of the 38th Annual Freq. Cont. Symp., pp 3-7, 1984.

[19] J. F. Balascio and A. F. Armington, "Growth and Sweeping of High Quality Quartz", 8th Quartz Devices Conf. and Exhib., Electron. Ind. Assoc., 2001 Eye St., NW, Wash, DC 20006, pp 221-228, 1986.

[20] J. F. Balascio and N. C. Lias, "Manufacturing Technology for High Quality Swept-Culture Growth Quartz", Final Report, Motorola, Inc., Carlisle, PA 17013, USAF Contract F19628-82-C-0067, March 13, 1986.

[21] A. F. Armington, J. J. Larkin, J. J. O'Connor, J. E. Cormier and J. A. Horrigan, "Effect of Seed Treatment on Quartz Dislocations", in Proc. of the 37th Annual Symp. on Freq. Control, pp 177-180, 1983.

[22] R. J. Baughman, "A Comparison of Quartz Crystals Grown from Fused Silica and from Crystalline Nutrient", in Proc. of the 36th Annual Freq. Cont. Symp., pp 82-89, 1982.

[23] G. W. Arnold, "Defects in Quartz Crystals", in Proc. of the 11th Annual Freq. Cont. Symp., pp 112-129, 1957.

[24] J. W. Nielsen and F. G. Foster, "Unusual Etch Pits in Quartz Crystals", Am. Mineral, vol 45, pp 299-210, Mar-Apr., 1960.

[25] F. Iwasaki, "Line Defects and Etch Tunnels in Synthetic Quartz", J. Cryst. Growth, vol 37, pp 91-100, 1977.

[26] T. R. Meeker and A. J. Miller, "The Temperature Coeff. of Frequency of AT-Cut Resonators Made from Cultured r-Face Quartz", in Proc. of the 34th Annual Freq. Cont. Symp., pp 85-92, 1980.

[27] J. R. Hunt, "Etch Pits and Channels in Swept At and SC-Cut Quartz", in Proc. of the 41st Annual Freq. Cont. Symp., pp 183-191, 1987.

[28] G. R. Johnson and R. A. Irvine, "Etch Channels in Single Crystal Cultured Quartz", in Proc. of the 41st Annual Freq. Cont. Symp., pp 175-182, 1987.

[29] A. R. Lang and V. F. Miuskov, "Dislocations and Fault Surfaces in Synthetic Quartz", J. Appl. Phys., vol 38, pp 2477-2483, 1967.

[30] U. Alter and G. Voigt, "Direction Change of Dislocations on Passing a Growth Sector Boundary in Quartz Crystals", Cryst. Rsch. Technol., vol 19, pp 1619-1623, 1984.

[31] J. G. Gaultieri and J. R. Vig, "Sweeping and Irradiation Studies in Quartz", in Proc. of the 38th Annual Freq. Cont. Symp., pp 42-49, 1984.

[32] J. G. Gaultieri, "The Influence of Temperature and Electric Field on the Etch-Channel Density in Swept Cultured Quartz", in Proc. of the 39th Annual Freq. Cont. Symp., pp 247-254, 1985.

[33] J. R. Vig, J. W. LeBus and R. L. Filler, "Chemically Polished Quartz", in the Proc. of the 31st Annual Freq. Cont. Symp., pp 131-143, 1977.

[34] J. Gaultieri, "Further Studies on Electrode Diffusion Suppressed Swept Quartz", in Proc. of the 41st Annual Freq. Cont. Symp., pp 192-198, 1987.

[35] S. Yamashita, A. Shinomiya and H. Kumasaki, "Content Correlation of Fe and Alkali Metals in Synthetic Quartz", J. Cryst. Growth, vol 30, pp 27-28, 1975.

[36] R. L. Barns, P. E. Freeland, E. D. Kolb, R. A. Laudise and J. R. Patel, "Dislocation Free and Low Dislocation Quartz Prepared by Hydrothermal Crystallization", J. Cryst. Growth, vol 43, pp 676-786, 1978.

[37] D. F. Croxall, I. R. A. Christie, J. M. Holtz, B. J. Isherwood and A. G. Todd, "Growth and Characterization of High Purity Quartz", in Proc. of the 36th Annual Freq. Cont. Symp., pp 62-65, 1982.

[38] A. Zarka, L. Liu and M. Buisson, "Influence de la Localisation Sectorielle du Germe sur la Qualite

Crystalline de Quartz de Synthese", J. Cryst. Growth, vol 54, pp 394-398, 1981.

[39] A. F. Armington and J. J. Larkin, "The Growth of High Purity, Low Dislocation Quartz", J. Cryst. Growth, vol 71, pp 799-802, 1985.

[40] A. F. Armington and J. J. Larkin, "The Growth of High Quality Quartz in Commercial Autoclaves", in Proc. of the 39th Annual Freq. Cont. Symp., pp 230-233, 1985.

[41] A. F. Armington, J. A. Horrigan, M. T. Harris and J. F. Balascio, "A Study of Dislocations and Inclusions in Alpha Quartz", in Proc. of the 41st Annual Freq. Cont. Symp., pp 213-215, 1987.

[42] J. F. Balascio and N. Lias, "Factors Affecting the Quality and Perfection of Hydrothermally Grown Quartz", in Proc. of the 34th Annual Freq. Cont. Symp., pp 65-71, 1980.

[43] J. Dowsett, D. F. G. Dwyer, F. Stern, R. A. Heinecke and A. H. Trulove, "Etch Processes of Bulk and Surfaces Wave Devices", in Proc. of the 39th Annual Freq. Cont. Symp., pp 301-309, 1985.

[44] J. Ashahara and Taki, "Physical Properties of Synthetic Quartz and Its Electrical Characteristics", in Proc. of the 26th Annual Freq. Cont. Symp., pp 93-103, 1972.

[45] J. C. Brice, "The Specification of Quartz for Piezoelectric Devices", in Proc. of the 38th Annual Freq. Cont. Symp., pp 487-495, 1984.

[46] G. R. Johnson and R. A. Irvine, "Etch Channel in Single Crystal Cultured Quartz", in Proc. of the 41st Annual Freq. Cont. Symp., pp 175-182, 1987.

[47] A. J. Cohen and G. G. Sumner, "Relationships Among Impurity Contents, Color Centers and Lattice Constants in Quartz", Am. Mineral., vol 43, pp 58-69, 1958.

[48] J. C. Brice, "The Lattice Constants of a-Quartz", J. Matls. Sci., vol 15, pp 161-167, 1980.

[49] R. D. Shannon and C. T. Prewitt, "Effective Ionic Radii in Oxides and Fluorides", Acta Cryst., vol B25, pp 925-946, 1969.

[50] W. Parrish, "Results of the IUCr Precision Lattice-Parameter Project", Acta. Cryst., vol 13, pp 838-850, 1960.

[51] W. J. Spencer and K. Haruta, "Defects in Synthetic Quartz", J. Appl. Phys., vol 37, pp 2557-2563, 1966.

[52] J. Yoshimura, T. Miyazaki, T. Wada, K. Kohra, M. Hosaka, T. Ogawa and S. Taki, "Measurement of Local Variations in Spacing and Orientation of Lattice Plane of Synthetic Quartz", J. Cryst. Growth, vol 46, pp 691-700, 1979.

[53] A. Zarka, L. Liu and M. Sauvage, "Studies of the Local Variations of Spacing and Orientation of Reticular Planes in Synthetic Quartz", J. Cryst. Growth, vol 62, pp 409-424, 1983.

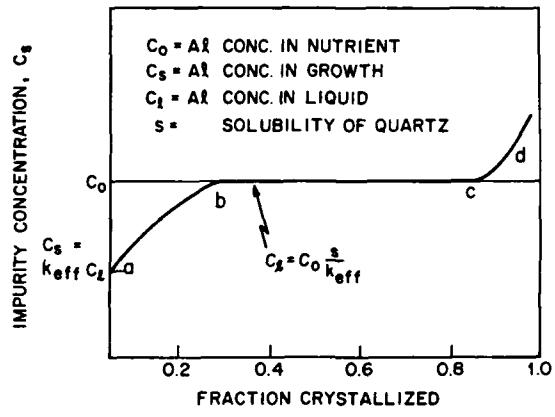



Fig. 1. Schematic of zone refining and hydrothermal nutrient recrystallization.

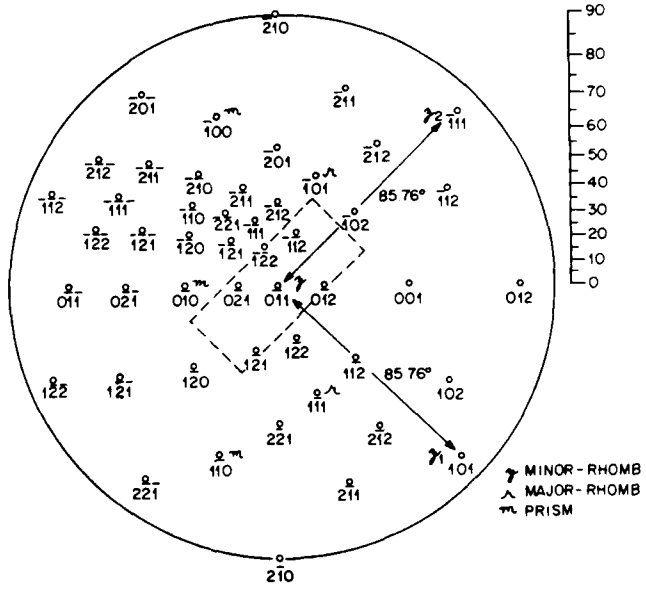



Fig. 3. Stereo projection of strategy for producing r-bar seed plates.

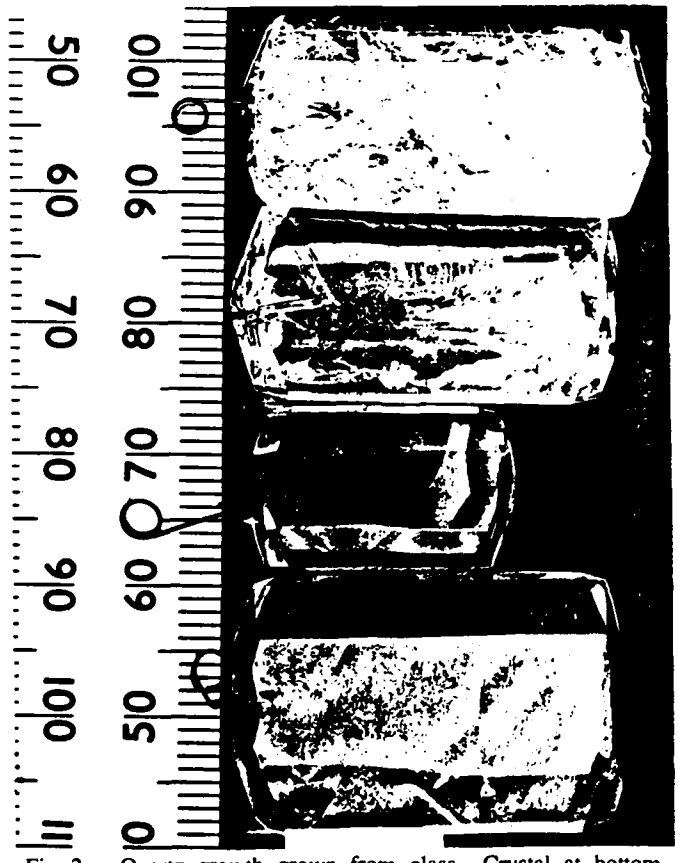


Fig. 2. Quartz growth grown from glass. Crystal at bottom (Q-1, Table 4) grown from crystalline nutrient. From bottom to top increasing fraction of glass nutrient. (Q-1, Q-2, Q-3, Q-4, Table 4). All crystals would be suitible for crushing and use as nutrient. Q-1 and Q-2 have macroscopic quality such that they could be used for devices.

42nd Annual Frequency Control Symposium - 1988

# RECENT EXPERIMENTS IN A SILVER LINED AUTOCLAVE

Robert Irvine, Jonathan Foise, Edward Leeson, and Gary Johnson

Sawyer Research Products, Inc.
35400 Lakeland Blvd.
Eastlake, Ohio 44094

## Summary

Eliminating impurities and lattice defects from single crystal cultured quartz ($SiO_2$) provide benefits to a variety of end-use quartz applications. Superior optical and electronic performance results from improved quartz. The culturing environment, growth technology and process control influence impurity contamination and imperfections in single crystal quartz. Reducing impurities and lattice defects depends on optimizing the process parameters. Under funding from DoD LABCOM, Sawyer has established a research program with the following objectives:

Develop hydrothermal quartz culturing equipment and processes capable of economically producing quartz which is:

a) etch processable and radiation hard for electronic applications,

b) low in strain, inclusions and fringe distortion for optic/laser applications.

Specific goals are:

Etch Channel Density: $<10/cm^2$
Inclusion Density: $<0.05/cm^3$

Impurity Concentrations in PPB:

| | |
|---|---|
| Aluminum: | <200 |
| Lithium: | <300 |
| Sodium: | <500 |
| Potassium: | < 50 |
| Iron: | <100 |

Other objectives include the elimination of seed veil, seed splicing, the search for useful quartz characterization methods, and defining the limits of existing quartz culturing technology.

## Introduction

In mid-1986, Sawyer Research Products, Inc. started this research effort with funding from DoD LABCOM under contract DAAL01-86-C-0012. It was defined as a multiphased effort to cover a period of 23 months. The initial phase covers the production of the highest quality quartz given state of the art sodium carbonate process equipment. The second phase seeks to establish definitive relationships between process variables and quartz quality. Given the results obtained in Phase I and II, the third phase includes the production of high purity quartz based on optimizeed process variables.

## Discussion

A significant portion of the research effort involves the design, fabrication and use of a silver liner in full scale production autoclaves. Liner research addresses the question of culturing environment purity and the impact on crystal growth.

Two liners were designed to provide isolation from impurities associated with the steel walls of the autoclaves and scale normally found on internal hardware. Experiments with a pressure balanced and weld-seal design were planned as part of the research. The pressure balanced liner was completed in June 1987.

FIGURE 1

Figure 1 shows the full size liner which required the use of 4000 t.o. pure silver, including solid silver internal hardware and closure heads. The liner can be installed in any of Sawyer's 330 mm by 3 meter autoclaves. It can be removed between runs for thorough cleaning and reconditioning.

During the period June-October 1987, the first liner was installed and run through a series of calibration experiments. The focus of the calibration was temperature control, heat flow and run management. Actual performance was compared to design calculations. Based on the results, the liner was modified to achieve the temperature and heat flow characteristics necessary for run management.

CH2588-2/88/0000-127 $1.00 C 1988 IEEE

Critical to the management of crystal growth is the accurate measurement of internal temperatures. Normally accomplished with steel encapsulated thermocouples (TC), a pure silver sheath was designed to isolate the steel in the autoclave. This design permitted the measurement of internal axial temperature profiles without risk of contaminating the process. Calibration runs verified that the TC isolation arrangement was suitable for temperature measurement. The heat flow characteristics of the liner in the autoclave were studied in subsequent calibration runs.

The use of silver liners in crystal growth is common. Several theories discuss the impact of heat conduction through the liner on heat flow and internal temperatures. The liner design permitted the investigation of this phenomena. The heat flow and temperature characteristics of an unlined autoclave (AC) are illustrated in Figure 2a. Steady state temperature drops across the AC wall and associated power input were established to achieve a $\Delta T$ of 6°C. Installation of the liner, Figure 2b, resulted in a substantial increase of power input, 55% vs. 30%, and larger temperature drop across the wall to achieve a $\Delta T_{(maximum)}$ of 4°C.

When the axial temperature profile of this calibration run is examined in Figure 3, a sharp transition is depicted between the AC dissolving and growing chambers. Large heat conduction up the liner body would degrade this transition. Heat convection in the pressure balance fluid outside the liner was identified as the possible cause of the high power/low $\Delta T$ conditions.

A seal shown in Figure 4 was welded to the liner O.D. at the baffle elevation to restrict the convection flow fluid between the liner and the AC wall. Baffles internal to the liner are normally 10-30% open area. The O.D. baffle was installed with 3-4% open area, sufficient to allow fluid to pass during the loading of the run. The temperature and heat flow characteristics of this configuration are depicted in Figure 2c. Although temperature drops across the wall and power input are slightly higher than under unlined conditions, it is apparent that convective flow, rather than conduction, was responsible for limited $\Delta T$ and increased power input requirements. The axial temperature profile of the calibration run with restricted outside flow is illustrated in Figure 5. A sharp transition occurs at the baffle level and a $\Delta T$ of 7°C was attained well below the operating limits of the system.

## HEAT FLOW CHARACTERISTICS

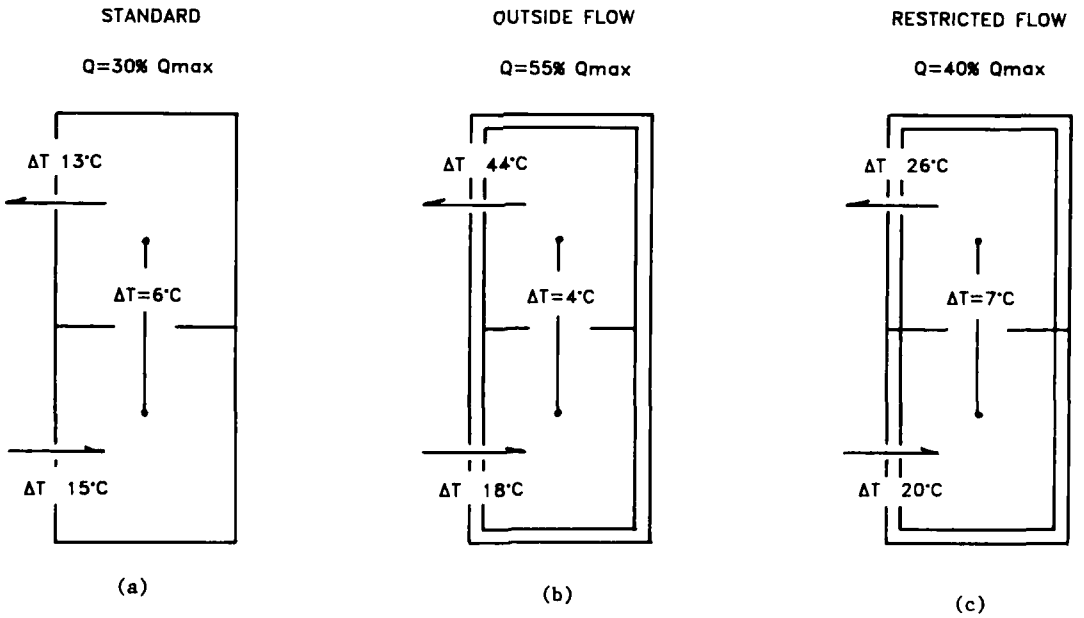



FIGURE 2

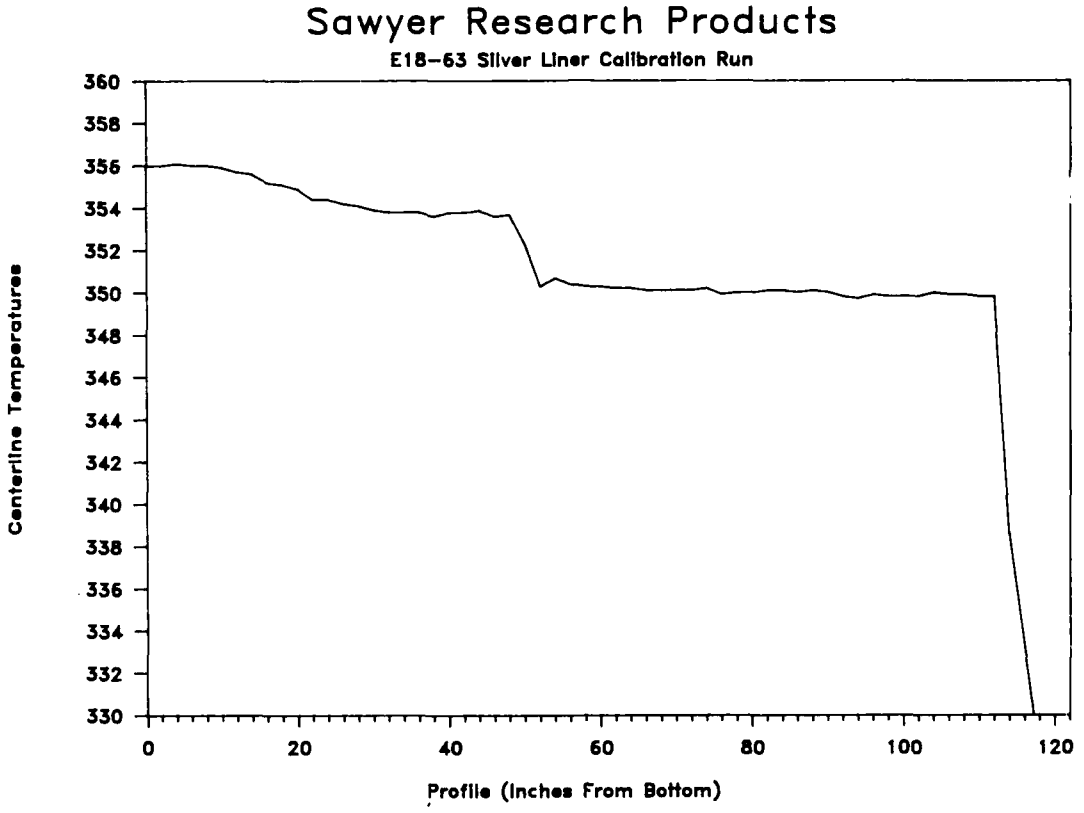



FIGURE 3

FIGURE 4

The calibration runs, modifications to the liner and temperature profile data, demonstrated the run management capability of the lined AC system. After removal and cleaning, the liner was installed in an AC for the first high purity, pressure balanced production run.

Preparation for the production run included high purity inputs to the process. Table I summarizes the ICP-MS analytical results of impurity levels in the sodium carbonate powder, solution and nutrient quartz. Impurity levels for solution sampled from inside the liner after the run are included for comparison to the original inputs. A total of 145 kg high purity, previously cultured pure-Z quartz was used for nutrient. Seeds numbered 84 and were Z plates 38 mm in the X direction. The internal TC assembly accommodated 2 TC and extended to 150 mm below the baffle level. The liner was pressure balanced.

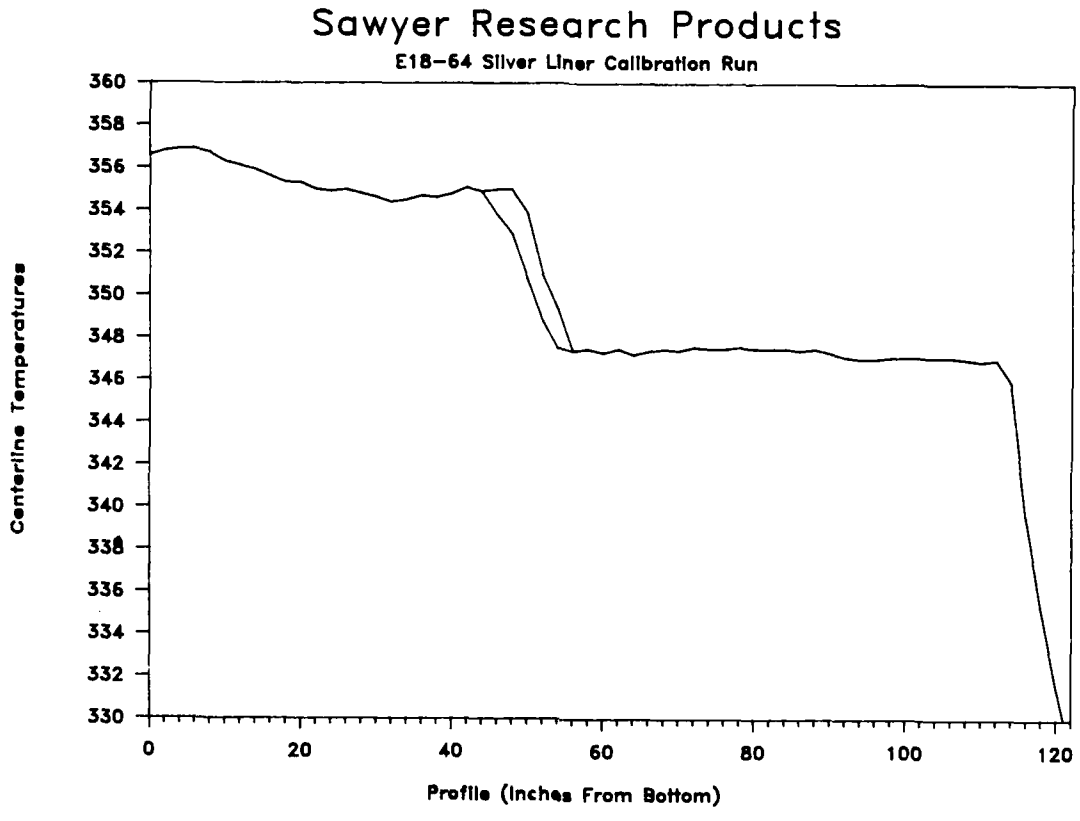



FIGURE 5

## PRODUCTION RUN INPUTS (PPM)*

| | Al | Ca | Fe | Li | Ag | K | Mg | Na |
|---|---|---|---|---|---|---|---|---|
| $Na_2CO_3$ Powder | 0.25 | 5.82 | <0.75 | <0.02 | NT | NT | 0.075 | NT |
| $Na_2CO_3$ Solution | <0.05 | NT | <0.06 | 0.005 | NT | NT | 0.006 | NT |
| Pure Z Nutrient | 0.84 | <0.67 | <0.10 | 0.41 | NT | NT | <0.009 | NT |
| Solution After Run | <0.006 | 7.44 | <0.11 | 2.94 | <0.002 | 1.19 | <0.003 | NT |

*ALL TESTS CONDUCTED WITH ICP-MS
NT = NOT TESTED

TABLE I

The length of the run was 97 days. Figure 6 illustrates the temperature management over the duration of the run. A target growth temperature of 346°C was attained. Note the dramatic increase in temperature drop across the AC walls experienced near the conclusion of the run. Little change in the $\Delta T$ of the process occurred in spite of high heating and cooling rates across the walls of the AC. Axial temperature profiles were recorded during the run at the 4th, 56th and 95th day of operation. The 4th day profile is shown in Figure 7. This profile indicates a stable growing chamber temperature, sharp transition at the baffle and a $\Delta T$ of 4°C. While some fluctuation in growth chamber temperature exists in Figure 8, the 56th day, the transition is sharp. A $\Delta T$ of 9.5°C is attained. The profile on the 95th day, Figure 9, shows marked instability in the growing chamber, a poor transition at the baffle and little change in the $\Delta T$ from the previous profile, Figure 8.

By positioning thermocouples near the baffle, these temperature fluctuations were studied over time and recorded on a strip chart. Figure 10 depicts the strip chart data for thermocouples located at, and 300 mm above, the baffle elevation in the AC. With this data, it is possible to relate fluctuations to time and estimate fluid velocity in the AC. At these elevations, the temperature fluctuations are nearly 180 degrees out of phase with a period of approximately 37 seconds. This translates to a fluid velocity at the baffle of 8.15 mm/second. This value is of the same order of magnitude as other estimates. The significance of fluid velocity will be studied in later phases of the program. This measurement serves as the baseline.

The power input to the process was monitored over the total duration of the run. Figure 11 traces the power input, expressed as a percent of the total system capability. The increase in power input towards the end of the run correlates with the increase in temperature drop across the AC walls shown in Figure 6. As an important process variable, the impact of power on crystal quality will be illustrated by later comparison to the IR trace of the cultured crystal.

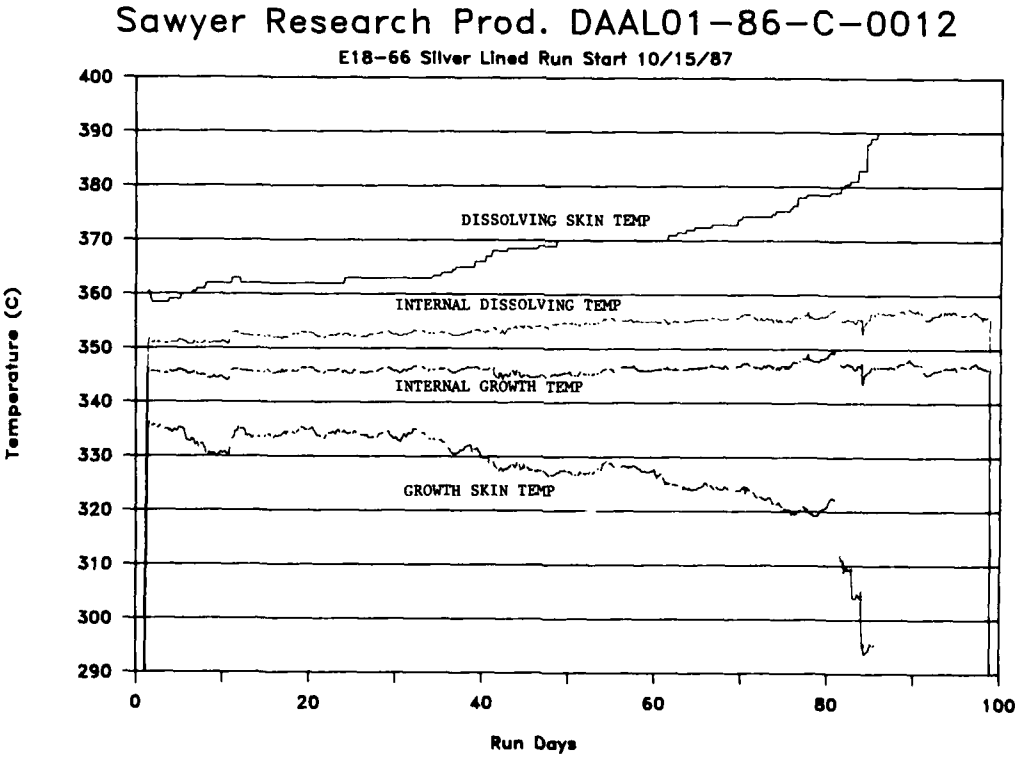



FIGURE 6

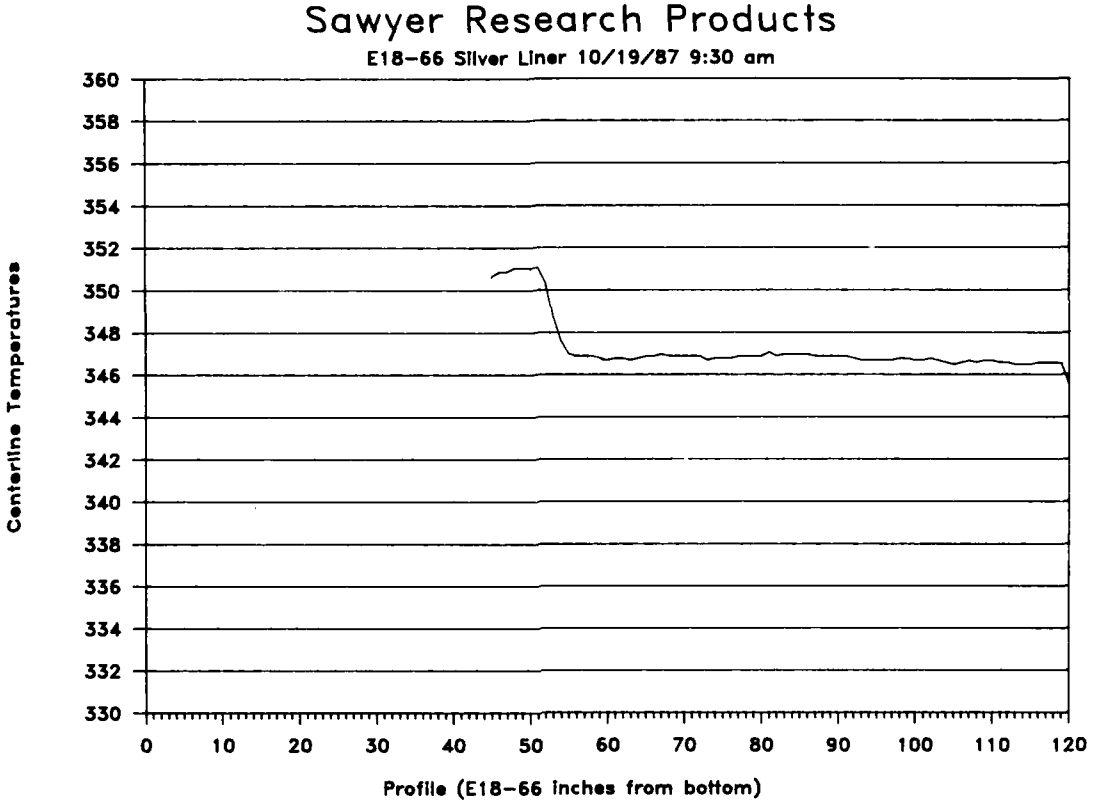



FIGURE 7

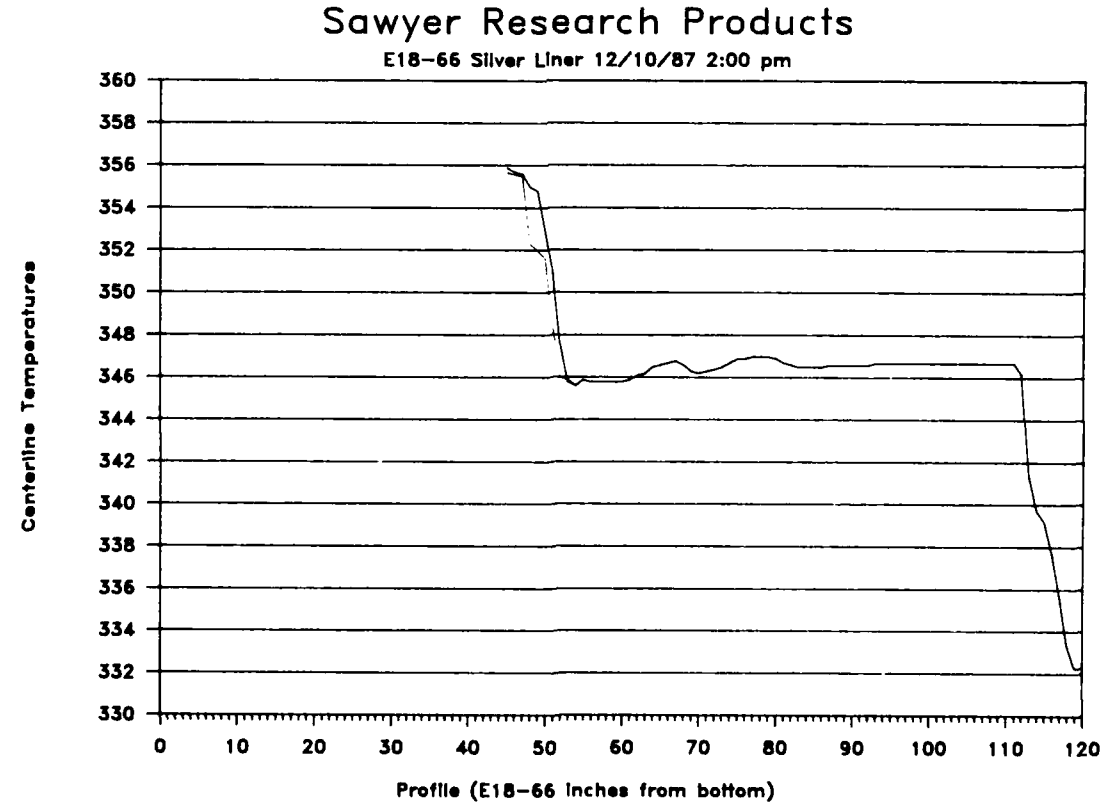



FIGURE 8

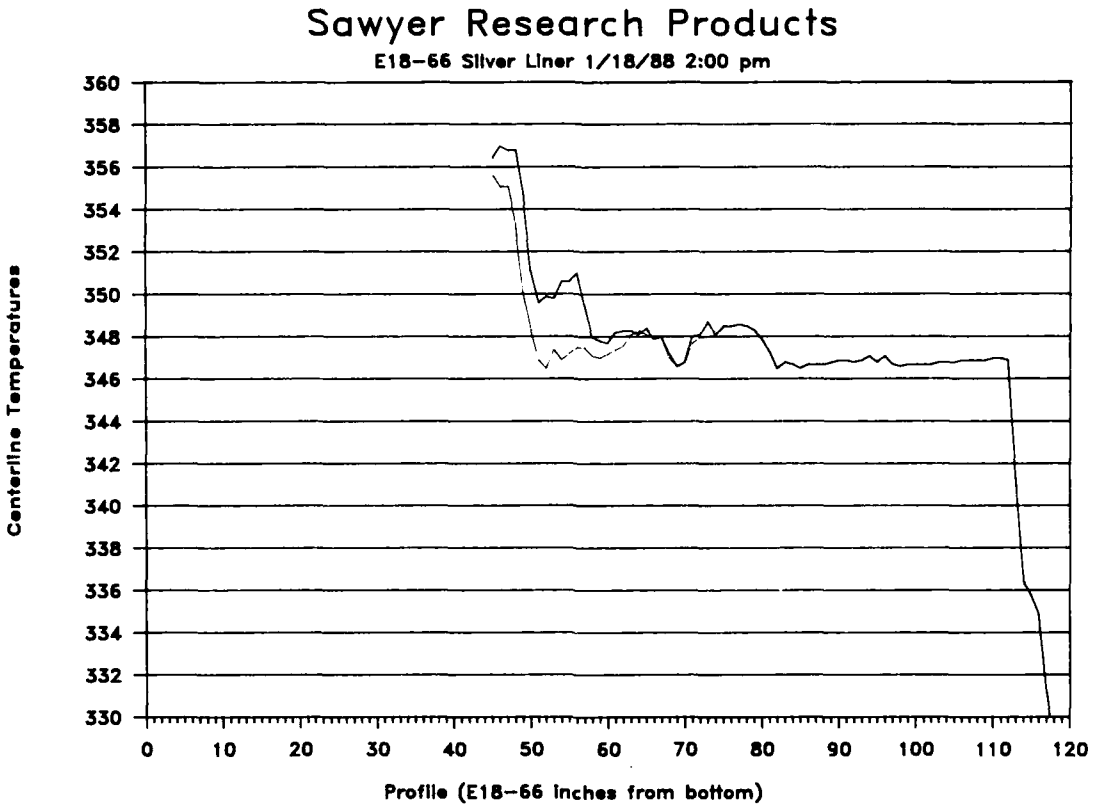



FIGURE 9

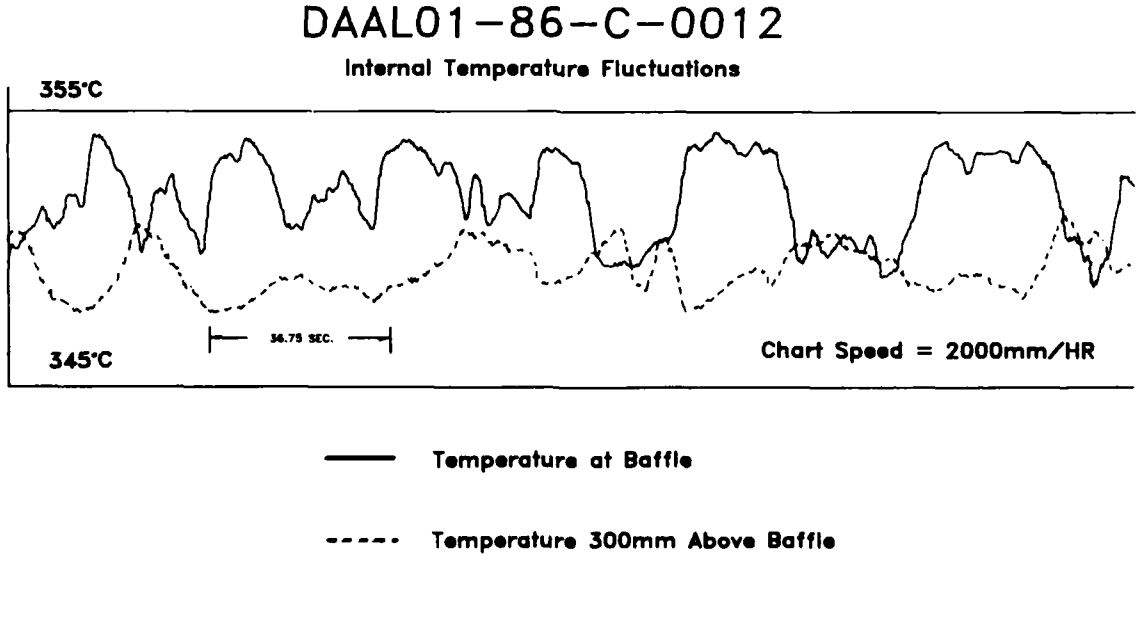



FIGURE 10

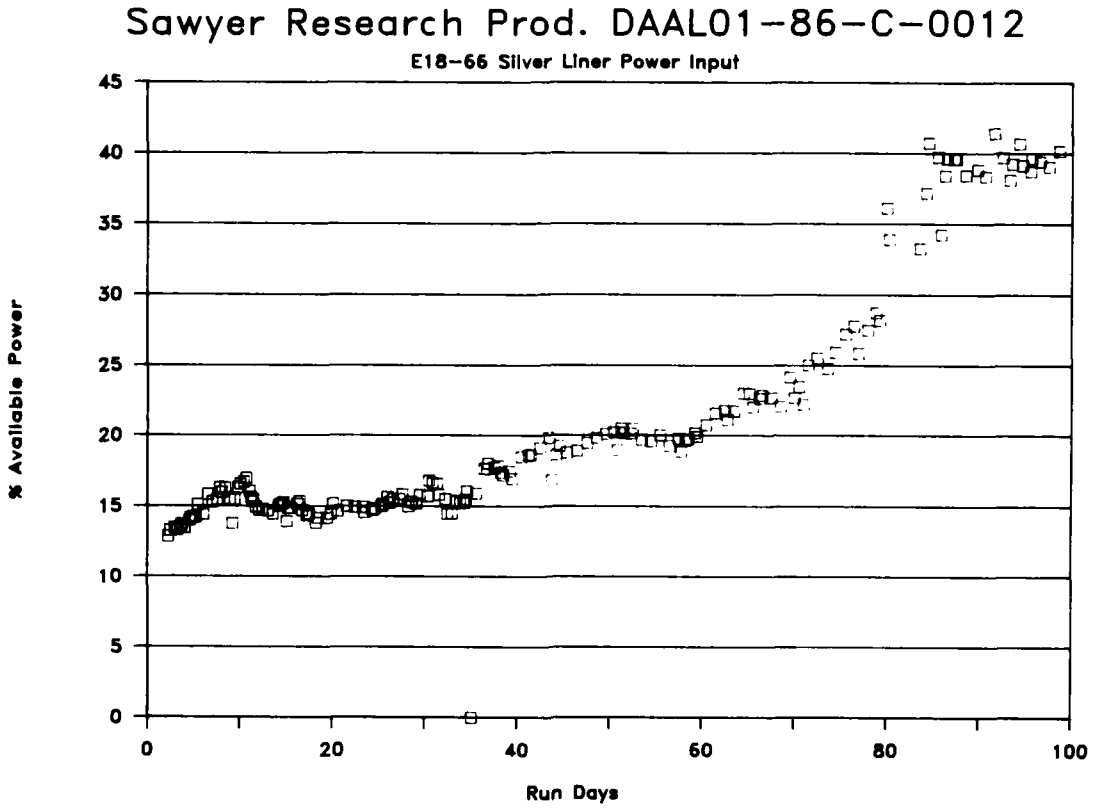



FIGURE 11

The general run management results are listed in Table II. Comparison to targets reveals that the run was on target with the exception of pressure and growth rate. The pressure shortfall may be attributable to an error in free volume measurement. The growth rate increase can be directly related to the elevated power input to the run.

## RUN MANAGEMENT RESULTS

| | Target | Actual |
|---|---|---|
| Growing Chamber Temperature | 346°C | 346.5°C |
| Dissolving Chamber Temperature | 350-358°C | 351-358°C |
| ΔT Start | 4.5°C | 5.5°C |
| ΔT End | 12.0°C | 9.7°C |
| Pressure | 12,000 PSI | 9,200 PSI |
| Run Days | 95 | 97 |
| Growth Rate | 0.41 mm/DAY | 0.56 mm/DAY |

TABLE II

Opening the AC revealed a collapsed seed rack. Measurements indicate the rack compressed 400 mm from the original height. Shown in Figure 12, it is likely that the collapsed rack caused the temperature fluctuations and poor profile characteristics described earlier. The solid silver seed rack design was insufficient to carry the weight of grown crystals. A silver plated steel rack will be used on subsequent experimental runs. This adds the risk of Fe contamination and will be monitored as part of future experiments.

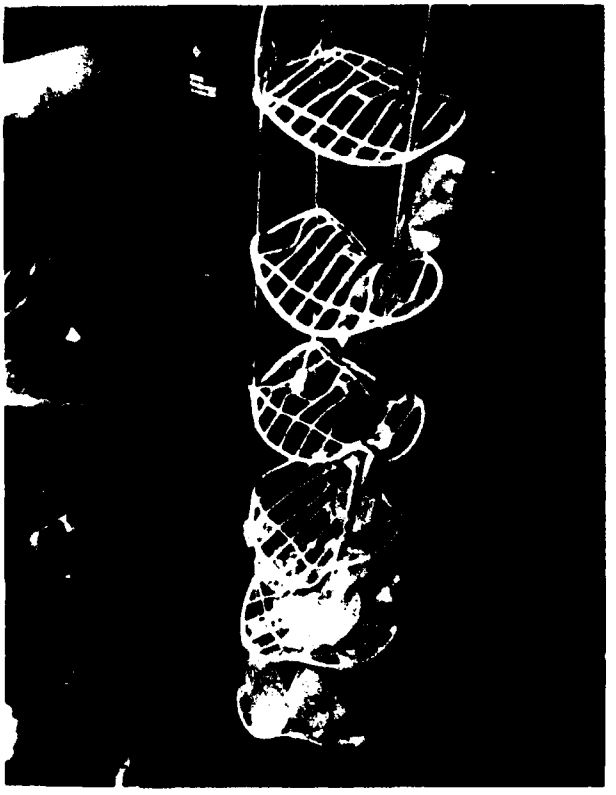


FIGURE 12

Characterizing the grown crystals was part of the effort. In total, 121 kg of crystal was cultured. Twenty (20) crystals were well formed with the remainder adversely affected by varying degrees of intergrowth with adjacent crystals or the seed rack. The crystals averaged 1.53 kg with a mean Z dimension of 55 mm.

The ICP-MS analysis of solution after the run in Table I indicated higher levels of calcium and lithium. Since the AC carried residual lithium in the scale and no lithium was added internally to the liner, this may be evidence of exchange between inside and outside fluids. Table III summarizes a comparison of the impurity levels found in the pure Z crystal growth and the targets established by the program. With the exception of lithium, all will require substantial reduction to meet the targets. The sodium level indicated is an order of magnitude higher than normally experienced in Sawyer quartz. Additional samples are being tested to confirm the sodium levels. Independent analysis for aluminum is in process by EPR.

## IMPURITY ANALYSIS (PPM)*

| | Al | Li | Na | K | Fe |
|---|---|---|---|---|---|
| Target | 0.2 | 0.3 | 0.5 | 0.04 | 0.1 |
| Pure Z Crystal | 0.7 | 0.3 | 64.4 | 0.3 | 1.8 |

*ALL TESTS CONDUCTED WITH ICP-MS

TABLE III

Etch channel density for a sample crystal was determined using the EIA standard test procedure with the results illustrated in Figure 13. Examination of etch channel distribution suggests that the seed was relatively low in etch channels. The majority of channels may have formed at the interface of the seed and growing crystal. X-ray topographic analysis should reveal additional information regarding the generation of dislocations.

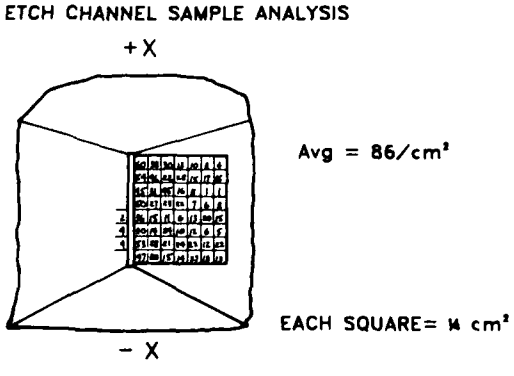



FIGURE 13

The inclusion levels, Figure 14, determined by microscopic analysis were low, but above the targeted values. The morphology of the bulk inclusions differed sharply from those normally found in Sawyer quartz. SEM-EDAX analysis may highlight any elemental differences which exist between the inclusions found in this quartz and quartz cultured in normal production runs. Figure 15 (100X) is representative of bulk inclusion morphology found in standard production quartz. Figure 16 (100X) illustrates the morphology of inclusions generally found in the silver lined run.

## INCLUSIONS

| SIZE | DENSITY |
|---|---|
| 25-70 MICRONS | 1.4/cc |
| 70-100 MICRONS | 0.3/cc |
| >100 MICRONS | 0.4/cc |
| TOTAL | 2.1/cc |
| TARGET | 0.05/cc |

FIGURE 14

FIGURE 15

FIGURE 16

Another notable difference between these and standard production crystals is the composition of the seed veils. Figure 17 recalls the standard seed veil comprised of a bubble plane, plates and debris on the seed surface. The seed veil of the silver lined run, Figure 18, is characterized by spike-like or dendritic projections from the seed surface. There is little debris on the seed surface. Pressure at the culturing conditions may have contributed to this veil morphology. SEM analysis may contrast the chemical composition of these seed veil morphologies.

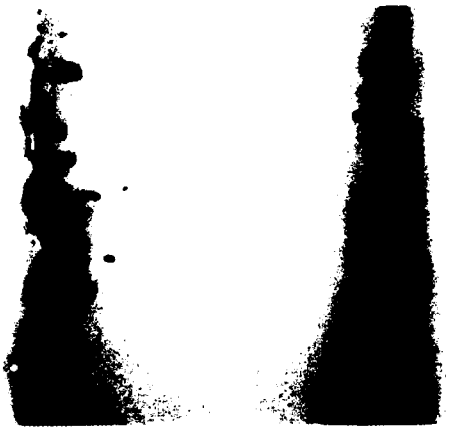


FIGURE 17

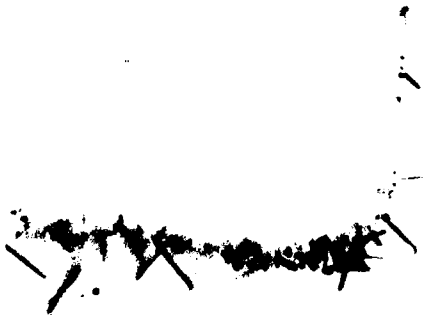


FIGURE 18

A quick review of Figure 11 helps one relate the dependence of crystal growth rate and IR absorption on power input in a run. The 3500/3800 $cm^{-1}$ IR absorption trace is reproduced in Figure 19. The spatial scan across the Z width of the crystal sample shows the direct relation between power and absorption. The apparent rapid growth and high power levels may have contributed to the impurity uptake in this quartz. The spatial distribution of impurities may substantiate this theory.

A final characterization method was borrowed from Sawyer's optical customers and applied to the quartz grown in this study. Shadowgraph imaging techniques have been used for years to contrast slight variations in refractive index in materials. The application of this technique to quartz reveals structural defects. The low cost and convenience of the method make it an attractive alternative or supplement to x-ray topographic techniques. Thick samples, up to 50 mm, have been successfully imaged with this technique which represents a nondestructive method of quartz analysis. Though the resolution is well below that of topographs, it provides a good qualitative view of the defect level in quartz. The results can be obtained immediately with minimal sample preparation. Figure 20 is a schematic of the technique and Figure 21 is the result of a shadowgraph image taken of a crystal Y-cut from the silver lined run. Note the detail associated with defects inherent in the seed and those generated at the surface of the seed. Also visible are the characteristic growth fronts in the -X material. There is sufficient resolution to view similar growth fronts in the Z material. The pattern of defects radiating in the Z direction is identical to those exhibited by dislocations shown in x-ray topographs. Shadowgraph images of quartz produced for this study will constitute part of the material characterization.

The research project enters the second phase associated with parametric testing. The emphasis is on developing a relation between power and growth characteristic in quartz. The second liner, a weld-seal design is in process. The internals have been redesigned to eliminate the structural problems experienced with the solid silver construction. Quartz produced in the first silver lined run is scheduled for further testing in electronic and optical applications.

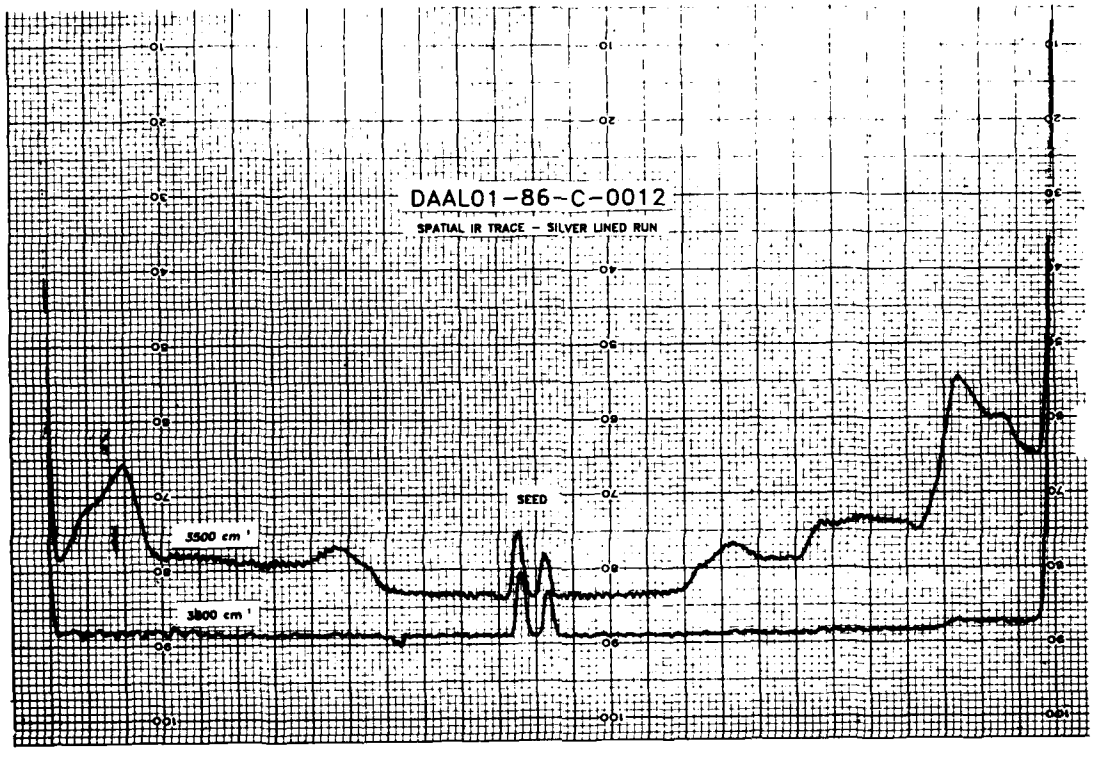



FIGURE 19

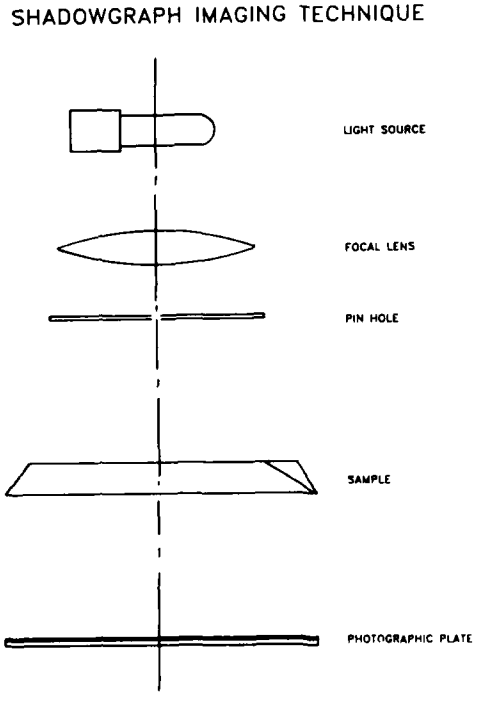



FIGURE 20

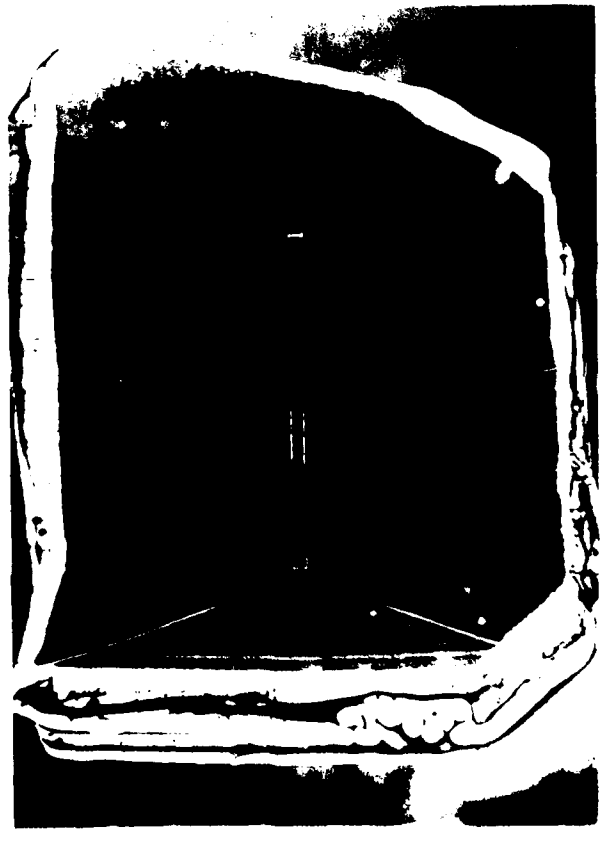


FIGURE 21

42nd Annual Frequency Control Symposium - 1988

# NEW ADVANCES IN CRYSTAL GROWTH OF HIGH PURITY BERLINITE: A NEW SOLVENT THE SULFURIC ACID

E. Philippot[1], A. Goiffon[1], J.C. Jumas[1], C. Avinens[1], J. Detaint[2], J. Schwartzel[2], A. Zarka[3]

[1] Laboratoire de Physico-Chimie des Matériaux Inorganiques - Université de Montpellier 34060 - FRANCE.

[2] Centre National d'Etudes des Télécommunications PAB/BAG/MCT Bagneux 92220 - FRANCE.

[3] Laboratoire de Minéralogie et de Cristallographie Université de Paris VI 75005 - FRANCE.

## Abstract

Systematic investigations of the solubility of berlinite in sulfuric acid as a function of temperature and solvent concentration have demonstrated that this solubility is much higher in sulfuric acid than in all other known solvents.

The crystal growth parameters were determined for the reverse temperature gradient method in 375cm$^3$ platinum lined autoclaves for different values of: the crystal growth temperature (Tc), the vertical temperature difference (ΔT=gradient), the acid concentration.

Infrared spectrometry was used to monitor the water concentration in crystals: this concentration was found to be principally dependent on the growth rate and the acid concentration. X-ray topography has shown a very low dislocation density.

Electrical measurements of resonators of several orientations have shown that the Q.f product can exceed 3.5 $10^{12}$ and the first and second order temperature coefficient of Y rotated resonators are lower than those of quartz resonators of corresponding orientations. The AT cut displays a cubic temperature behaviour and it was verified that k is for the AT cut twice that of AT quartz.

## Introduction

Since the work of Chang and Barsch[1] many works have been devoted to growing and characterizing berlinite crystals and to modelling and experiment with bulk and surface wave devices made with this material. It was successively shown that this material was a close piezoelectric analog of quartz, that possesses similar cuts with zero frequency temperature coefficients, either for bulk or surface wave applications.

In the past ten years, many efforts were made to develop hydrothermal methods that can produce large crystals of high perfection needed for piezoelectric applications[2-9]. Unfortunatly all crystals show an excessive amount of water incorporated during growth. Recently the importance of reducing the water solubility in crystals was emphasized in two papers[10,11] that show the large influence of this phenomenon on device properties (insertion losses, or Q factor, temperature coefficients ...). Transmission electron microscopy method[12] was used after high temperature precipitation of water to assess the $H_2O$ content of crystals as a function of some growth parameters: temperature and pressure.

According to this "concept", following these investigations, we have demonstrated, last year, the feasability of obtaining very high Q crystals with thermal behaviour superior to this of quartz. These growth cycles have been performed in $H_3PO_4$ acid at T>240°C. Unfortunatly, a major drawback of this method is the difficulty of transposition at an industrial scale:

- very small change of berlinite solubility in $H_3PO_4$ acid when T>230°C.
- seed dissolving before thermal equilibrium of autoclaves.

Instead of trying to modify drastically growth parameters, we have looked for another solvent. After some tests, we have choosen sulfuric acid, $H_2SO_4$. A preliminary study of berlinite in this medium with different concentrations shows a retrograde solubility, Figure 1, just as with previously considered acids.

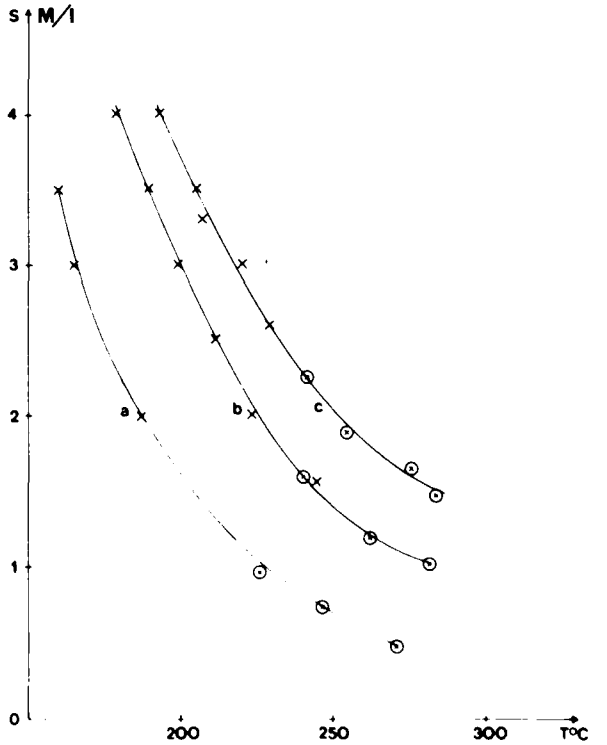



Figure 1 : $AlPO_4$ solubility in $H_2SO_4$ a=5M b=7.5M c=9M.

CH2588-2/88/0000-138 $1.00 C 1988 IEEE

If we compare this solubility to that in other acids, with the same concentration, $H_2SO_4$ shows a much higher solubility of berlinite, Figure 2.

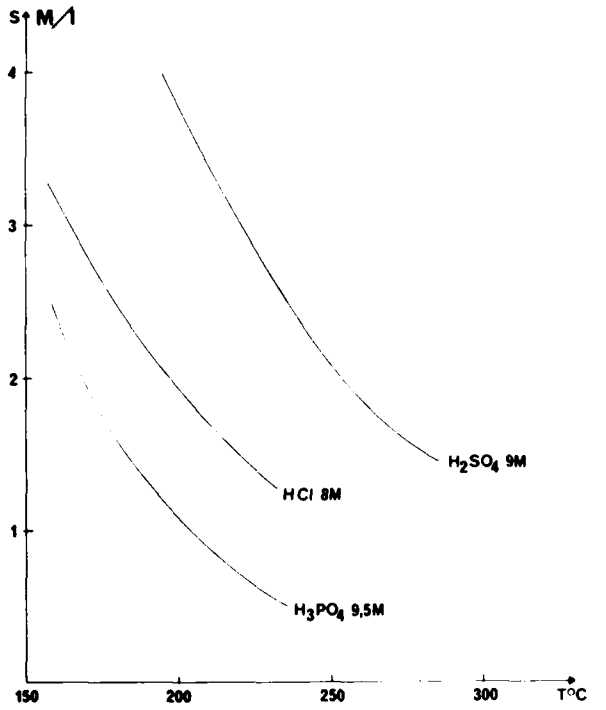



Figure 2 : Compared solubility of berlinite in $H_2SO_4$, HCl and $H_3PO_4$ acids

In these conditions, we have made some crystals from seeds with varied orientations (chiefly X, Z and Y). The crystals obtained were clear and of high cristalline perfection. Shape of crystals is quite different of that obtained in phosphoric acid, Figures 3a and b. Morphology difference must be explained from different growth rates.

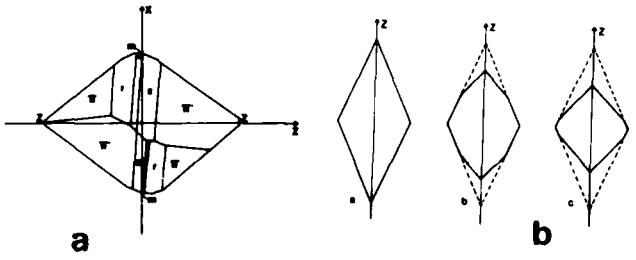



Figure 3a : Shape of berlinite crystals grown in sulfuric acid.

Figure 3b : Compared shapes of berlinite crystals grown in
a=$H_3PO_4$ b=$H_3PO_4/H_2SO_4$ c=$H_2SO_4$

In order to determine the best crystal growth conditions of berlinite in $H_2SO_4$ solvent, we have undertaken a systematic study in terms of:

- crystal growth temperature, Tc.
- temperature difference between cold and hot zone, ΔT.
- solvent concentration, 6 and 9M.

## I. CRYSTAL GROWTH EXPERIMENTS IN $H_2SO_4$ ACID

All these experiments were conducted through the reverse temperature gradient method[2,3,6], using two 375cm$^3$platinum lined autoclaves with baffle aperture 8.5%. To keep a low pressure during the growth, the filling is only 80-85% to avoid a "full" autoclave at Tc.

Though experiments are not yet quite completed, it can be emphasized some interesting results, in particular for growth rates Vx and Vz. The Figure 4 displays the Vx and Vz rates as a function of ΔT for a constant temperature, Tc=230°C. In the considered ΔT range, Vx and Vz rate variations versus ΔT are linear and their slopes can be computed: $a_{Vx}$=0.14 = 2x($a_{Vz}$=0.07). This

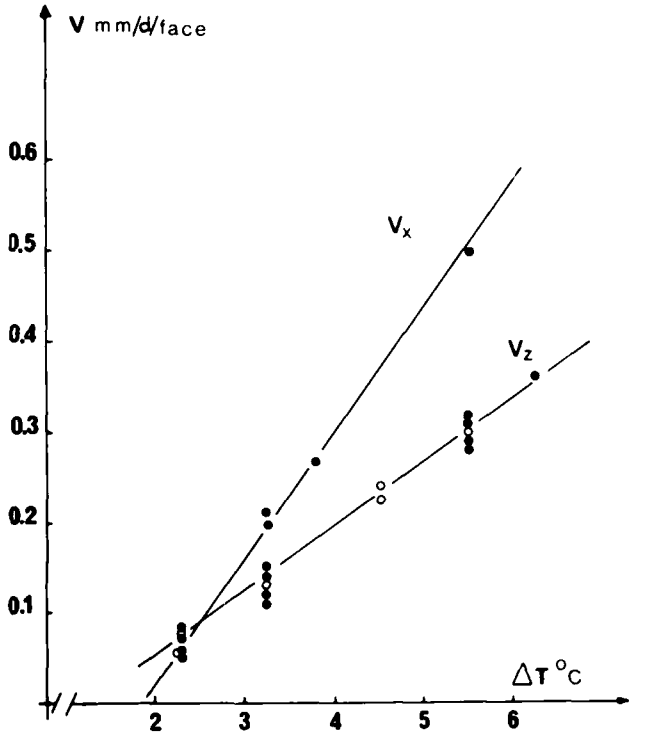



Figure 4: Growth rates, Vx and Vz, in $H_2SO_4$ 6M at 230°C
●=growth rate normal to the seed(Vx-Xseed)
o=growth rate parallel to the seed(Vx-Zseed)

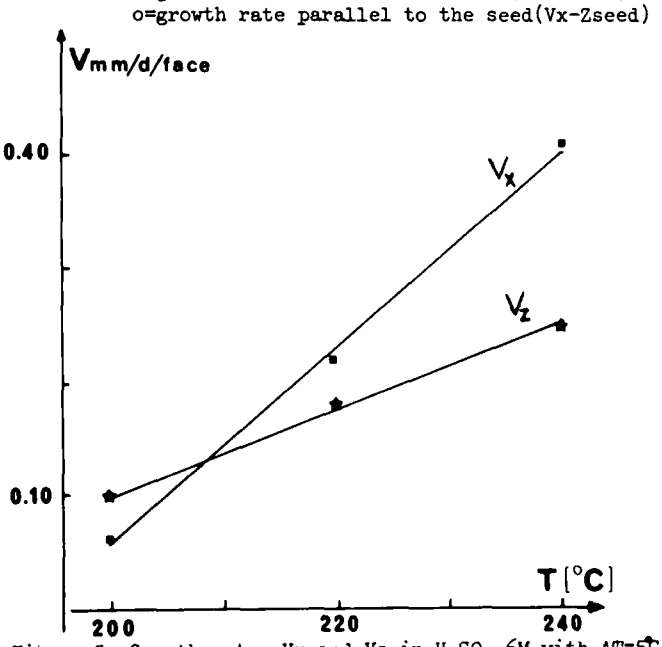



Figure 5: Growth rates Vx and Vz in $H_2SO_4$ 6M with ΔT=5°C

difference explains the existence of a point where Vx=Vz=0.08mm/d/face. The higher sensitiveness of Vx to ΔT can be seen again when considering the Vx change with respect to Tc, with the same ΔT, Figure 5.

We have found the same results in sulfuric acid, $H_2SO_4$ 9M. But in this case the rate variations are not always linear and show a maximum value for ΔT=10°C, Figure 6. In this case also, slopes can be comuted for linear parts and it can be found the same values at 223 and 241°C: $a_{Vx}$=0.05 and $a_{Vz}$=0.025 ($a_{Vx}$=2± $a_{Vz}$). It should be observed that increasing the value of Tc, from 223 to 241 C, gives a growth rate twice bigger. The same evolution can be observed if Vx and Vz are plotted versus Tc with constant ΔT=5°C, Figure 7.

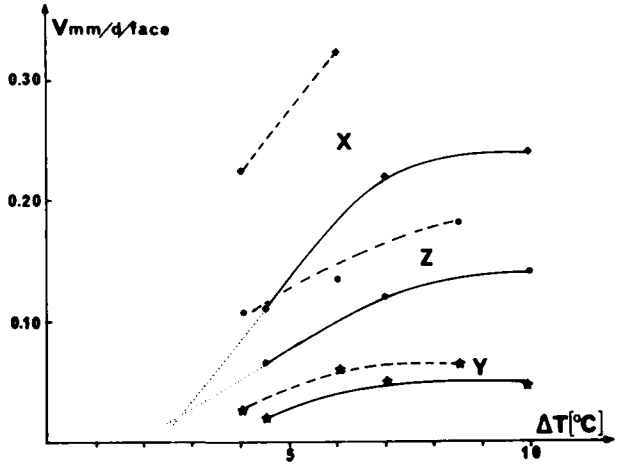



Figure 6 : Growth rates Vx, Vz and Vy in $H_2SO_4$ 9M.
Lines for Tc=223°C, dashes for Tc=241°C, dots for extrapolated Vx=Vz for Tc=223°C.

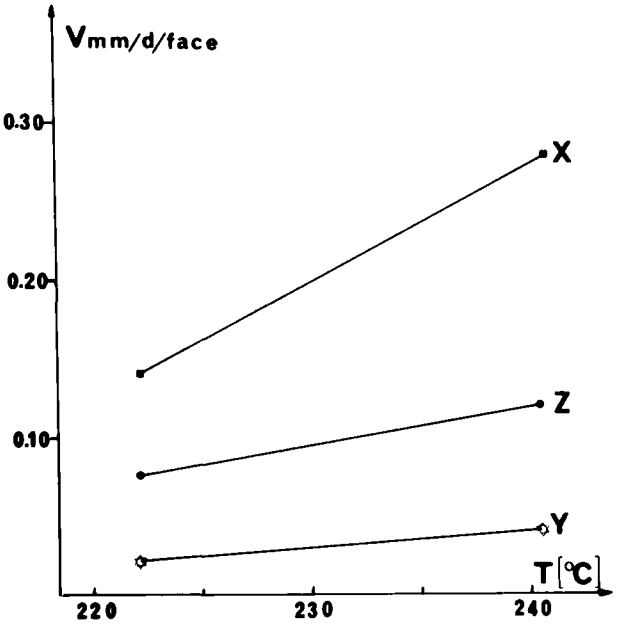



Figure 7 : Growth rates, Vx, Vy and Vz in $H_2SO_4$ 9M with constant ΔT=5°C.

What conclusions can be drawn from these experiments? They are of several kinds and show a clear difference between crystal growth of berlinite in orthophosphoric acid and sulfuric acid.

Generally speaking, the growth rates, Vx and Vz, are much more sensitive to temperature in $H_2SO_4$: growth temperature, Tc, and vertical temperature difference, ΔT. That sensitivity imposes a very accurate regulation of these parameters during crystal growth. This sensitiveness had not be encountered in $H_3PO_4$ solvent and increases with lower acid concentration:

$H_2SO_4$ 6M $a_{Vx}$ = 0.14 $a_{Vz}$ = 0.07

$H_2SO_4$ 9M $a_{Vx}$ = 0.05 $a_{Vz}$ = 0.025

but the ratio $a_{Vx}/a_{Vz}$ is constant whatever the acid concentration, $a_{Vx}/a_{Vz}$=2.

In the other hand, the Vx growth rate is generally higher than the Vz one. It is the opposite of what happens in the orthophosphoric acid where Vz is always much higher than Vx. For exemple:

$H_3PO_4$ 6.5M Tc=170°C Vx=0.12 and Vz=0.40mm/d/face

$H_2SO_4$ 6M Tc=230°C Vx=0.50 and Vz=0.30mm/d/face

for the same ΔT=5-6°C.

The growth rate anisotropy is much stronger in the $H_3PO_4$ solvent and explains the difficulty to obtain good crystal growth from Z-seeds. It can be observed liquid inclusions in the direction perpendicular to the seed plane. This anisotropy explains also different morphologies observed.

Nevertheless, in sulfuric acid, it exists a narrow temperature range where Vz>Vx and then a temperature where Vz=Vx. The corresponding value of Tc can be expressed in terms of ΔT, Figure 8. It can be observed that values are independent of acid concentration, 6 or 9M.

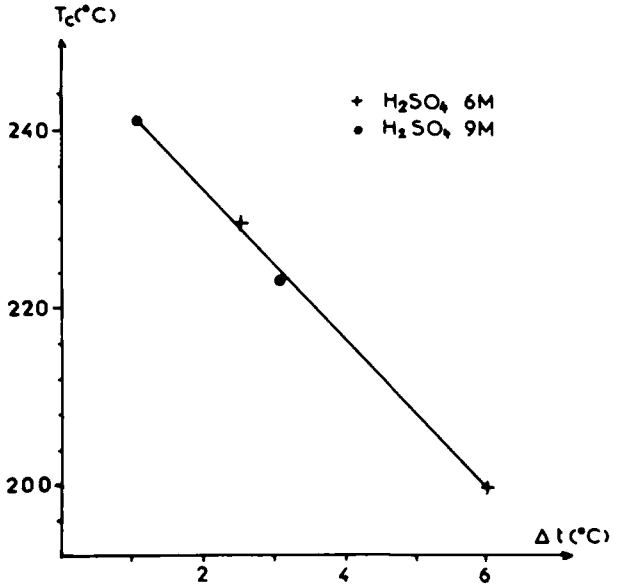



Figure 8 : Growth temperature where Vx=Vz in terms of ΔT.

Finally an increasing value of acid concentration decreases strongly the growth rate with the same ratio for Vx and Vz: $V_{9M}/V_{6M}$ = 1/3, Figure 9.

## II. INFRARED SPECTROMETRY MEASUREMENTS

As water content, $c[H_2O]$, in the crystals and their acoustic attenuation are drastically related, we have performed a systematic investigation by infrared spectrometry of water content in terms of crystal

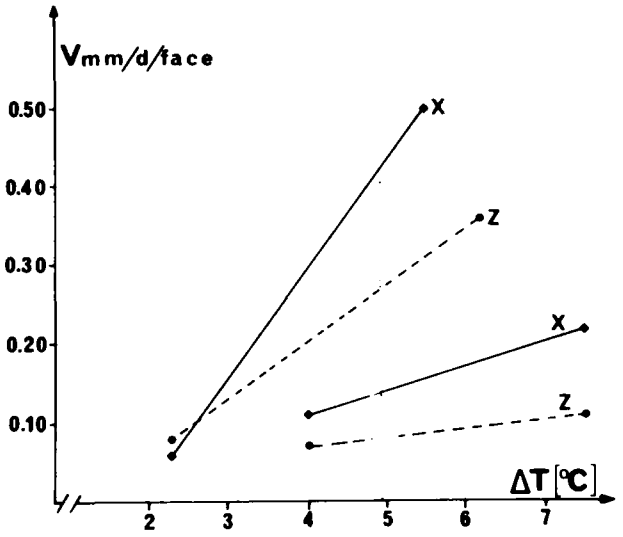



Figure 9 : Growth rates, Vx and Vz, for approximatively the same Tc≈230°C. In both cases $V_{6M} \approx 3\times V_{9M}$.

growth characteristics. Samples were slices taken on both sides of the seed. Results are certainly pessimistic because measurements have been undertaken from small crystals with no more than 2mm growth thickness on both sides.

Water concentration, expressed in ppm, has been computed from the relation proposed by Steinberg et all[11]:

$$c[H_2O] = 48.73\alpha \quad \text{with} \quad \alpha = \frac{1}{d} \log (T_{3800}/T_{3500})$$

Nevertheless, it can be assessed water content depends on: growth rate, V, growth temperature, Tc, and acid concentration. In Figure 10, we have reported some results for two acid concentrations, 6 and 9M. It is very interesting to notice that $c[H_2O]$ is lower in $H_2SO_4$ 6M than in $H_2SO_4$ 9M for the same crystal growth temperature, Tc≈240°C. One explanation is that the "water" term is probably incorrect because water molecule cannot be inserted in the $AlPO_4$ network. On the other hand, if we considere the $H_2SO_4$ dissociation:

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

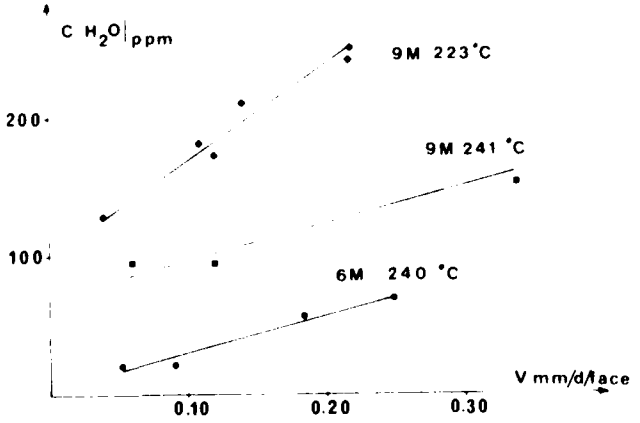



Figure 10 : Water content, $c[H_2O]$ ppm in terms of growth rates, Tc≈223 and 240°C in $H_2SO_4$ 6 and 9M.

In this case, when acid concentration is increasing $[H_3O]^+$ concentration increases also and can be inserted in the $AlPO_4$ network as an impurity.

In sulfuric acid 6M, "water" content is less than 60ppm for V=0.25mm/d/face and can be only 20ppm for V≈0.10mm/d/face, comparable to that of a good quartz, Figure 11. On the other hand, one measurement has been made directly on a crystal of 7.5mm thick and growth rate, V=0.25mm/d/face, in this case the $c[H_2O]$ is in the 40ppm range.

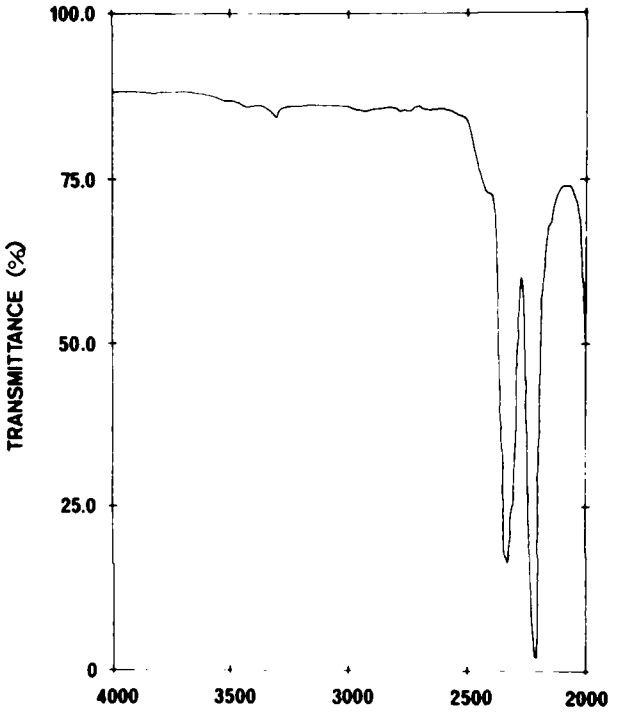



Figure 11: Infrared absorption of "dry" berlinite

## III. CHARACTERIZATION BY X-RAY TOPOGRAPHY

After "water" content, now we have to check the cristalline quality of berlinite grown in sulfuric acid solvent. In Figure 12 is presented an X-ray topograph of a Z slice of berlinite cut in a crystal grown in $H_2SO_4$ 6M, Tc 240°C, with the reverse temperature gradient technique from a Z seed. The sample was taken out of the seed and only the part concerning the X growth zone is given here. The principal features of this topograph are:

- The horizontal striations, parallel to the large side of the sample, which are due to the growth bands inthe X region of the crystal.
- The very low density of dislocations in the whole sample, less than 20.

The excellent quality of this crystal is comparable to the best samples of synthetic quartz and confirm sulfuric acid solvent as an excellent one for berlinite crystal growth.

Figure 12 : X-ray topograph of a Z-cut slice of berlinite. λ MoKα. 101 reflexion. Magn. x 4. Only the X-zone part of the slice is shown

## IV. PIEZOELECTRIC CHARACTERIZATION OF CRYSTALS

Bulk wave resonators were made, Figure 13, nearly at each step of the crystallogenesis studies. Their electrical properties were measured[13]. Among them a special attention was devoted to the frequency constants, the temperature coefficients of the resonance frequencies, RFTC, and the Q factors since we have demonstrated in a previous study[14,15] that these quantities are rather sensitive to the impurities.

### Thickness mode resonators

As in quartz, the most common impurity is "water" ($OH_3^+$ ?), this impurity limits the Q factor and has a large effect on the RFTC[16].

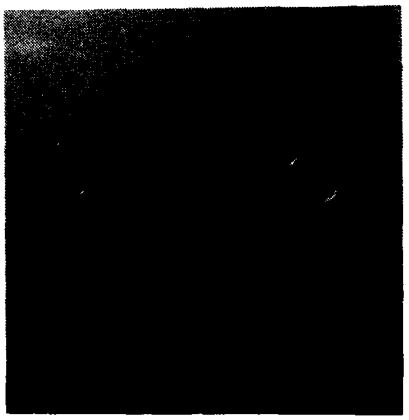


Figure 13 : Y-cut resonator in a MC6 holder.

In table 1 are displayed the results obtained with Y-cut resonators obtained from crystals grown on Y-cut seeds. As discussed in section I, the growth rate in the Y direction is very reduced in the $H_2SO_4$ solvent, so that we can expect that for these crystals the "water" concentrations are very close to the equilibrium concentration characterizing the growth thermodynamic parameters. This fact indicates the interest of such crystals and resonators to compare solvents for berlinite growth.

Table 1 : Properties of Y-cut resonators (crystals grown on Y seeds).

| CRYSTAL (solvent) | Q(5MHz) | NF (KHz.mm) | CTF1 $10^{-6}/°C^{-1}$ | CTF2 $10^{-9}/°C^{-2}$ | "water" cont. ppm |
|---|---|---|---|---|---|
| $H_3PO_4$ 180°C | 3500 | 1700 | +55 to ±65 | -60 to -180 | $\times 10^3$ |
| $H_2SO_4$ 200°C | 48400 | 1714 | 78.1 | 24.8 | 90/110 |
| $H_2SO_4$ 210°C | 53300 | 1710 | 79.0 | 45.4 | 20/45 |
| $H_2SO_4$ 220°C | 82500 | 1712 | 80.9 | 47. | 20/30 |
| $H_2SO_4$ 230°C | 51500 | 1714 | 81.3 | 47. | >40 |
| Computed properties | --- | 1688 | 63.5 | --- | --- |
| QUAPTZ Y | --- | ---- | 92. | 58. | --- |

In this table, it is assumed that the Q.f product is a constant so that Q factor can be compared at 5MHz. The resonators were actually made to have fundamental thickness shear frequencies in the range 6 to 10MHz. The Q values measured using the fundamental mode were rarely higher than 75000 at 5Miz. This value appears to be limited mostly by the design of the plane resonators used, but it can be noticed that the obtained values are much higher than those corresponding to crystals obtained in $H_3PO_4$ at low temperature. We can also observe that the resonance frequency constants ($N_{fr}$=fr.2h) display a slight sensitivity to the "water" concentration. The first & the second order RFTC are very sensitive to large "water" concentration[17]. The third order RFTC not given in table 1 remains sensitive to medium "water" concentrations. All these facts were already observed with crystals grown in $H_3PO_4$ at high temperature[14], but it can be noticed that the results obtained for Tc>250°C in $H_3PO_4$ are obtained here for Tc≈210°C, this demonstrates the interest of $H_2SO_4$.

A typical example of electrical response of a Y resonator is given in Figure 14. The thermal behaviour of resonators cut in crystal grown at 200 and 210°C is displayed on Figure 15. Since it was observed that the Q factors measured with the fundamental mode were principally limited by the type of resonators used (plane ones), we have attempted to measure a better approximation of the intrinsic Q factor of the material using higher overtone modes.

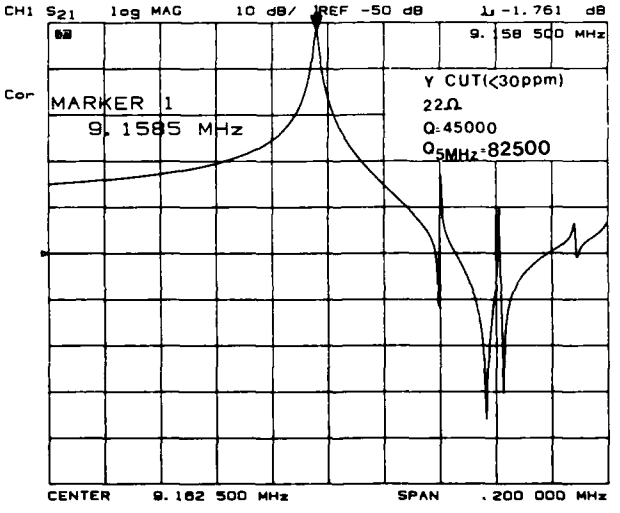



Figure 14 : Electrical response of a Y-cut resonator.

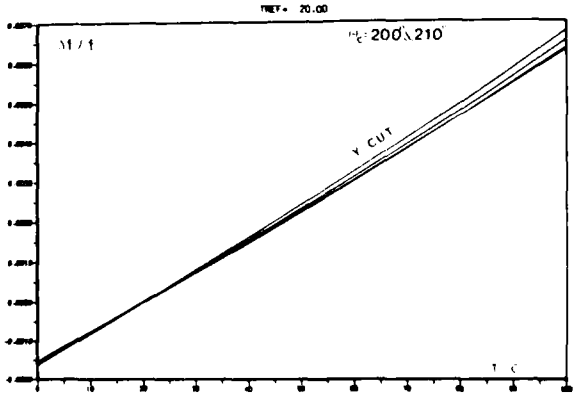



Figure 15 : Thermal behaviour of Y resonators.

In table 2 are given typical examples of Q factors (at 5MHz) measured with the 5th and 7th overtones. We can notice again the interest of the $H_2SO_4$ solvent.

Table 2 : Q factors (at 5MHz) for overtones of Y resonators cut in crystals grown on Y seeds.

| CRYSTAL | FUNDAMENTAL | OVERTONE 5 | OVERTONE 7 | "water" content |
|---|---|---|---|---|
| $H_2SO_4$ 210°C | 38200 | 840000 | 650000 | ≈20 ppm |
| $H_3PO_4$ 265°C | 65400 | 450000 | -- | ≈30 ppm SHM(250-277°C) |
| $H_3PO_4$ 240°C | 32000 | 300000 | -- | ≈100 ppm SHM(230-250°C) |

SHM = Slow heating method.

On Figure 16 the phase of $S_{21}$ for a high Q resonator operated at the fifth overtone is given. This plate was obtained in a crystal grown at 210°C. Plate of other orientations were cut in crystals grown on X,Y,Z seeds. As previously indicated, for the X and Z growth directions, the growth rates are usually much greater than for the Y direction so that larger "water" concentrations can be incorporated in crystals unless optimized growth conditions are used. In table 3 we report the results obtained for Y rotated resonators made with Y and Z seed crystals.

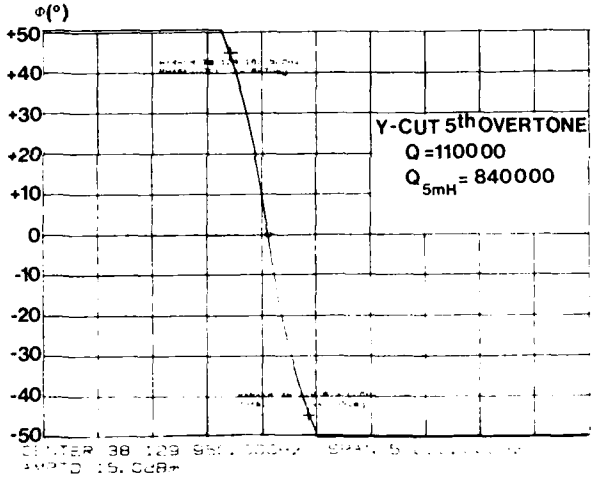



Figure 16 : Phase response for the fifth overtone of a Y resonator.

Table 3 : Properties of Y rotated resonators.

| CUT | CRYSTAL growth(°C)/seed | Q (5MHz) fund. / tal | CTF1 $10^{-6}$°C$^{-1}$ | CTF2 $10^{-9}$°C$^{-2}$ | "water" content ppm |
|---|---|---|---|---|---|
| Y+21°(r) | 220 / Y | 75000 | 61. | 21. | 20/30 |
| Y-21°(z) | 210 / Y | 50000 | 47. | 17. | ≈45 |
| Y-30° | 230 / Z | 17000 | 10. | | ≈160 |
| Y-32° | 230 / Y | 22000 | 5.4±0.1 | 2.6 | ≈150 |
| Y-38° | 230 / Y | 15000 | -21. | | ≈170 |
| Y+21° | QUARTZ | | 64. | 26. | |
| Y-21° | QUARTZ | | 58. | 49. | |

In this table we can observe that the AT cut of berlinite with low "water" concentration is situated near Y-32° whereas for "water" rich crystals it was found near Y-30° [17,18]. In table 3 as in table 1 we have given the experimental (or computed) first and second order RFTC for quartz, it can be noticed that for the corresponding cuts they are higher than the values measured for berlinite. We can also observe that, as already observed for "water" rich crystals [17] that the first and the second order RFTC vanish simultaneously near Y-32° so that the AT cut of dry berlinite has a cubical frequency temperature behaviour. From the measurements made for cuts near AT, we can expect for the third order RFTC of the AT cut of berlinite a value most probably in the range +30 to +60. $10^{-12}$ at 20°C. These values are in all cases much lower than the value observed for AT quartz (+109. $10^{-12}$).

In Figure 17 is represented the electrical response of a Y+21°20' (r-cut) resonator. The thermal behaviour of the resonance frequencies of several r-cut resonators are represented in Figure 18 where they are compared to those of Y-cut resonators.

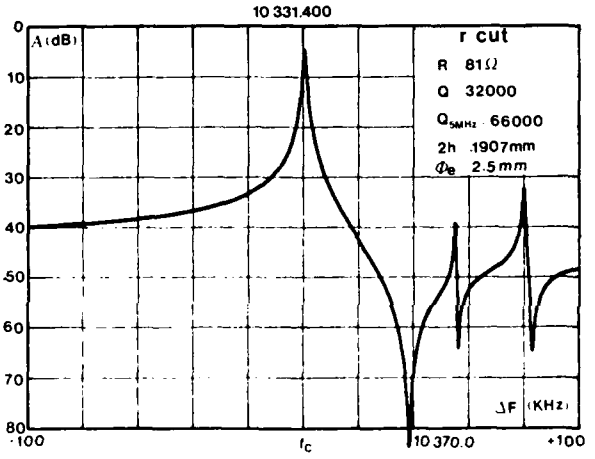



Figure 17 : Electrical response of a r-cut resonator.

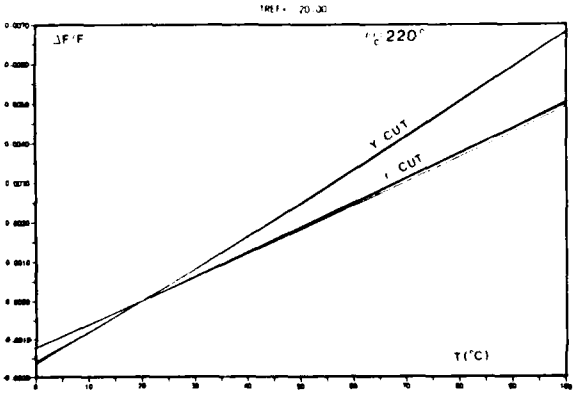



Figure 18 : Thermal behaviour of Y and r-cut.

Extensional resonators

Bars with the thickness normal to the X direction and length along Y were made. On Figure 19 the electrical response of such a resonator is given. We can notice that the Q factor is comparable to that of a quartz resonator of similar design. The measurements made on these resonators have permitted new measurements of some material constants. Their values are given in table 4.

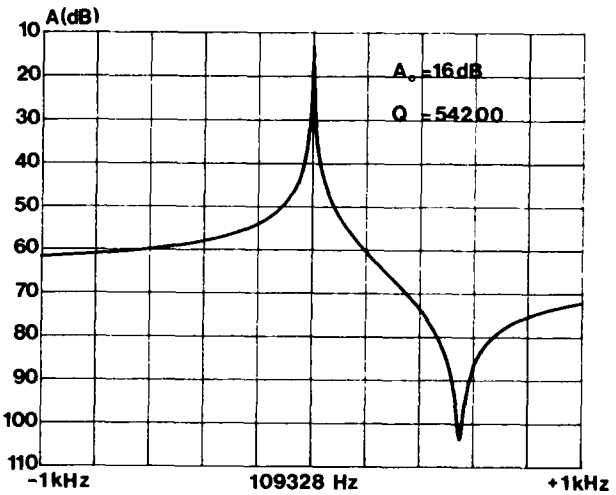



Figure 19 : Electrical response of a X+0° bar resonator.

Table 4 : Value of $S_{11}$ and $d_{11}$.

| | This study | Wang 1986 (18) | Ilraev 1986 (19) | Chang and Barach (1) |
|---|---|---|---|---|
| $S_{11}{}^{10^{-12}}_{MKSA}$ | 16.24 | 16.95 | 17.3 | 17.31 |
| $d_{11}{}^{10^{-12}}_{MKSA}$ | -3.21 | -3.30 | ---- | -5.29 |

The measured coupling coefficient is $k^2_{31}$=0.0134; this value is to be compared with an experimental value of $k^2_{31}$ = 0.00987 for a X+5° quartz resonator. The frequency temperature relation for such a resonator is given on Figure 20. In this case, the inversion temperature is 2°C and the second order T.C. at this temperatureis: -27 $10^{-9}°C^{-2}$(this value is lower than that measured for a X+5° quartz resonator of similar design). During this study it was observed that the sensitivity of the resonance frequency to the excitation level was much lower than what is observed for X+5° quartz. On Figure 21 we can observe that the resonance curve presents very little deformation even with the higher excitations. The frequency shift (one Hertz for the maximal excitation) has much probably a thermal origin.

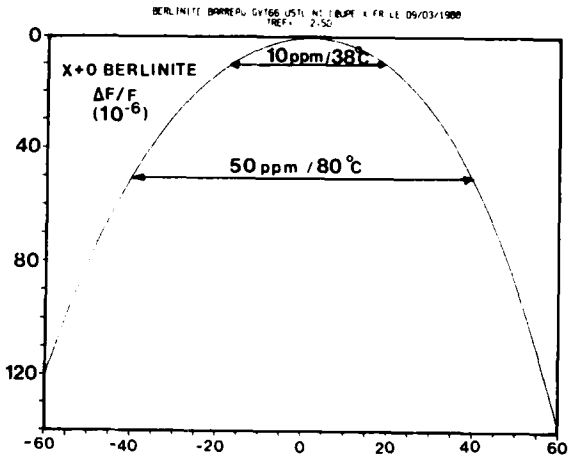



Figure 20 : Frequency temperature behaviour of a X+0° resonator.

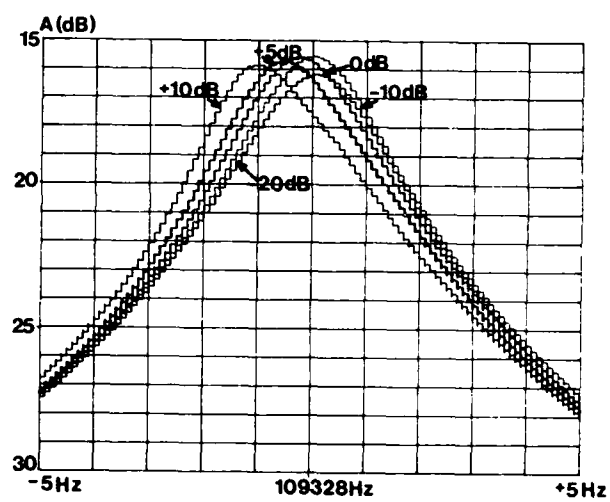



Figure 21 : Effect of the excitation level on the response of a X bar.

CONCLUSION

Sulfuric acid is a very good solvent for berlinite crystal growth. It presents the highest known solubility of berlinite without seed dissolving before thermal equilibrium of autoclaves. Under certain conditions of temperature and acid concentration, it can be possible to obtain "dry" berlinite with less than 20ppm of "water". On the other hand X-ray topography shows the excellent quality of the grown crystals, comparable to the best samples of synthetic quartz.

The electrical measurements have demonstrated that the crystals obtained in the $H_2SO_4$ solvent display high Q factors and very favourable thermal behaviour. More generally, they have confirmed the interest of this material which is characterized by a higher coupling coefficient than quartz (nearly twice for the AT cut) and better thermal behaviour. These experiments have also demonstrated that compensated cuts exist in berlinite for extensional bars.

In Figure 22 are reported two crystals grown in sulfuric acid. Now industrial crystal growth of berlinite in this solvent is investigated by SICN, Annecy, FRANCE.

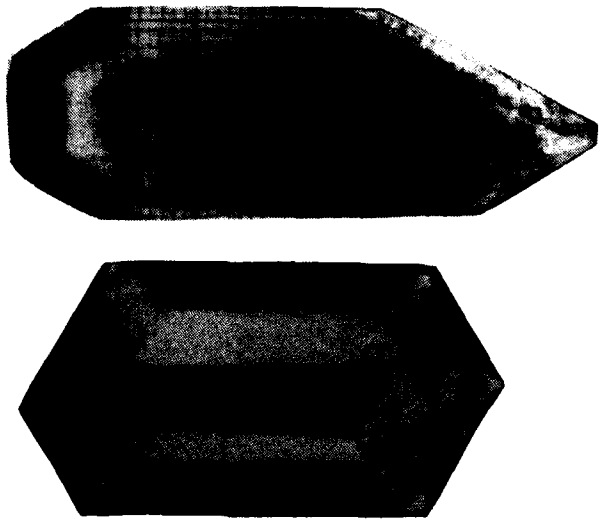


Figure 22 : Crystals grown at 230°C, $H_2SO_4$=6M, Z seeds

Aknowledgments : The authors whish to acknowledge the skilful assistance of H. Boissier, C. Cambié, D. Excoffon, C. Joly for experimental work and measurements. The authors are indebted to Professor M. Maurin, M. X. Buisson, R. Arnaud, R. Bignon and Y. Toudic for valuable indications during the course of this work.

Several parts of this work were supported by CNET contracts and CNRS.

REFERENCES

(1) Y.P. Chang, G.R. Barsch, IEEE Trans. Sonics Ultras, 23 n° 2 p 127 (1976).

(2) J.C. Jumas, A. Goiffon, B. Capelle, A. Zarka, J.C. Doukhan, J. Schwartzel, J. Détaint, E. Philippot J. Cryst. Growth vol 80 p 133 (1987).

(3) E.D. Kolb, R.A. Laudise J. Cryst. Growth vol 43 p 313 (1978).

(4) D.F. Croxall, T.R.A. Christie, B.J. Isherwood, A.G. Todd, J. Birch Europ. Conf. on Crystal Growth, Lancaster Univ. (1979).

(5) B.H.T. Chai, M.L. Shand, E. Buehler, M.L. Gilleo Proc. IEEE Ultrasonics Symp. p 577 (1979).

(6) K. Nagai, T. Ogawa, J. Okuda, J. Asamara, S. Taki Proc. First Int. Symp. Hydrothermal Reactions, p 496 (1982).

(7) J. Détaint, H. Poignant, Y. Toudic Proc. 34th Freq. Control Symp. p 93 (1980).

(8) B.H.T. Chai, E. Buehler, J.J. Flynn US Patent 4 324773 (1982).

(9) E. Philippot, J.C. Jumas, A. Goiffon, R. Astier French Patent 2559166 (1984).

(10) R.S. Naryanan, B.H.T. Chai, J.F. Vetelino, W. Soluch, R.S. Falconer, B. Chick Proc. IEEE Ultrasonics Symp. p 262 (1984).

(11) R.F. Steinberg, M.K. Roy, A.K. Estes, B.H.T. Chai, R.C. Morris Proc. IEEE Ultrasonics Symp. p 279 (1984).

(12) J.C. Doukhan, B. Boulogne, P. Cordier, E. Philippot, J.C. Jumas, Y. Toudic J. Cryst. Growth vol 84 p 167 (1987).

(13) J. Détaint, J.P. Aubry, B. Capelle, A. Zarka, J. Schwartzel, L. Michel, R. Lefèvre, E. Philippot to appear in Proc. Second Europ. Freq. and Time Forum (1988).

(14) J. Détaint, E. Philippot, B. Capelle, Y. Toudic, B. Boulogne, J.C. Jumas, A. Zarka, J.C. Doukhan Proc. Symp. Freq. Control and Synthesis p 314 (1987)

(15) J. Détaint, A. Zarka, B. Capelle, Y. Toudic, J. Schwartzel, E. Philippot, J.C. Jumas, A. Goiffon, J.C. Doukhan Proc. 40th Annual Freq. Control Symp. (1986).

(16) J.P. Hou, B.H.T. Chai Proc. IEEE Ultrasonics Symp. (1987).

(17) J. Détaint, E. Philippot, J.C. Jumas, J. Schwartzel, A. Zarka, B. Capelle, J.C. Doukhan Proc. 39th Annual Freq. Control Symp. (1985).

(18) H. Wang, B. Xu, X. Liu, J. Han, S. Shan, H. Li J. Cryst. Growth vol 79 p 227 (1986).

(19) A. Ilraev, V.E. Iopolitov, B.C. Imarov Kristallographia vol 31 n° 1 p 201 (1986).

42nd Annual Frequency Control Symposium - 1988

# A NEW MEASUREMENT OF THE BASIC ELASTIC AND DIELECTRIC CONSTANTS OF QUARTZ

B J James

GEC Hirst Research Centre, East Lane, Wembley, Middlesex HA9 7PP England

## ABSTRACT

Recent demands for high performance quartz device designs have highlighted inadequacies in previously available data sets, the most widely used of which is over 25 years old[1]. To remedy this situation a project was undertaken to measure the elastic and associated parameters of quartz to very high accuracy using a fully automated measurement system[2].

Important factors influencing the accuracy of the data obtained include

a) Indium cold weld bonding of transducers allowing a very repeatable, well characterised bond to be produced and modelled using a six-layer transmission line model.

b) Use of a comprehensive "angular spectrum of plane waves" diffraction analysis using a parabolic approximation to the slowness surface, for all modes.

c) Interferometric sample length measurement giving an accuracy of ±5ppm.

d) A new measurement of the density of quartz to an accuracy of ±50ppm.

e) Reference of all measured quantities to fundamental constants or National Standards of frequency, length, mass, temperature and resistance.

f) A new measurement of the dielectric moduli, to ± 0.2%, and their temperature coefficients over the range -60°C to +120°C was performed and the result incorporated in the determination of the elastic moduli.

Measurements were made of the properties of standard grade synthetic quartz and of Very High Purity Quartz grown by the GEC Hirst Research Centre, Wembley, England. Significant differences between the two grades of quartz were found to be fully consistent with an anelastic relaxation of the known point defects in impure quartz.

## INTRODUCTION

Historically, the majority of determinations of the elastic moduli of quartz have been based upon measurements of the frequency constant of singly and doubly rotated resonators, and most determinations of the temperature coefficients have been based upon measurement of the frequency vs. temperature behaviour of such resonators. It should be noted that in general, a rotated cut resonator experiences a change of effective orientation upon change of temperature and that this effect has not in general been allowed for in such determinations.

Data resulting from such determinations are subject to the notorious dependence of frequency and temperature coefficient of frequency of resonators upon such factors as fabrication technique, mass loading, piezoelectric shorting (plateback), resonator geometry, support structure, electrode type and size, drive level etc. Such data, although extremely useful for the calculation of characteristics of similar devices is often inadequate for accurate determination of the characteristics of different types of device at orientations far removed from those used originally.

An alternative method for elastic modulus determination is that of transit time measurement, and subsequent velocity determination, of an ultrasonic wave-packet in a parallel faced sample of material, the so called "pulse-echo" technique. This type of measurement generally requires the cementing of a piezoelectric transducer to the sample. This has traditionally been performed using an organic bond whose properties are ill defined and subject to large variations with temperature.

The aim of this project was to produce a completely general, universally consistent data set of the elastic moduli and related properties of quartz, and so for this reason, a pulse-echo type measurement was chosen in order to avoid the inaccuracies inherent in resonator type measurements. This method, and the propagation modes employed effectively eliminate the problem of variation of effective orientation with temperature found in rotated cut resonator measurements.

Special emphasis was placed upon eliminating the errors traditionally associated with pulse-echo measurements.

Quartz specimens orientated to within ±1' of the required direction were prepared with faces parallel to better than $10^{-6}$ radians. These had transducers attached via a thin film indium cold weld technique. The properties of the bond so produced are stable and completely characterizable, over the range of temperature and frequency used, by means of a comprehensive transmission line model.

An automated pulse-echo phase detection system similar to that used by Peterson, Chick and Junker[4] was used to gather an extremely large amount of transit time data which was then corrected for diffraction effects and used to accurately determine the parameters of the transducer bond. Interferometry was used to virtually eliminate length measurement error and thus a highly accurate velocity (V) measurement could be made.

The density ($\rho$) of the quartz used was measured hydrostatically so that the effective elastic moduli (C') for each mode could be determined via:

$$C' = \rho V^2$$

These effective elastic moduli are combinations of the six independent elastic moduli of quartz ($C_{ij}$), the piezoelectric moduli ($e_{ij}$) and the dielectric moduli ($\varepsilon_{ij}$).

The dielectric moduli were measured using a capacitance technique and literature values for the piezoelectric moduli[3] were used. A computer optimization routine was then developed to fit the best set of independent elastic moduli to the entire set of experimental data. To obtain the temperature coefficients the same apparatus was used to determine the variation of velocity with temperature, incorporating comprehensive corrections for the variation of all errors with

CH2588-2/88/0000-146 $1.00 C 1988 IEEE

temperature. The independent elastic moduli were calculated at each temperature and a polynomial was then fitted to the resulting data.

Absolute Transit Time Measurement

A schematic of the pulse-echo phase detection system is shown in Figure 1. An R.F. pulse is gated out of a

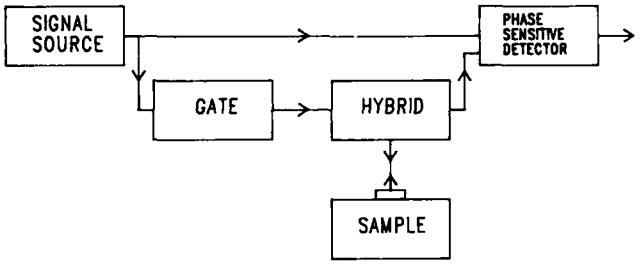



Figure 1 PULSE ECHO PHASE DETECTION

C.W. source and applied to the transducer on the specimen. A series of echoes is thus initiated which after suitable amplification is phase detected with the original C.W. As the frequency is changed, a series of quadrature frequencies ($\omega_q$) may be observed such that:

$$\omega_q \, T_{TOT} = (2n-1).\,\pi/2$$

where n= integer

$T_{TOT}$ = total time delay in system.

The group delay of the whole system is given by

$T_{TOT} = \frac{\partial\phi}{\partial\omega}$ which we may approximate by $\frac{\Delta\phi}{\Delta\omega}$ in order to find $T_{TOT}$ from our measured data[5].

A system was developed to locate and record these quadrature frequencies over a frequency range of 10-220 MHz over several overtones of the transducer, approximately 50% of the frequency spectrum being utilized. A facility was also incorporated to enable automatic switching of the detector between echoes. Temperature was controlled to 0.05°C using a fully calibrated platinum resistance thermometer. A block diagram of this system is shown in Figure 2.

The data output from this system, a series of quadrature frequencies, is manipulated to give a data file of $T_{TOT}$ vs. frequency for each of several echoes

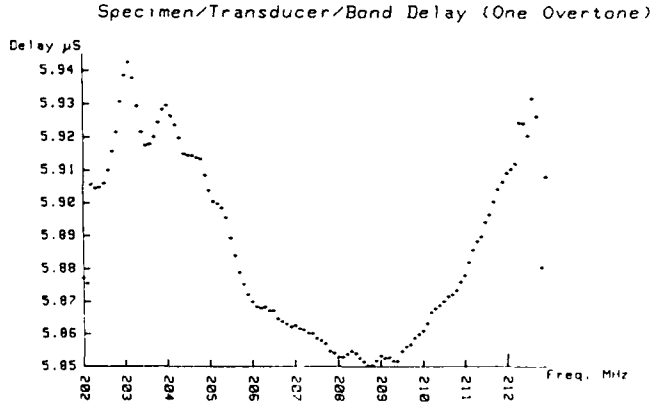



Figure 3

measured (see figure 3).

The group delay $T_{TOT}$ for echo m at frequency f is given by:

$$T_{TOT}(f,m) = mT_{RT}+(m-1)T_\Theta(f)+2T_\phi(f)+T_e(f)-T_d(f,m)$$

where $T_{RT}$ = round trip delay time in sample

$T_\Theta(f)$ = group delay due to reflection at bond

$T_\phi(f)$ = group delay due to insertion of signal at transducer

= group delay due to extraction of signal at transducer

$T_e(f)$ = group delay through all electronics

$T_d(f,m)$ = group delay due to diffraction phase advance

The diffraction delay is calculated using an angular spectrum of plane waves analysis and a parabolic approximation to the slowness surface. The $T_d(f,m)$ term is thus removed.

At each of a series of interpolation frequencies ($f_i$) a least squares straight line fit is performed on the diffraction corrected $T_{TOT}(f_i,m)$ vs. m data to give a slope of $T_{RT}+T_\Theta(f_i)$ and intercept $2T\phi(f_i)+T_e(f_i)-T_\Theta(f_i)$. The transducer reflection contribution to group delay is calculated using the transmission line model for the bonded transducer shown in Figure 4. Hence the specimen round trip delay time $T_{RT}$ may be found, see Figure 5.

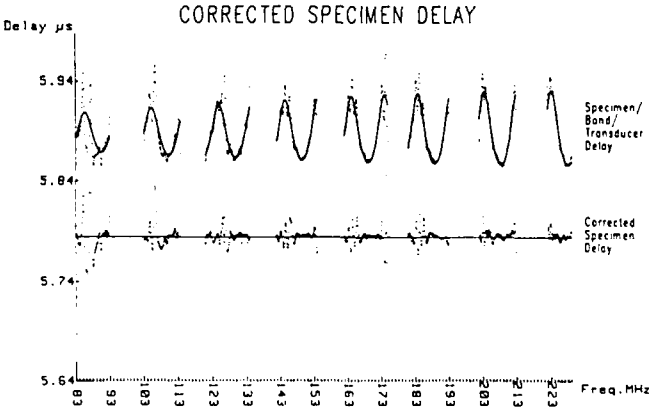



Figure 5

Temperature Variation of Transit Time

For temperature coefficient measurement, the automated pulse-echo phase detection system is phase locked to a particular quadrature point on a particular echo (dotted components in Figure 2), as the temperature of the specimen is stepped from -60°C to +120°C. The total phase ($\Phi$) remains constant, therefore

$$T(t) = \Phi/2\pi f(t)$$

where T(t) = Total phase delay at temperature t

f(t) = Phase locked frequency at temperature t

The measured data is normalised to the absolute transit time measurement and then corrected for transducer insertion and reflection phase delay as a function of frequency and temperature and electronic phase delay as a function of frequency.

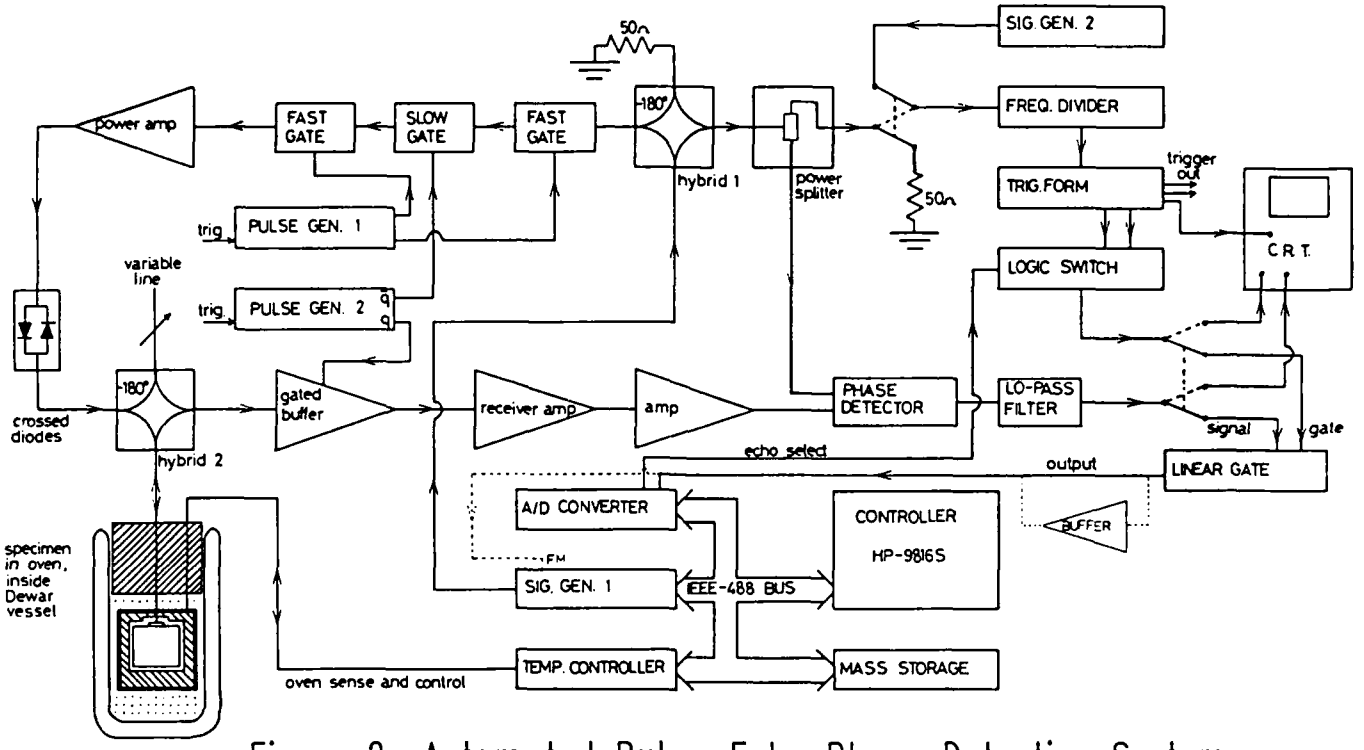



Figure 2 Automated Pulse-Echo Phase Detection System

Calculation of the Elastic Moduli

Interferometric length measurements and hydrostatic density measurements, suitably transformed using the expansion coefficients, were then used to calculate velocity for each mode j, and hence an effective elastic modulus for that mode via:

$$C_{exp_j} = \rho V_j^2 \quad \text{at } t_{ref}$$

This was done for each of the modes in Table 1. A "best" set of actual independent elastic moduli Cpq was then found by minimising the variance of the data set with respect to Cpq, where:

$$\text{Variance} = \sigma^2 = \frac{1}{n} \cdot \sum_{j=1}^{n} \frac{1}{(\Delta C_{exp_j})^2} \cdot (C_{exp_j} - C_{CALCj})^2$$

where $C_{CALC_j}$ are the effective elastic moduli in Table 1 calculated from the Cpq, the dielectric moduli measured in this analysis, $\varepsilon^s_{jk}$, and the piezoelectric moduli which are the result of a critical analysis of data available in the literature[3]. The piezoelectric and dielectric constants are transformed to their values at $t_{ref}$ by means of their temperature coefficients. $\Delta C_{exp_j}$ is the standard error associated with the value for mode j, and n is the number of modes measured.

All the measured data was thus used in the determination of the independent elastic moduli weighted only by the measurement uncertainty for each mode.

Table 1 Propagation Modes Used In This Analysis

| Prop. Dir. | Pol. Dir. | Effective Elastic Modulus |
|---|---|---|
| X | X | $C_{11} + e_{11}^2/\varepsilon_{11}$ |
| X | Y+ | $(C_{11} - C_{12} + 2C_{44})/4 - (((C_{11} - C_{12})/2 - C_{44})^2 + 4C_{14}^2)^{1/2}/2$ |
| X | Z+ | $(C_{11} - C_{12} + 2C_{44})/4 + (((C_{11} - C_{12})/2 - C_{44})^2 + 4C_{14}^2)^{1/2}/2$ |
| Y | X | $(C_{11} - C_{12})/2 + e_{11}^2/\varepsilon_{11}$ |
| Y | Y | $(C_{44} + C_{11})/2 + ((C_{44} - C_{11})^2 + 4C_{14}^2)/2$ |
| Z | X | $C_{44}$ |
| Z | Y | $C_{44}$ |
| Z | Z | $C_{33}$ |
| Y+45° | X | $(C_{44} + (C_{11} - C_{12})/2 + 2C_{14})/2 + (e_{11} + e_{14})^2/2(\varepsilon_{11} + \varepsilon_{33})$ |
| Y+45° | Y+45° | $(C_{11} + C_{33} + 2C_{44} - 2C_{14})/4$ $+ ((C_{11} - C_{33} - 2C_{14})^2/4 + (C_{44} + C_{13} - C_{14})^2/2)^{1/2}/2$ |

Temperature Coefficients of the Elastic Moduli

The parameters of the transducer and bond were calculated as a function of temperature and frequency and the electronic time delay was measured as a function of frequency. These factors were used to correct the temperature variation of transit time data to obtain true specimen transit time as a function of temperature. Hence it was possible to find $C_{exp_j}$ as a function of temperature. These values were then normalized to the absolute values. At each experimental temperature, the "best" set of independent elastic moduli was found via the same optimization method used to find the absolute moduli. Finally, a polynomial was fitted to the data for each independent elastic modulus vs. temperature.

Measurement of the Dielectric Moduli and their Temperature Coefficients

History

The only reliably documented determination of the dielectric moduli of quartz is given by Bottom[6], who determined the values:

$\varepsilon^T_{11} = 39.93 \times 10^{-12}$ Fm$^{-1}$ and $\varepsilon^T_{33} = 40.73 \times 10^{-12}$ Fm$^{-1}$ both $\pm 0.09 \times 10^{-12}$ Fm$^{-1}$. Bechmann, Ballato and Lukaszek quote values of $\varepsilon^T_{11} = 39.97 \times 10^{-12}$ Fm$^{-1}$ and $\varepsilon^T_{33} = 41.03 \times 10^{-12}$ Fm$^{-1}$ (where $\varepsilon^T_{11}$ = dielectric modulus at constant stress).

The only temperature coefficients for these values are given by Landolt-Börnstein[7] who give linear temperature coefficients of +28 p.p.m. °C$^{-1}$ and +39 p.p.m. °C$^{-1}$ for $\varepsilon_{11}$ and $\varepsilon_{33}$ respectively for the range 25°C to 100°C Whether this is for constant stress or for constant strain is not stated. No values exist for temperature coefficients below room temperature.

## Absolute Measurement

Parallel plate capacitors were formed from quartz plates nominally 20 mm x 20mm x 0.5 mm thick whose plate normals were orientated to within ±2' of the c-axis ( z-axis) and a-axis (x-axis) respectively. The plates were made parallel to better than 10" of arc. One side was completely metallized and the other was plated with a series of varying diameter electrodes.

The value of the capacitor thus formed ($C_{TOTAL}$) is given by:

$$C_{TOTAL} = \pi r^2 \varepsilon_o . \varepsilon^T_r / t + 2\pi\alpha r$$

r = electrode radius
t = sample thickness
$\varepsilon^T_r$ = relative dielectric modulus at constant stress
$\varepsilon_o$ = permittivity of free space = $8.85416 \times 10^{-12}$ Fm$^{-1}$
$\alpha$ = constant of proportionality for fringing field = $8 \times 10^{-4}$ F/m (measured in this analysis).

A plot of $C_{TOTAL}/\pi r^2$ vs. 1/r therefore gives a straight line with slope $2\alpha$ and y-intercept $\varepsilon_o \varepsilon_r / t$.

The capacitance of several samples was measured for each of the two axes. Four different electrode diameters were used for each sample.

Capacitance was measured at several frequencies using a Hewlett Packard 4275A Multi-frequency LCR meter using a four wire terminal pair configuration and software offset compensation. Temperature of the sample was measured using a platinum resistance temperature sensing element and controlled to a stability of ±.05°C. Dielectric modulus was calculated for each sample, and an error weighted average over all samples was then formed.

## Temperature Coefficients of the Dielectric Moduli

In order to measure the variation of capacitance with temperature, an evacuable brass vessel was constructed, with appropriate shielding, to house the sample. This was used to eliminate surface effects due to condensation. With a sample installed, the vessel was evacuated, flushed and back filled with dry nitrogen. Measurements on the empty vessel showed negligible variation of stray capacitance with temperature.

Measurements were made on several samples from -60°C to +120°C. The capacitance vs. temperature data was then corrected for variation of sample thickness, electrode diameter and fringing field with temperature. The data was normalized to the measured absolute value of the dielectric moduli and a polynomial was fitted to the resultant curve.

Finally a calculation was made of the value of $\varepsilon^S_{11}(t)$ (the dielectric modulus at constant strain) using the dielectric data measured in this analysis and the piezoelectric data due to Cook and Weissler[3].

## Results

The moduli, their temperature coefficients and the density measured in this analysis were combined with the piezoelectric moduli and their temperature coefficients and the thermal expansion coefficients[8] due to other workers, to produce a data set of the physical properties of quartz.

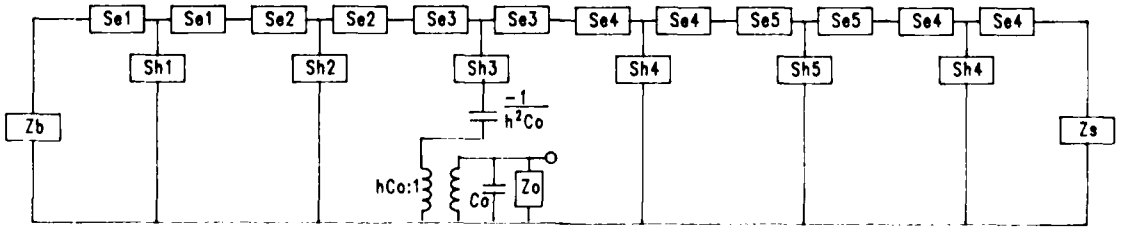
The accuracy of the data set thus assembled was tested by using it to determine the frequency vs. temperature behaviour of various well documented singly and doubly rotated bulk wave quartz resonators, including the AT cut; this is a very sensitive test, as the elastic moduli are based entirely on pulse echo measurements, without the use of any resonator data.

The results of the f vs. t predictions for various rotated-cut resonators showed good agreement with published experimental values for values of rotation angle for inflection, polynomial coefficients and rate of change of polynomial coefficients with rotation angle, in short, good agreement was found for all functional parameters except inflection temperature, which was found to be approximately 13°C low for the AT-cut ($\phi = 0$, $\Theta = -35°\ 12'$) and approximately 8°C high for the SC-cut ($\phi = 22.0$, $\Theta = -34°\ 02'$).

The value of the inflection temperature is considerably more sensitive to the absolute values of the elastic moduli than all of the other resonator parameters. Because of this dependence it was possible to vary the

Figure 4 TRANSMISSION LINE MODEL OF BONDED TRANSDUCER

$$Se_i = jZ_i \tan\theta_i \qquad Sh_i = \frac{-jZ_i}{\sin 2\theta_i} \qquad \theta_i = \frac{\pi t_i}{\lambda_i}$$



## Table 2 The New Data Set for Quartz

| Parameter (f) | Units | f (25°C) | Δf (uncertainty) | $T_f^{(1)}$ | $T_f^{(2)}$ | $T_f^{(3)}$ |
|---|---|---|---|---|---|---|
| Elastic Moduli | $Nm^{-2}$ | | | | | |
| $C_{11}$ | | $8.6790X10^{10}$ | $1.3X10^{7}$ | $-4.3591X10^{-5}$ | $-1.1098X10^{-7}$ | $-9.9879X10^{-11}$ |
| $C_{12}$ | | $6.7901X10^{9}$ | $3.7X10^{7}$ | $-2.6401X10^{-3}$ | $-3.2587X10^{-6}$ | $-2.4878X10^{-9}$ |
| $C_{13}$ | | $1.2009X10^{10}$ | $1.8X10^{8}$ | $-5.9227X10^{-4}$ | $-1.2185X10^{-6}$ | $6.8606X10^{-10}$ |
| $C_{14}$ | | $1.8116X10^{10}$ | $6.3X10^{7}$ | $1.0259X10^{-4}$ | $-3.0802X10^{-8}$ | $-4.5928X10^{-10}$ |
| $C_{33}$ | | $1.0579X10^{11}$ | $2.1X10^{7}$ | $-1.9032X10^{-4}$ | $-1.6531X10^{-7}$ | $-1.9089X10^{-11}$ |
| $C_{44}$ | | $5.8212X10^{10}$ | $1.8X10^{7}$ | $-1.7159X10^{-4}$ | $-2.5449X10^{-7}$ | $-2.4495X10^{-10}$ |
| $C_{66}$ | | $4.0000X10^{10}$ | $1.7X10^{7}$ | $1.7679X10^{-4}$ | $1.5618X10^{-7}$ | $1.0280X10^{-10}$ |
| Dielectric Moduli | $Fm^{-1}$ | | | | | |
| $\varepsilon_{11}^T \varepsilon_o$ | | $3.992X10^{-11}$ | $1X10^{-13}$ | $7.06X10^{-6}$ | $1.632X10^{-8}$ | $-2.206X10^{-10}$ |
| $\varepsilon_{33}^T \varepsilon_o$ | | $4.104X10^{-11}$ | $1X10^{-13}$ | $1.88X10^{-5}$ | $3.29X10^{-8}$ | $-1.92X10^{-10}$ |
| $\varepsilon_{11}^S \varepsilon_o$ | | $3.916X10^{-11}$ | —— | $1.47X10^{-5}$ | $1.45X10^{-8}$ | $-2.10X10^{-11}$ |
| Piezoelectric Moduli | $Cm^{-2}$ | | | | | |
| $e_{11}$ | | 0.1711 | $9.4x10^{-4}$ | $-1.6x10^{-4}$ | 0 | 0 |
| $e_{14}$ | | 0.0406 | —— | $-1.44x10^{-3}$ | 0 | 0 |
| Expansion Coeff's | | | | | | |
| $\alpha_{11}$ | | $l_x$ (25°C) | | $1.377X10^{-5}$ | $1.303X10^{-8}$ | $-6.329X10^{-12}$ |
| $\alpha_{33}$ | | $l_z$ (25°C) | | $7.483X10^{-6}$ | $9.405X10^{-9}$ | $-5.440X10^{-12}$ |

$$\varepsilon_o = 8.85416X10^{-12}\ Fm^{-1}$$

Density $\rho$ (25° C) = $2648.38 \pm 0.03\ kgm^{-3}$

$$T_f^{(n)} = \frac{1}{n!}\frac{df^n}{f(25)\ dT^n} \qquad f(T) = f(25)\left[1+\sum_n T_f^{(n)}(T-25)^n\right]$$

absolute values of the measured elastic moduli within their experimental error bounds to produce a new data set optimized to the calculation of AT and SC cut resonators. It should be noted that this optimization process resulted in a shift of the values of the effective elastic moduli of only a few parts in $10^5$.

These results support the hypothesis that the original effective moduli values (the measured values) and the f vs. t data are representations of the sample physical system, i.e. there is no significant difference between the actual elastic moduli values implied by the f vs. t curves and the measured effective elastic moduli. The fact that such a small change in the elastic moduli can cause such a large change in the inflection temperature (~13°C before optimisation, ~19°C after optimisation) indicates the great sensitivity of the f vs. t curves to elastic moduli values.

The final data set including the piezoelectric moduli and expansion coefficients due to other workers is shown in Table 2. The temperature coefficients in Table 2 are defined as $T_f^{(n)} = \frac{1}{n!}\frac{d^n f}{f dT^n}$ so that the expansion of f(T) is then given by

$$f(T) = f(To).(1 + \sum T_f^{(n)}(T-To)^n)$$

Comparison with other Data Sets

The values for elastic moduli determined in this investigation do not differ drastically from those determined by other workers as may be seen from Table 3 (in the column labelled RSQ).

Table 3 $C_{ij}(25°C)\ 10^9 Nm^{-2}$

| ij | KAY | Bech. | McSk | S-Yak | RSQ | VHPQ |
|---|---|---|---|---|---|---|
| 11 | 86.83 | 86.72 | 86.80 | 86.81 | 86.790 | 86.822 |
| 12 | 7.09 | 6.88 | 7.04 | 7.05 | 6.790 | 7.182 |
| 13 | 11.93 | 11.88 | 11.91 | 12.0 | 12.01 | —— |
| 14 | -18.06 | -17.92 | -18.04 | -18.05 | 18.116 | 19.008 |
| 33 | 105.94 | 107.1 | 105.75 | 105.762 | 105.787 | 105.720 |
| 44 | 58.26 | 57.89 | 58.20 | 58.18 | 58.212 | 58.286 |
| 66 | 39.87 | 39.92 | 39.88 | 39.88 | 40.00 | 39.82 |

The change in sign for $C_{14}$ for this determination is due to the use of the 1978 IEEE convention[9] on piezoelectricity, all previous analyses use the 1949 convention[10]. Similarly, there is general agreement between the values of first order temperature coefficients determined by various workers. There is understandably a lower agreement between the values for second order temperature coefficients, and even less

Table 4

$T_1(C_{ij})\ 10^{-6}\ °C^{-1}$

| ij | KAY | BBL | AEKW | S-Mil | S-Yak | RSQ | VHPQ |
|---|---|---|---|---|---|---|---|
| 11 | -44.3 | -48.5 | -49.6 | -50.8 | -44.1 | -43.6 | -43.7 |
| 12 | -2567 | -2678 | -2525 | -2468 | -2651 | -2640 | -2477 |
| 13 | -492 | -550 | -651 | -568 | -500 | -592 | —— |
| 14 | 98 | 101 | 89 | 100 | 108 | 103 | 90 |
| 33 | -188 | -160 | -192 | -159 | -194 | -190 | -191 |
| 44 | -172 | -177 | -172 | -169 | -174 | -172 | -170 |
| 66 | 180 | 178 | 167 | 175 | 186 | 177 | 176 |

$T_2(C_{ij})\ 10^{-9}\ °C^{-2}$

| ij | KAY | BBL | AEKW | S-Mil | S-Yak | RSQ | VHPQ |
|---|---|---|---|---|---|---|---|
| 11 | -407 | -107 | -107 | -184 | -104 | -111 | -109 |
| 12 | -7245 | -2679 | -3204 | -4095 | -2369 | -3259 | -3070 |
| 13 | -596 | -1150 | -1021 | -1350 | 30 | -1219 | —— |
| 14 | -13 | -48 | -19 | 32 | -80 | -31 | 25 |
| 33 | -1412 | -275 | -162 | -156 | -106 | -165 | -169 |
| 44 | -225 | -216 | -261 | 264 | -209 | -254 | -257 |
| 66 | 201 | 118 | 164 | 204 | 96 | 156 | 159 |

$T_3(C_{ij})\ 10^{-12}\ °C^{-3}$

| ij | KAY | BBL | AEKW | S-Mil | S-Yak | RSQ | VHPQ |
|---|---|---|---|---|---|---|---|
| 11 | -371 | -70 | -74 | -320 | -17 | -100 | -91 |
| 12 | 4195 | -1126 | -1251 | -5111 | 9416 | -2488 | -11250 |
| 13 | -5539 | -750 | -240 | -2884 | 20000 | 686 | —— |
| 14 | -825 | -590 | -521 | -468 | -1400 | -459 | -1923 |
| 33 | -243 | -250 | 67 | -266 | 650 | -19 | -36 |
| 44 | -190 | -218 | -194 | -242 | 390 | 245 | -200 |
| 66 | -777 | 21 | 29 | 127 | -850 | 103 | 916 |

agreement for third order coefficients, due to the increasing sensitivity of the fitted polynomial coefficients to experimental results, as the order increases. A comparison with the published values of the temperature coefficients of the elastic moduli may be seen in Table 4 (in the column labelled RSQ).

The workers referred to in Tables 3 and 4 are as follows:

AEKW ; Adams, Enslow, Kusters and Ward 1970[11]
Bech. ; Bechmann 1958[12]
Bech. 1956 ; Bechmann 1956[13]
BBL ; Bechmann, Ballato and Lukaszek 1962[1]
KAY ; Koga, Aruga and Yoshinaka 1958[14]
McSk ; McSkimin 1962[15]
S-Mil ; Smagin and Milshtein 1975[16]
S-Yak ; Shevel'ko and Yakovlev 1977[17]

It should be noted that all other workers consider the temperature coefficients of the dielectric moduli to be zero, and all but Shevel'ko and Yakovlev and Koga, Aruga and Yoshinaka consider the temperature coefficients of the piezoelectric moduli to be zero. These latter groups of workers used their experimental results (variation of delay-line delay and resonator frequency respectively, with temperature) to calculate the temperature coefficients of the piezoelectric moduli. The effects of the omitted temperature coefficients are thus included in the temperature coefficients of the elastic moduli making the data sets universally inconsistent. The results of this new research comprise the only universally consistent data set available.

Results of Calculation using the new Data Set

The results of the prediction of AT resonator characteristics using the new data set can be seen in Table 5 under 'This Research'. The results of the calculation of AT characteristics using other workers' data sets are also shown. Characteristics shown are the resonant frequency $f_o$, the polynomial coefficients a,b,c the inflection temperature $T_i$ and the angle dependence of the coefficients. These results are taken from Fenlon, Dworsky, Lawrence and Ryback[18]. The first column, Bechmann 1956, is the average for 12 experimental determinations of AT characteristics using different grades of natural and synthetic quartz. The column labelled 'Empirical Data' is derived from empirical data on AT resonators available from the Piezoelectric Department of the GEC Hirst Research Centre. These two columns therefore provide what is probably the 'best' representation of the actual characteristics of an AT-cut resonator.

The characteristics in column 2 (BBL 1962) are in good agreement with the 'real' AT characteristics, but this is probably due to the fact that BBL considered their AT data to be 'in perfect agreement' and thus used their AT data directly to calculate a considerable part of their data set, giving a lower weighting to other rotations. The other data sets used manage to predict the AT characteristics with varying degrees of success.

It should be noted that the AT characteristics calculated in the column marked 'This Research' are, apart from the optimisation of the inflection temperature $T_{inf}$, the results of delay line transit time measurements along pure and quasi-pure mode directions in quartz, in which no information from the AT direction was included. The fact that the predicted characteristics of a highly sensitive rotated resonator, the AT cut, agree so well with actual characteristics, reflects the accuracy of the fundamental constant data set derived in this research.

Comparison of Predicted and Empirical AT Resonator Characteristics

| AT Property | Bechmann 1956 | BBL 1962 | KAY 1958 | AEKW 1970 | S-Mil 1975 | S-Yak 1977 | Empirical Data | This Research |
|---|---|---|---|---|---|---|---|---|
| fo MHz | 16.600 | 16.603 | 16.600 | 16.612 | 16.712 | 16.598 | | 16.600 |
| a $10^{-6}°C^{-1}$ | 0.00 | -.54 | 2.30 | -.45 | 1.52 | 2.66 | 0.208 | 0.044 |
| b $10^{-9}°C^{-2}$ | -1.7 | -.49 | 23.84 | -2.78 | 171.0 | 25.2 | -8.155 | -5.56 |
| c $10^{-12}°C^{-3}$ | 106.0 | 108.4 | -235.0 | 99.1 | 119.8 | 151.9 | 108.2 | 97.11 |
| Tinf °C | 28.0 | 26.5 | 58.7 | 34.4 | -450.0 | -30.0 | 26.0 | 19.08 |
| Δa/Δθ | -5.15 | -5.13 | -6.09 | -4.59 | -4.86 | -5.13 | -4.89 | -4.80 |
| Δb/Δθ | -4.5 | -4.7 | -5.9 | -6.0 | 1.4 | -4.7 | -4.67 | -5.54 |
| Δc/Δθ | -7.3 | -2.0 | 8.5 | -2.0 | -3.0 | 21.1 | -1.94 | -4.00 |
| ΔTinf/Δθ | | 14.3 | -7.2 | 20.35 | -15.62 | 18.3 | 16.8 | 19.72 |

Table 5

## Very High Purity Quartz (VHPQ)

As part of a research programme aimed at producing very high quality quartz devices with low ageing rates and increased radiation hardness, workers at the GEC Hirst Research Centre routinely produce quartz, via a hydrothermal growth process, with impurity content typically 10 times lower than the best available commercial quartz, and with dislocation densities typically $10^2$–$10^3$ times lower[19].

Table 6 Typical Quartz Defect Concentrations

| Impurity | VHPQ(ppm) | RSQ(ppm) |
|---|---|---|
| Al | .003–.03 | .8 |
| Na | .01–.5 | .01–5.0 |
| Fe | .01 | .01 |
| Dislocations | $10cm^{-2}$ | $10^3$– $10^4 cm^{-2}$ |

Typical concentration levels of some of the more important impurities for this quartz are compared with those for standard synthetic quartz in Table 6. Samples of VHPQ were made available for this research, and a total of five acoustic velocity determinations were performed for various propagation modes sufficient to determine all elastic moduli and their temperature coefficients, except for $C_{13}$.

Significant differences were found in the absolute elastic moduli of VHPQ compared to RSQ (see Table 3). A comparison of the 25°C elastic moduli values can be seen in Figure 6 in which the lowest value on the error bar pairs for each modulus has been normalized to zero. This clearly shows a significant difference in the values of the independent moduli $C_{14}$ and $C_{12}$ (and hence $C_{66}$). These differences may be explained in terms of an inelastic relaxation of the standard synthetic quartz moduli.

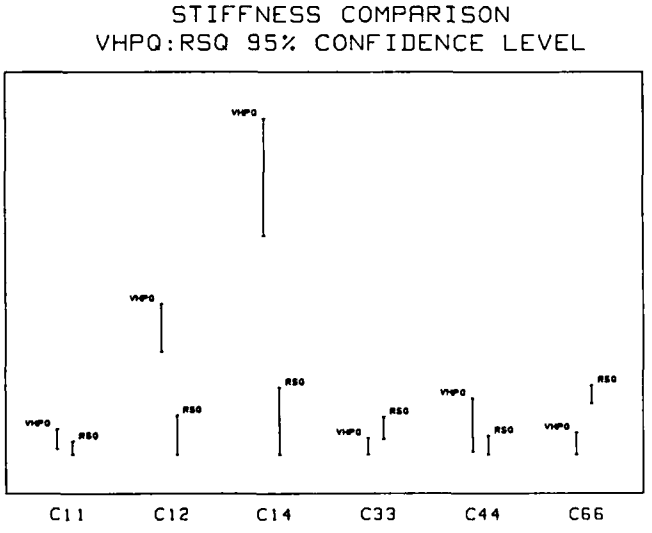



Figure 6

## Anelastic Relaxation

For a crystal containing impurities an anelastic relaxation of the elastic moduli is possible and the effective elastic modulus C' is given by:

$$C' = C_u - (C_u - C_R)/(1 + \omega^2\tau^2)$$

where $C_u$ = Unrelaxed Modulus
$C_R$ = Relaxed Modulus $< C_u$
$\Delta C = C_u - C_R$
$\omega = 2\pi$ times the applied frequency f

and $\tau$ is the relaxation time of the impurity given by $\tau = \tau_o \exp(\Delta H/kT)$

$\tau_o$ is the inverse of a fundamental jump attempt frequency, $\Delta H$ is the enthalpy of activation for the jump, usually called the activation energy and kT is Boltzmann's constant times the absolute temperature[20].

It can be seen from the above that $C' = C_u$ for $\omega\tau \gg 1$ and $C' = C_R$ for $\omega\tau \ll 1$. If several defect species are present $\Delta C = \sum \Delta C_i$ where $\Delta C_i$ is the modulus relaxation due to a particular defect. $\Delta C$ is the total relaxation between $T = 0$ and $T = \infty$ (for constant $\omega$) or between $\omega = \infty$ and $\omega = 0$ (for constant T).

The elastic modulus relaxation is accompanied by an acoustic loss peak indicating a maximum in the phase lag $\phi$ between stress and strain such that

$$\tan\phi = Q^{-1} = D\,\omega\tau/(1 + \omega^2\tau^2)$$

and D = relaxation strength.

These relationships are shown in Figure 7.

## Discussion

Many workers have studied the defect structure of quartz using acoustic and dielectric loss and electrical and thermal conductivity techniques. Consequently the relaxation times and activation energies for quartz defects have been calculated and confirmed and data for some of the more important defects are shown in Table 7.

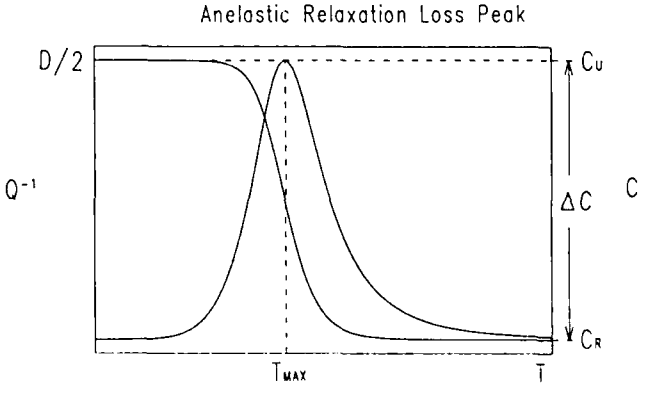



Figure 7

Table 7 Quartz Defect Characteristics

| Defect | T(K) @ 5 MHz | E(meV) | $\tau_o$(s) |
|---|---|---|---|
| $Al^{3+}$-$Na^+$ | 53 | 57 | $1.65\times10^{-13}$ |
| $Al^{3+}$-$Na^+$ | 135 | 130 | $4.44\times10^{-13}$ |
| $Al^{3+}$-hole | 23 | 8 | $8.3\times10^{-10}$ |
| $Al^{3+}$-hole | 100 | 90 | $1.0\times10^{-12}$ |
| $Al^{3+}$-hole | 135 | 110 | $2.7\times10^{-12}$ |
| Dislocations | 21 | 6.7 | $7.69\times10^{-10}$ |

Defects with a lower symmetry than quartz may cause an anelastic relaxation of the elastic moduli. Not all of the moduli are allowed to relax however. Those symmetrised moduli[20] (linear combinations of actual moduli which remain invariant under certain symmetry operations) that may possibly relax in the presence of a defect of certain symmetry have been determined by Nowick and Heller[21] and constitute the anelastic selection rules.

The only possible relaxations in quartz are for the compliances $(S_{11}-S_{12})$, $S_{44}$ and $S_{14}$. The $Al^{3+}-M^{+}$ ($M^{+} = Li^{+}$, $Na^{+}$,$K^{+}$, $H^{+}$,) defect in quartz is one of triclinic symmetry which allows all of these compliances to relax. The only other defect to allow a relaxation is one of monoclinic symmetry, which also allows all three of the above moduli to relax.

Dislocation induced relaxations are also allowed, however due to the great variety of possible dislocations and lack of exact knowledge as to their nature, no similar 'selection rules' may be determined for dislocation relaxations.

The combinations of elastic stiffnesses that are allowed to relax for a triclinic defect in quartz are given by the following relationships:

$$(S_{11}-S_{12}) = C_{44}/A$$
$$S_{44} = (C_{11}-C_{12})/A$$
$$S_{14} = -C_{14}/A$$
$$A = C44 \,.\, (C_{11}-C_{12}) - 2C_{14}^{2}$$

It may be seen from the above that, of the six independent elastic stiffnesses of quartz all but $C_{13}$ and $C_{33}$ are allowed to relax due to the presence of an $Al^{3+}-M+$ point defect.

Doherty et al[22] have measured the acoustic loss as a function of temperature for 5 MHz fifth overtone AT-cut bulk acoustic wave resonators fabricated from VHPQ. The results show a virtually flat response with no discernible acoustic loss peaks attributable to defect relaxation and a constant value for $Q^{-1}$ of approximately $5.0\times10^{-7}$ from 10K to 300K. The effective shear modulus for the AT-cut is given by

$$C'_{66} = \tfrac{1}{2}(C_{11}-C_{12})\cos^{2}\Theta + C_{44}\sin^{2}\Theta - C_{14}\sin 2\Theta$$

and $\Theta \approx -35.2°$ (1978 Convention)

and so includes all the moduli that are allowed to relax by the anelastic selection rules. The fact that no relaxation phenomena are seen for the AT resonator fabricated from VHPQ show the exceptional purity of this material, and imply that it may be very close to 'perfect' quartz in its physical properties.

If we assume the values measured for the elastic moduli of VHPQ to be those for 'pure' quartz, we may then calculate values for the elastic modulus change due to defects in RSQ (since for VHPQ, $C_U = C_R = C$).

Nowick and Berry[20] have calculated the expected relaxation in terms of the components of the strain tensor and the defect concentration. Unfortunately, insufficient data is available to calculate the magnitude of the relaxation expected between VHPQ and RSQ. We may however calculate the sign of the expected changes, and these are as follows:

$$\delta(S_{11}-S_{12}) = +ve$$
$$\delta(S_{44}) = +ve$$
$$\delta(S_{14}) = +ve \text{ or } -ve$$
$$\delta(S_{33}) = 0$$

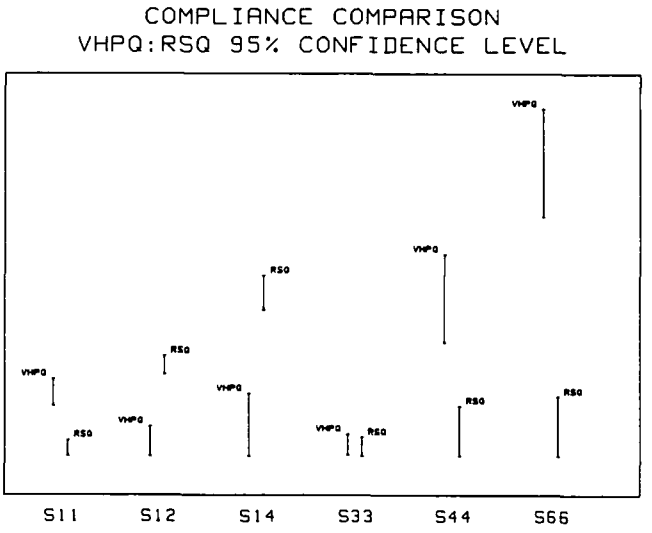



Figure 8

The values obtained experimentally are:

| | | |
|---|---|---|
| $\delta(S_{11}-S_{12})$ | $= +3.207\times10^{-13}$ | 2.20% |
| $\delta S_{44}$ | $= +3.229\times10^{-13}$ | 1.61% |
| $\delta S_{14}$ | $= -3.215\times10^{-13}$ | 7.10% |
| $\delta S_{33}$ | $= +5.035\times10^{-15}$ | 0.05% |

Which agree qualitatively with the results expected from anelastic relaxation theory. These results are depicted graphically in Figure 8 (normalised as per Figure 6). These results were calculated using the value for $C_{13}$ measured for RSQ to transform both RSQ and VHPQ stiffnesses to compliances. Figure 8 shows fairly convincingly that the differences between VHPQ and RSQ obey the anelastic selection rules for the moduli determined in this analysis.

The difference in the effective shear modulus $C'_{66}$ for the AT-cut resonator fabricated from VHPQ and from RSQ is of the order of $9.4\times10^{8}$ $Nm^{-2}$ or 3% giving a difference in frequency constant of ~1.5%. This difference is small when compared to the spread of frequency constants used for different resonator geometries, which may be of the order of 16%. If this modulus change were due to a single relaxation, the height of the $Q^{-1}$ loss peak would be of the order of $1.6\times10^{-2}$. The largest 53K $Al^{3+}-Na^{+}$ relaxation peaks seen for poor quality synthetic quartz are of the order of $10^{-3}$ making the $Q^{-1}$ peak value predicted from this analysis slightly high. However, we do not expect a single relaxation. If the total elastic modulus relaxation measured in this analysis were divided amongst, say, a 21K dislocation peak, a 53K $Al^{3+}-Na^{+}$ peak and a 135K $Al^{3+}-Na^{+}$ peak we would obtain very similar loss peak heights to those found in practice.

ACKNOWLEDGEMENTS

This work was supported by SERC and MoD and forms part of a study carried out jointly at the Hirst Research Centre and at the Royal Holloway and Bedford New College, London University.

The author would like to thank Dr. Mike Lea of RHBNC for the supervision of this project and for his encouragement, and Dr Robert Peach of GEC Hirst for his diffraction analysis and valuable discussions.

REFERENCES


1. R. Bechmann, A.D. Ballato, and T.J. Lukaszek, "Higher Order Temperature Coefficients of the Elastic Stiffnesses and Compliances of Alpha Quartz", Proc. IRE, 50, 1812-1822, 1962

2. B.J. James, Ph. D Thesis "Determination of the Elastic and Dielectric Properties of Quartz", Royal Holloway and Bedford New College, London Uni. Spring 1987

3. R.K. Cook and P.G. Weissler, "Piezoelectric Constants of Alpha and Beta-Quartz at Various Temperatures", Phys. Rev. 80(4), 712, 1950

4. G.L. Peterson, B. Chick and W. Junker, "Error Correction Methods in Gated Amplifier Absolute Velocity Measurements and Comparison with the Pulse-Echo-Overlap Technique", Ult. Symp. Proc. 650, 1975

5. E.H. Young Jr., "Discussion of Time Delay in Reference to Electrical Waves", IRE Trans. Ult. Eng. 13, July 1962

6. V.E. Bottom, "Dielectric Constants of Quartz", J. Appl. Phys. 43 (4), 1493, 1972

7. Landolt-Börnstein, New Series, Group III, Vol. 11, ed. K.H. Hellwege, Springer-Verlag, Berlin, 1979

8. R.A. Kirby, T.A. Hahn and B.D. Rothrock in "American Institute of Physics Handbook", ed. D.E. Gray, McGraw-Hill, New York, 1972

9. IEEE Standard on Piezoelectricity (1978), ANSI/IEEE Std.176, 1978

10. IEEE Standard on Piezoelectricity (1949), IRE 14 S1, 1949

11. C.A. Adams, G.M. Enslow, J.H. Kusters and R.W. Ward, "Selected Topics in Quartz Research", Proc. 24th Ann. Symp. Frequ. Contr., 55-63, 1970

12. R. Bechmann, "Elastic and Piezoelectric Constants of Alpha-Quartz", Phys. Rev. 110, 1060-1061, 1958

13. R. Bechmann, "Frequency-Temperature Angle Characteristics of AT-Type Resonators Made of Natural and Synthetic Quartz", Proc. IRE 44, 1600-1607, 1956

14. I. Koga, M. Aruga and Y. Yoshinaka, "Theory of Plane Elastic Waves in a Piezoelectric Medium and Determination of Elastic and Piezoelectric Constants of Quartz", Phys. Rev. 109, 1467-1473, 1958

15. H.J. McSkimin, "Measurement of the 25°C Zero-Field elastic moduli of quartz by high frequency Wave Propagation", J. Acoust. Soc. Amer. 34, 1271-1274, 1962

16. A.G. Smagin and B.G. Mil'shtein, "Elastic Constants of Alpha-Quartz Single Crystal", Sov. Phys. Crystallogr. 19, 514-516, 1975

17. M.M. Shevel'ko and L.A. Yakovlev, "Precision Measurement of the Elastic Characteristics of Synthetic Piezoelectric Quartz", Sov. Phys. Acoust. 23, 187-188, 1977

18. P. Fenlon, L. Dworsky, J. Lawrence and D. Ryback, "The Physical Constants of Quartz", Proc. 5th Quartz Crystal Conference, 176, 1983

19. D.F. Croxall, I.R.A. Christie, J.M. Holt, B.J. Isherwood and A.G. Todd, "Growth and Characterization of High Purity Quartz", Proc. 36th A.F.C.S., 62, 1982

20. A.S. Nowick and B. Berry, "Anelastic Relaxation in Crystalline Solids", Academic, New York, 1972

21. A.S. Nowick and W.R. Heller, "Dielectric and Anelastic Relaxation of Crystals containing Point Defects", Advan. Phys. 14, 101, 1965

22. S.P. Doherty, S.E. Morris, D.C. Andrews and D.F. Croxall, "Radiation Effects in Synthetic and High Purity Synthetic Quartz: Some Recent Infrared, Electron spin resonance and acoustic loss results", Proc. 36th A.F.C.S., 66, 1982

42nd Annual Frequency Control Symposium - 1988

# POSSIBLE MECHANISMS FOR THE INTRODUCTION OF HYDROGEN INTO ALPHA-QUARTZ DURING SWEEPING

John G. Gualtieri

US Army Electronics Technology and Devices Laboratory
(LABCOM), Fort Monmouth, NJ 07703-5000

## ABSTRACT

Previous air-sweeping experiments suggested that metal electrodes in contact with the quartz influence or catalyze the introduction of hydrogen into the crystal.[1] We proposed that hydrogen was being dissolved into the metal electrode where it became charged and was subsequently driven into the quartz Z-axis channels by the electric field.

Additional hydrogen-introduction mechanisms are suggested as a result of recent sweeping experiments. If the metal anode is porous, protons can be removed directly from atmospheric water molecules in an electrochemical reaction which results in the liberation of oxygen and the transport of electrons to the metal. Restricting the penetration of air through the metallization to the quartz has resulted in a vacuum/inert-type of sweeping.

Air-sweeping experiments have confirmed that the introduction of hydrogen: 1) is not increased by using electrode metals such as Pd, V, and Ni, which readily dissolve hydrogen, 2) requires quartz-metal contact, 3) does not correlate with Z-surface damage, and 4) is not suppressed by reduction in field strengths. By contrast, electrode-metal indiffusion is increased by Z-surface damage and high electric fields.

Fourier transform infrared spectrometry (FTIR) was used to determine the increased hydrogen content due to changes in sweeping conditions. The indiffusion of electrode metal was studied using Rutherford backscattering techniques.

## INTRODUCTION

Previously, we found a reduction in sweeping efficiency when suppression-type electrodes were used.[1] We suggested that the reduction in sweeping efficiency was probably the result of interference with the dissolution and diffusion of hydrogen in and through the metal electrode into the quartz. If H-diffusion through the metal is the principal rate-limiting mechanism, then Pd, V, and Ni metallizations should be superior sweeping electrodes. The reason is that hydrogen diffuses through these metals much more easily than through Au and Pt, which are normally used as sweeping electrodes.

Our objective in this study was to investigate hydrogen-dissolving metallizations as sweeping electrodes to increase our understanding of the possible mechanism or mechanisms of hydrogen introduction into alpha-quartz during sweeping. Our approach was to vary the sweeping conditions and monitor the increased hydrogen content of the quartz and the indiffusion of electrode metal.

## EXPERIMENTAL

Cultured quartz from three U.S. suppliers was used in this investigation. The quartz was lumbered into Y-bars after seed removal.[2] Various evaporated metallizations were applied to polished Z-surfaces of the quartz samples. Sweeping was carried out in air, forming gas (95% $N_2$ + 5% $H_2$), and ultra high-purity (UHP) $N_2$. A gas flow rate of 100 cc/min and temperature ramped sweeping were used.[3] Fourier transform infrared spectrophotometry was used to obtain room-temperature infrared absorption due to OH-vibrations, before and after sweeping. The spectrophotometer was a Perkin-Elmer model 1760. In order to estimate the increased hydrogen content due to sweeping, normalized-spectral-difference spectroscopy was applied to spectra taken before and after sweeping. A detailed description of this technique was given previously.[1] In addition, Rutherford backscattering (RBS) was used to study the indiffusion of Au into quartz during sweeping. The technique of RBS involves directing a beam of high-energy $^4He$ ions at a sample. The primary effect of the scattering is to change the energy of the $^4He$ by backscattering through the conservation of momentum. By analyzing the different energies at which backscattered particles emerge, the target atomic masses, as well as their depth, can be determined.

## EXPERIMENTAL OBSERVATIONS

### High-Field Sweeping

We reported that suppression of electrode-metal diffusion by degradation of the metal-electrode contact to the quartz, resulted in suppression of hydrogen indiffusion.[1] We now have observed that sweeping with good metal contact and high electric fields (~ 1000 V/cm) results in both hydrogen and electrode-metal indiffusion. In Fig. 1, we have plotted a series of infrared scans, taken with the

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

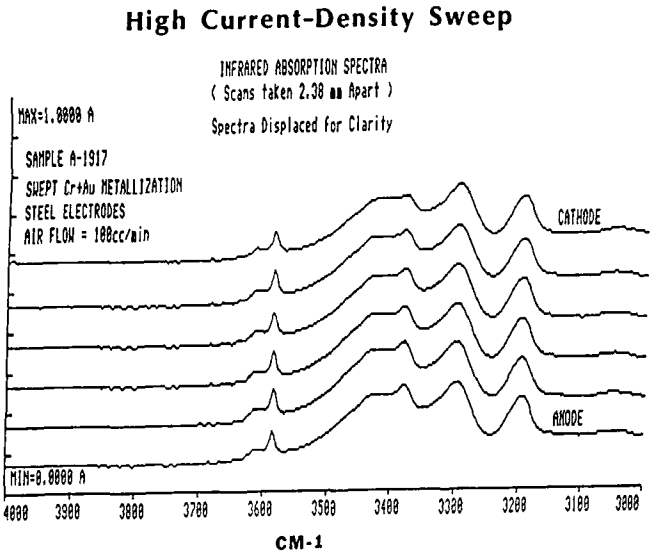



Fig. 1. A series of infrared scans, taken at room temperature. The IR beam was unpolarized and in the X-direction. The scans adjacent to the anode and cathode are labelled. The field was 1000 V/cm and the maximum current density was 4.37 $\mu A/cm^2$.

infrared beam in the X-direction, of a sample swept in air and at high field (high current densities). The scans were taken at successive positions between the anode and cathode. The band at 3378 $cm^{-1}$, which is due to aluminum in a silicon site compensated by interstitial hydrogen (Al-OH), is fairly uniform across the bar. This band is absent in unswept cultured quartz, where the Al-impurity is compensated by interstial alkali-metal ions. The band at 3585 $cm^{-1}$, indicative of as-grown OH, is also uniform across the bar. In vacuum or inert-atmosphere sweeping, this band is largely depleted in the anode region, because the crystal uses as-grown OH as an internal source of hydrogen for compensation of the aluminum impurity.[4,5] This depletion in the anode region is evident in Fig. 2. The uniformity of the 3585 $cm^{-1}$ band and the increased absorption in the 3378 $cm^{-1}$ region (Fig. 1) show that hydrogen has been introduced into the quartz from an external source.

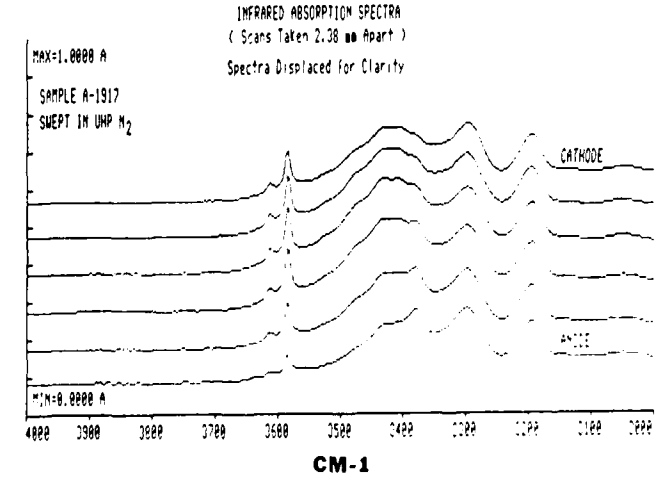



Fig. 2. Similar to Fig. 1., except that UHP N2 was used with a 100 cc/min flow rate. Note the depletion of the 3585 $cm^{-1}$ (as-grown OH) band in the anode region.

The RBS data, shown in Fig. 3, reveal that gold has also diffused into the sample during the sweeping. Also in Fig. 3, the large bands, centered at about 1.75 MeV, indicate the presence of gold at both the anode and cathode surface. If there were no indiffusion of gold, this large band would be close to rectangular in shape. The long tails extending to lower backscattered energies demonstrate that gold is deeply inside the bar under both surfaces. The small band at 1.5 MeV represents chromium at the metallized surface. The band at 1.25 MeV indicates Cr beneath the layer of gold. The (150 Å thick) Cr metallization was evaporated onto the quartz before the evaporation of the (1000 Å thick) Au metallization. It is apparent that Cr has migrated to the surface.

Low-Field Sweeping

We have found that sweeping does not proceed (no hydrogen is introduced) if only an electric field is present and if there is no current flow. At low current densities (produced by fields typically below 250 V at

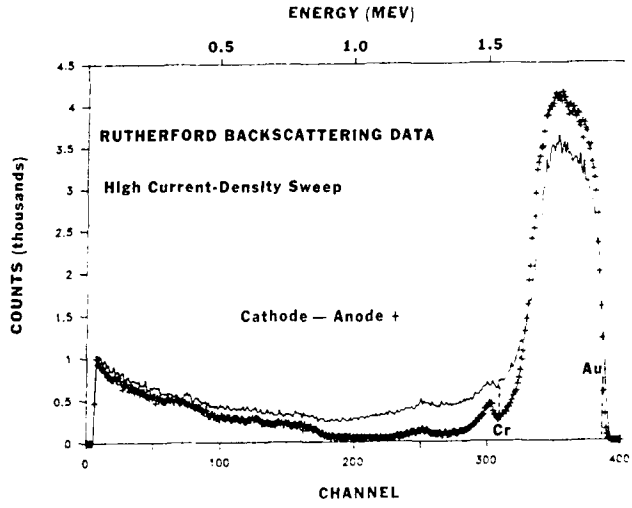



Fig. 3. RBS data for the high-current-density air-swept sample of Fig. 1. Gold is present under both surfaces (long tails extend to low backscattered energies).

500°C), we have found that hydrogen is introduced, but the electrode-metal indiffusion is suppressed.[3] To see that the the hydrogen introduction is unaffected by lowering the field, we show again, in Fig. 4, a series of infrared scans similar to Fig. 1. The band at 3378 $cm^{-1}$ (Al-OH) is represented uniformly across the bar, and the band at 3585 $cm^{-1}$ (as-grown OH) is not depleted in the anode region. The RBS data for this bar, shown in Fig. 5, reveal that little gold has diffused into the bar. Note the low-yield tail in the 2.5 MeV region.

Other Observations

To decide if the gold was migrating into the quartz as an uncharged species, we subjected Cr/Au metallized samples to three annealing cycles. In each cycle, samples were heated: in air, from room temperature to 540°C in 24 hours, kept at this temperature for 24 hours, and cooled back to room temperature in

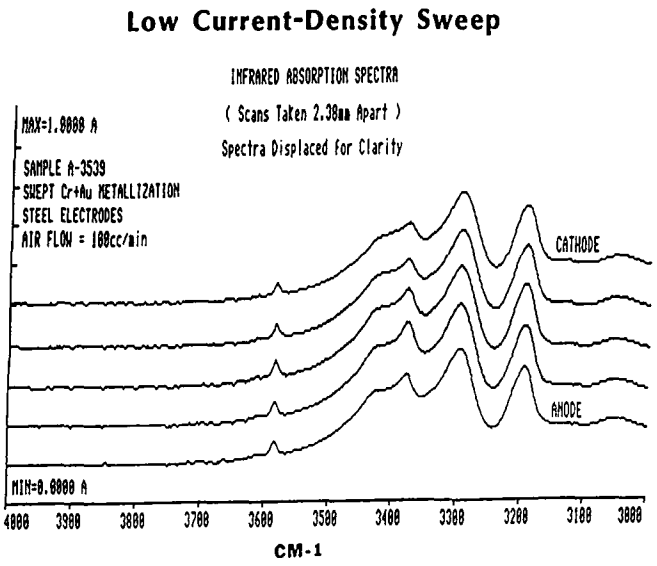



Fig. 4. A series of infrared scans similar to Fig. 1. Both 3585 $cm^{-1}$ (as-grown OH) and 3378 $cm^{-1}$ (Al-OH) bands are uniform across the bar. The field was 200 V/cm and the current density was 0.70 $\mu A/cm^2$.

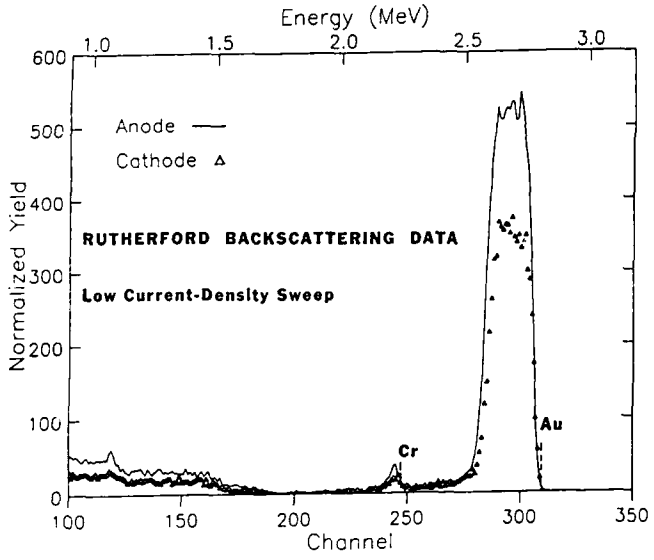



Fig. 5. RBS data for the low-field, low-current-density sample. The low-yield tail in the 2.5 MeV region indicates that very little gold has diffused into the sample.

24 hours. The RBS results are given in Fig. 6. The gold indiffusion is slight, as indicated by the low yield (small tail) in the 2.5 MeV region.

We have observed that the indiffusion of electrode metal nucleates the crystallization of inclusions.[6,7] The inclusions grow to very large size and cause fracturing of the quartz. Precipitation of inclusions and fracturing are eliminated when electrode-metal indiffusion is suppressed. Figure 7(a) is a photomicrograph of a large inclusion which was formed as a result of sweeping. In Fig. 7(b), the photomicrograph is overexposed to reveal the fainter image of the associated fracture.

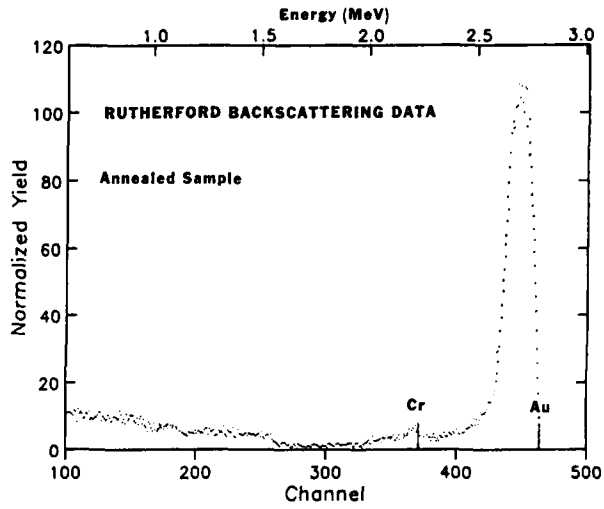



Fig. 6. RBS data for a quartz sample, with evaporated 150 Å Cr/1000 Å Au electrodes, that was annealed using three temperature-ramped annealing runs. The gold indiffusion is slight, as indicated by the low-yield tail in the 2.5 MeV region.

(A)

(B)

Fig. 7. Photomicrographs of: (a) a large inclusion formed during sweeping and (b) the same inclusion using overexposure to reveal the associated fracture.

## DISCUSSION

### Hydrogen Dissolution Mechanism

Previously, we suggested that metals in contact with quartz catalyze the introduction of hydrogen during sweeping.[1] We speculated that the mechanism involved the adsorption of molecular hydrogen on the metal surface and diffusion of some of the hydrogen into the metal, where it became positively charged. The sweeping field could then drive protons from the metal into the quartz Z-axis channels at the metal-quartz contact areas. Since we have found that no hydrogen is introduced without current flow, we assume charge is transferred through an electrochemical process.[8,9] Thus,

$$H_2(g) \underset{\text{metal surface}}{\rightleftharpoons} H_2 \text{ (adsorbed)}$$

$$H_2 \text{ (adsorbed)} \longrightarrow 2H^+ + 2e^-.$$

For this mechanism to be operative, there should be some hydrogen-introduction dependence on the ability of different metallizations to dissolve hydrogen. Also, catalytic activity is a function of pretreatment of the catalytic surface.[10] Thus, if dissolution of hydrogen by the metal surface is influencing the sweeping, there should be some dependence of increased hydrogen content upon surface roughness.

TABLE I

Metallization-type vs. Increased H+ Content

| Sample* | H-Solubility | Metallization (1000 Å) | Swept - Unswept Integrated Absorption ($cm^{-2}$) | Increased H-Content x $10^{-17}$ ($cm^{-3}$) |
|---|---|---|---|---|
| A | hs/hs | V/Pd | 48.3 | 7.00 |
| | s | Pt | 48.7 | 7.06 |
| | ss/is | Cr/Au | 57.7 | 8.37 |
| B | hs | Pd | 23.0 | 3.33 |
| | ss/is | Cr/Au | 30.6 | 4.44 |
| C | hs | Ni | 14.8 | 2.15 |
| | ss | Al | 16.9 | 2.45 |
| | ss | Cr | 13.4 | 1.94 |
| | hs/is | Ta/Au | 16.5 | 2.39 |

hs = highly soluble, s = soluble, ss = slightly soluble, is = insoluble. [11]

*A, B, and C represent samples from different suppliers. Comparisons should be made only within each sample type.

In Table I, we have listed the increased hydrogen content, obtained from infrared difference spectroscopy, using metallizations with different hydrogen-dissolution properties.[11] The sweeping experiments were carried out using quartz from three different suppliers (listed as A, B, and C). Note that there is no dependence of increased hydrogen content on ability of the metallization to dissolve hydrogen. The result of sweeping experiments, attempting to show a dependence of increased hydrogen content on the surface roughness of Cr/Au metallized surfaces, is given in Table II. The null result is in contrast to similar experiments, which showed a correlation of surface roughness and indiffusion of electrode metal during sweeping.[12] This means that reduction of surface roughness through mechanical or chemical polishing is another way to suppress electrode-metal indiffusion without affecting the hydrogen indiffusion.

Since increased hydrogen content, resulting from sweeping, does not correlate with either metallization type or surface roughness of metallization, we suspect that the hydrogen-introduction depends on another mechanism.

### Electrode Porosity

Initially, if the evaporated metal surfaces are porous, and further physical and chemical changes due to conditions of the sweeping experiment increase the porosity, then electrochemical processes involving the atmosphere, electrode, and the quartz (solid electrolyte) are possible. When metal foil electrodes are used for sweeping, there are, also, large areas of free-quartz surface accessible to the gas phase.

TABLE II

Z-Surface Finish vs. Increased H+ Content

[Sample A - Cr/Au Metallization]

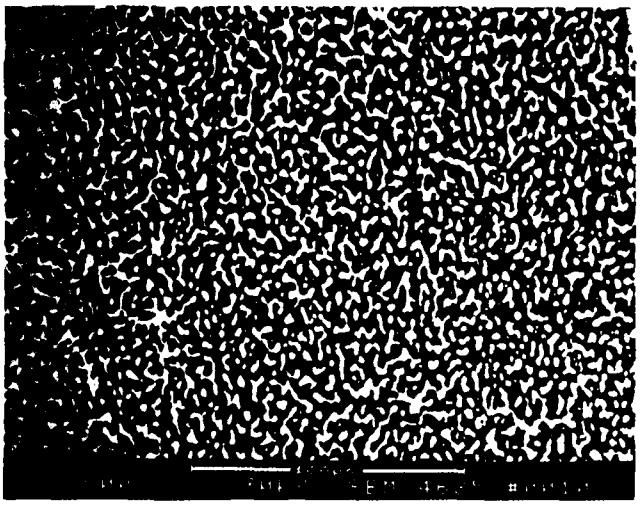
| Surface Treatment | Roughness rms Deviations (μM) | Swept - Unswept Integrated Absorption ($cm^{-2}$) | Increased H-Content x $10^{-17}$ ($cm^{-3}$) |
|---|---|---|---|
| Sawn | 1.96 | 23.9 | 3.47 |
| Lapped (12μM) | 0.49 | 27.3 | 3.96 |
| Polished (1μM) | 0.02 | 23.5 | 3.41 |
| Chemically Polished | 0.01 | 22.8 | 3.31 |

In Fig. 8, we show photomicrographs of a Cr/Au electrode surface before (a) and after (b) a temperature-ramped sweeping run (1000 V/cm and a maximum temperature of 537°C). The metallization before sweeping shows numerous holes, i.e., places where the three phases of atmosphere, electrode, and quartz can coexist. After sweeping, the porosity is extensive. An attempt was made to reduce the increased porosity, produced as a result of sweeping, by increasing the Cr-metal thickness of the Cr/Au metallization from 150 Å to 1000 Å. The increased Cr metallization produced an entirely different air-sweeping result. The sweeping resembled vacuum or inert atmosphere-type sweeping. Compare Fig. 9 to Fig. 2, especially the depletion of the 3585 $cm^{-1}$ (as-grown OH) band in the anode region. This depletion indicates that the atmosphere was not penetrating the thicker Cr metallization. A photomicrograph of the bar with the increased Cr-metal thickness after sweeping is shown in Fig. 10. Intense color-center striae in the Z-direction are evident. The small arrow on the right points to the end of the Cr/Au metallization. Note that the color-center boundary also ends there. Porous graphite electrodes were found to be most effective in preventing the formation of color centers in electrolytic coloration of alkali halides.[13]

We found that porous electrodes are used extensively with electrochemical reactions in which one of the reagents is a gas. Because transport processes occurring over long distances are slow, such reactions proceed efficiently in the neighborhood of all three phases or triple points where the gas phase,

$H_2O$ or $H_2$, the metal, and the electrolyte all come together. It has been shown that the current density is increased substantially at those points of contact between the platinum electrode and solid electrolyte (stabilized zirconia) from which generated gas can escape

(A)



(B)

Fig. 8. Photomicrograph of a Cr/Au electrode surface (a) before and (b) after a temperature-ramped sweeping run. Before sweeping, the metallization contains many holes. After sweeping, extensive porosity develops into island formations.

more easily. Lines were scratched in both cathode and anode to increase the boundary of the triple junction.[14] To determine if this technique would aid in sweeping quartz, we performed the following experiment: We removed the +Z-region of a lumbered Y-bar and then cut this bar in half by cutting perpendicular to the Y-axis. The resulting Y-bars were swept under identical conditions. Evaporated chrome-gold electrodes were used along with a flowing air atmosphere. The only

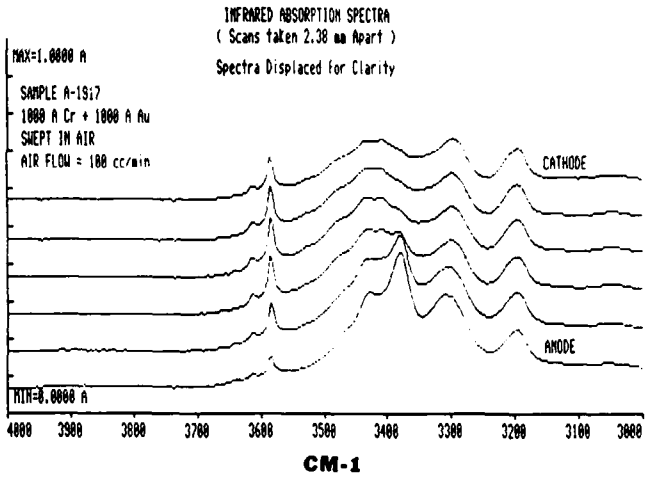



Fig. 9. A series of infrared scans similar to Fig. 1. Comparison with Fig. 2 indicates that both show depletion of the 3585 $cm^{-1}$ (as-grown OH) band in the anode region. This air-swept (thick electrode) sample resembles a $N_2$-swept sample.

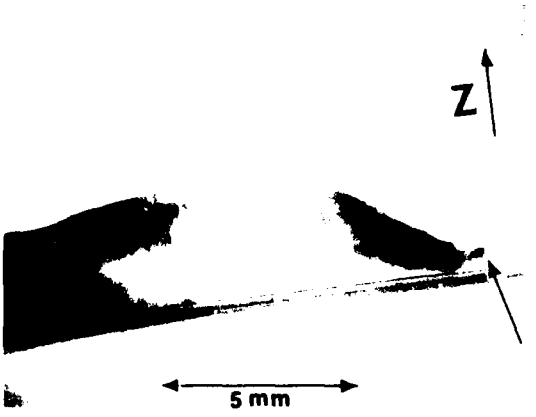



Fig. 10. A photomicrograph showing color-center striae formed as a result of air-sweeping with (1000 A Cr + 1000 A Au)-evaporated electrodes. The arrow points to the color-center boundary and to the end of the metallization.

difference between the two sweeping runs was that several thin shallow scratches were made in both anode and cathode metallizations of one of the bars. The result of this treatment produced an increase in transported charge of approximately 15 percent.

In air, there is about 250 times more hydrogen available, due primarily to water vapor than to free hydrogen. To decide which species is more active at the anode, we monitored the sweeping current while changes were made in the composition of the flowing gas. Table III shows the result of a series of gas changes on the sweeping current. Water vapor, labelled "+ $H_2O$," was added by bubbling the gas through water. In Table III, increases (decreases) in current correlate better with increases (decreases) in $H_2O$ content, rather

than $H_2$ content. For this reason, $H_2O$ is favored as the consumable gas during air-sweeping. Hydrogen is more likely the consumable gas when sweeping is performed in hydrogen-rich atmospheres such as forming gas.

TABLE III

Sweeping Current vs. Gas Composition

[Sample A - Cr/Au Metallization]

| Step | Atmosphere | Current(μA)* | [$H_2$]ppm+ | [$H_2O$]ppm+ |
|---|---|---|---|---|
| (1) | UHP $N_2$ | 18.4 | 1 | 1 |
| (2) | UHP $N_2$ +$H_2O$# | 19.2 | 1 | 24,500 |
| (3) | Air($H_2O$) | 20.1 | 100 | 24,500 |
| (4) | Air($H_2O$)+$H_2O$# | 20.1 | 100 | 24,500 |
| (5) | Forming Gas | 18.6 | 50000 | 1 |
| (6) | Forming Gas +$H_2O$# | 19.3 | 50000 | 24,500 |

* After a 3 min. equilibration time.

+ ppm values are approximate.

# Water vapor was added by bubbling the gas through water.

## Electrochemical Mechanisms

The following mechanisms seem reasonable, given the prior observations. With water vapor availability and porous electrodes, protons can be formed at the anode by the oxygen evolution reaction

$$H_2O(g) \text{ -----> } 2H^+ + 1/2\ O_2(g) + 2e^-,$$

i.e., water vapor is consumed at the triple junctions, protons enter the quartz, and electrons are supplied to the metal. In a hydrogen-rich atmosphere, the electrolytic reaction

$$H_2(g) \text{ -----> } 2H^+ + 2e^-$$

would be highly probable. We would expect these processes to occur together with electrochemical oxidation of the metal electrode (when high potentials are used)

$$M \text{ -----> } M^{z+} + Ze^-.$$

Both the anodic hydrogen and metal introduction reactions would require a cathodic reaction such as

$$2H^+ + 1/2\ O_2(g) + 2e^- \text{ -----> } H_2O(g),$$

where protons exiting the quartz would unite with oxygen and electrons would be supplied by the metal at the triple points. Alkali-metal ions would also be expected to be reduced at the cathode. If insufficient oxygen is available, then the reaction to be expected is a hydrogen evolution reaction

$$2H^+ + 2e^- \text{ -----> } H_2(g).$$

Confirmation of these mechanisms will have to await the construction of a two-compartment gas cell. A quartz sample disc, in the process of being swept, would separate the anode and cathode compartments. In this way, evolving gases could be analyzed and current densities could be monitored as a function of gas species introduced at each compartment.

## CONCLUSIONS AND FUTURE PLANS

Good metal and electrical contact is essential for hydrogen introduction during sweeping. Hydrogen indiffusion also is enhanced by porous electrodes which probably facilitate electrochemical reactions. Electrode-metal indiffusion is suppressed by polished Z-surfaces and the use of low electric fields ( $\lesssim$ 250 V at 500°C ) during sweeping. In the future, we intend to find the optimum field/temperature values, which will allow uniform sweeping without the deleterious effects of electrode-metal indiffusion. Confirmation of the proposed electrochemical reactions is another objective. We also plan to investigate glow-discharge electrolysis as a means of implanting hydrogen into quartz without electrode contact.[15]


## ACKNOWLEDGEMENT

The author thanks William Washington for cutting and polishing of samples, Melvin Wade for the evaporated metallizations, Robert Pfeffer and Christopher Wrenn for the RBS work, Donald Eckart for the SEM photographs, and Dr. John Vig for helpful suggestions.


## REFERENCES


1. John G. Gualtieri, "Further Studies on Electrode-Diffusion-Suppressed-Swept Quartz," Proc. 41st Annual Frequency Control Symposium (AFCS), pp. 192-198, 1987.
2. John G. Gualtieri and John R. Vig, "Sweeping and Irradiation Studies in Quartz," Proc. 38th AFCS, pp. 42-49, 1984.
3. John G. Gualtieri, "The Influence of Temperature and Electric Field on the Etch-Channel Density in Swept-Cultured Quartz," Proc. 39th AFCS, pp. 247-254, 1985.
4. Ferdinand Euler, Herbert G. Lipson, Alfred Kahan and Alton F. Armington, "Characterization of Alkali Impurities in Quartz," Proc. 36th AFCS, pp. 115-123, 1982.
5. Gerda B. Krefft, "Effects of High-Temperature Electrolysis on Coloration Characteristics and OH-Absorption Bands in Alpha-Quartz," Radiation Effects 26, pp. 249-259, 1975.
6. R. D. Maurer, "Effect of Catalyst Size in Heterogeneous Nucleation," J. Chem. Phys. 31, pp. 444-448, 1959.
7. G. W. Arnold, "Near-Surface Nucleation and Crystallization of an Ion-Implanted Lithia-Alumina-Silica Glass," J. Appl. Phys. 46, No. 10, pp. 4466-4473, October, 1975.
8. N. Boes and H. Zuchner, "Electrochemical Methods for Studying Diffusion, Permeation and Solubility of Hydrogen in Metals," J. of Less-Common Metals, 49, 223-240, 1976. Reprinted in Hydrogen in Metals, Elsevier Sequoia S.A., Lausanne 1976.
9. M.E. Martins, J.J. Podesta and A.J. Ariva, "Chemical Evidence of Hydrogen Sorption Processes on Potential Cycled Gold Electrodes," Electrochimica Acta 32,

No. 7, pp. 1013-1017, 1987.
10. F. A. Lewis, "The Palladium / Hydrogen System," Academic Press, London and New York, 1967, p. 12.
11. Z. M. Turovtseva and L. L. Kunin, Analysis of Gases in Metals, Consultants Bureau, New York, 1955, p. 4.
12. John G. Gualtieri and Donald W. Eckart, "The Influence of Surface Finish and Metallization on Electrode Electromigration in Alpha-Quartz During Sweeping," Proc. 40th AFCS, pp. 115-120, 1986.
13. G. Heiland, Z. Physik 128, 144, 1950.
14. S. V. Karpachev and A. T. Filyaev, "The Problem of Electrochemical Kinetics in the Case of a Solid Electrolyte," Electrokhimia, 2, pp. 1215-1216, 1966.
15. Shizuo Taniguchi, Tamio Endo and Koichi Sugiyama, "Depth Distributions of ESR Centers Induced in Fused Quartz by Gas-Discharges," Jap. J. of Appl. Phys. 25, No. 9, pp. L722-L724, September, 1986.

42nd Annual Frequency Control Symposium - 1988

# EVALUATION OF RESONATORS FABRICATED FROM HIGH QUALITY QUARTZ

J. J. Martin, Augusto R. Lopez
Department of Physics
Oklahoma State University
Stillwater, OK 74078

A. F. Armington
RADC/ESM
Hanscom AFB, MA 01731

and

J. F. Balascio
Motorola, Inc.
Carlisle, PA 17013

## ABSTRACT

Previously, the authors have reported on the production and properties of high purity, low dislocation quartz. In these studies, etch channel densities less than 10/cm$^2$ were also routinely achieved by using X-growth seeds and by sweeping. Aluminum contents below 3 ppm were routinely obtained by combining slow growth with cultured quartz nutrients. The lowest aluminum concentrations which were below 0.5 ppm were obtained by using Z-growth cultured quartz as the nutrient. The low aluminum content stones also showed very small OH growth-defect bands. The 5 MHz mechanical Q's ranged from 1.3 million for a crystal grown from a natural quartz nutrient to 2.5 million for crystals grown from cultured nutrients in commercial autoclaves. A slightly higher mechanical Q of 2.7 million was found for crystals grown from Z-growth cultured nutrient in research vessels. Sweeping consistently raised the mechanical Q of the 5 MHz 5th overtone AT-cut test blanks. For instance, sweeping raised the mechanical Q of one run from 2.28 million to 2.7 million. The room temperature mechanical Q seems to be limited by a lithium related defect which is removed by sweeping. EPR, low temperature IR, and low temperature acoustic loss measurements were used to determine the aluminum concentrations. The aluminum content ranged from 23 ppm for material grown from natural quartz down to 0.11 ppm for material grown from Z-mined cultured quartz. Acoustic loss measurements were made on unswept, Na-swept, and air or hydrogen-swept resonators. The 53 K Al-Na loss peak was not observed in the unswept resonators; this result confirms that the $Li_2CO_3$ added to the mineralizer excluded sodium as the charge compensator for the substitutional aluminum. The unswept crystals showed small anelastic loss peak at 305 K which limits the mechanical Q of the sample. A larger loss peak near 335 K was observed in the Na-swept samples; this peak is in addition to the usual 53 K Al-Na loss peaks. The 53 K Al-Na peak was used to estimate the aluminum content in the Na-swept samples. Air or hydrogen sweeping removed the 305 K lithium related peak. Irradiation at room temperature increased the lithium and sodium room temperature loss peaks. As expected, irradiation of the Na-swept crystals caused positive frequency offsets while smaller negative offsets were observed for the unswept crystals. No irradiation induced loss peaks were observed in the air or hydrogen swept samples. These samples showed smaller negative offsets than the unswept crystals taken from the same run. This indicates that sweeping may be important for radiation hardness even in low-aluminum material.

## INTRODUCTION

The need for high purity, low dislocation density quartz is growing with the increasing use of quartz oscillators in aerospace applications and with the increased use of very high frequency resonators and other devices that are fabricated by etching techniques. This increased need has stimulated research into quartz growth techniques[1-5]. Armington and Balascio[1,2] have reported on growth techniques for the production of low dislocation density, low aluminum, and low OH content cultured quartz in commercial size autoclaves. Low dislocation material can be routinely produced by the use of seed taken from the X-growth region of a synthetic quartz crystal. These runs yielded quartz with etch channel densities less than 10/cm$^2$ whereas conventional seeds produced densities in the 200-400/cm$^2$ range. Air sweeping or electrodiffusion[1,6] also resulted in etch channel densities less than 10/cm$^2$. These results are very important for the production of the very thin bulk vhf resonator crystals and other quartz devices fabricated using photolithographic techniques. Brice[7] has reviewed the selection of quartz for a number of commercial applications. He has pointed out in particular that the yield of "good" devices increases greatly when high quality material is used.

The reduction of impurities, principally aluminum, can be achieved both by reducing the growth rate[8] and by the use of high purity nutrient[1,4,5]. The aluminum related point defects are particularly important since they seem to be the cause of most of the radiation-induced frequency offsets[9]. Martin has reported on the acoustic loss peaks associated with the aluminum centers[10]. Martin and Armington[8] have reported that the OH-related growth-defects (as-grown OH) and the aluminum content increase rapidly for crystal growth rates greater than about 0.6 mm/day for quartz grown in a sodium hydroxide mineralizer. The carbonate process seems to require a slower growth rate of around 0.3 mm/day. In general, the higher the purity of starting material the higher the purity of the resulting crystal. Two nutrients have been found to be most effective for the growth of high purity quartz. These are 1. synthetic quartz mined from the Z-growth region[1,4] and 2. quartz crystallized from high purity fused quartz[5]. The first is based upon the segregation of impurities into the X-growth region of the quartz stone. The second is base upon the intrinsic very high purity of synthetic fused quartz. This second technique requires an additional autoclave run.

We report here a set of evaluation tests made on resonators fabricated from quartz grown from several different types of cultured nutrients in both commercial and research size autoclaves. This work is a continuation of the earlier reports[1-3]. The aluminum

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

content was determined by three different physical techniques: electron paramagnetic resonance, EPR; low temperature IR absorption measurements on swept material, and acoustic loss measurements on sodium-swept resonator blanks. Low temperature IR measurements were also used to monitor the as-grown OH bands (OH growth-defects). The room temperature mechanical Q was measured on 5 MHz 5th overtone AT-cut resonator blanks. Acoustic loss versus temperature spectra were taken on unswept, air or hydrogen-swept, and on Na-swept blanks to characterize the relation between the defect related loss and the growth conditions. The effect of room temperature irradiation on the acoustic loss spectra was also investigated.

## EXPERIMENTAL PROCEDURE

Cultured quartz for this project was grown in commercial size and experimental size autoclaves. Both sodium hydroxide and sodium carbonate mineralizers were employed. In all cases, lithium carbonate was was added to the mineralizer. The details of the growth techniques have been described by Armington and Balascio[2] and by Balascio and Armington[1]. Samples designated ME-- were grown in 4 inch diameter vessels with hydroxide mineralizer. Samples designated MH-- and MC-- were grown in 10 inch autoclaves using hydroxide and carbonate as the mineralizer respectively. Samples designated as X-- were grown at Hanscom in a 4 inch vessel with hydroxide mineralizer. Three grades of starting materials were used for the nutrient in the growth runs. Natural quartz was used for one run, MC1-7. Cultured quartz from previous standard commercial runs was employed for most of the work reported here; this starting material is designated CUL. QTZ. I. CUL. QTZ. III was the highest grade nutrient used in these runs. It was regrown cultured quartz with the +X-growth region removed. CUL. QTZ. III was used for runs X67, X78, and ME24-48. The aluminum content of the three nutrient grades was 21 ppm, 7 ppm, and 2.4 ppm for natural quartz, CUL. QTZ. I., and CUL QTZ. III respectively[1]. The growth rates for most of the runs were below the 0.6 mm/day[8] threshold above which the impurity content increases rapidly. Both standard seeds and seeds taken from the +X-growth region of cultured stones were used.

The aluminum concentration was directly measured by atomic absorption, AA, analysis for the earlier runs and by an inductively coupled plasma source mass spectrometer, ICP/MS, for the later runs. We will refer to this type of analysis as chem analysis. The OSU electron paramagnetic resonance, EPR, test was used as the main physical determination of the aluminum concentration[11, 8]. Additional physical aluminum determinations were also made using low temperature infrared absorption measurements and low temperature acoustic loss measurements. These techniques are described below. All test samples were taken from the Z-growth region of the stone.

Standard polished plano-convex AT-cut 5 MHz 5th overtone resonator blanks were fabricated from the Z-growth region of the stones. Most of the blanks were manufactured by Piezo Crystal while a few of the earlier ones were made by K-W Mfg. Both 14 mm and 15 mm diameter blanks were employed. So that a comparison between swept and unswept resonators could be made additional blanks were made from either air swept bars from the same growth run or several the finished blanks from the unswept material were hydrogen swept. The air or hydrogen swept samples have the Al-Li center converted into the Al-OH center which is IR active and can be used to determine the aluminum content. Additional blanks from each run were Na-swept. This replaces the lithium which charge compensates the aluminum with sodium. The resulting Al-Na center has a strong acoustic loss peak which can be used to determine the aluminum concentration.

Liquid nitrogen temperature infrared, IR, absorption scans were made on samples from each of the growth runs. These samples were 14 or 15 mm diameter 5 MHz 5th overtone AT-cut resonator blanks. The samples were mounted with the X-axis vertical in a liquid nitrogen cryostat. The cryostat had $CaF_2$ windows. A Beckman model 4240 spectrophotometer was used to make the scans over 3700 $cm^{-1}$ to 3100 $cm^{-1}$ range. The spectrophotometer slit was set to 0.12 mm and the scans were made at a rate of 20 $cm^{-1}$/min for improved resolution. IR scans on as-grown samples were used to characterize the OH growth-defect bands at 3581, 3437, 3398, and 3350 $cm^{-1}$. Similar scans were also made on air swept or on hydrogen swept samples from each run. Air or hydrogen sweeping converts the aluminum-alkali centers into Al-OH centers which has two absorption bands; a strong band at 3367 $cm^{-1}$ and a weaker band at 3306 $cm^{-1}$. We have used the 3367 $cm^{-1}$ band as one physical method of determining the aluminum content. Based upon low temperature IR scans on a number of swept 5 MHz 5th overtone AT-cut blanks with aluminum contents ranging from below 1 to about 70 ppm we find that the aluminum content, c, in atomic ppm is given by $c = 46A$ where A is the peak absorbance (optical density) of the 3367 $cm^{-1}$ Al-OH band. The reader should be cautious when attempting to transfer this result to other spectrophotometers since the observed width (and height) of the band is affected by the resolution of the instrument.

The room temperature mechanical Q of the quartz was determined for each run. The mechanical Q was measured on unswept 14 mm or 15 mm diameter 5th overtone contoured polished AT-cut blanks taken from the Z-growth region of the stones. No significant difference was observed between the 14 mm and 15 mm diameter blanks. The log-decrement method was employed for the Q measurements with the blanks mounted in an evacuated gap holder. We have used the same measuring system and similar gap holders for acoustic loss measurements over the 4 to 400 K temperature range. We often observe Q's of 100 million at low temperatures; consequently, we assume that the measuring system does not significantly limit the room temperature mechanical Q. The Q of air or hydrogen swept blanks from each run was also measured.

The acoustic loss versus temperature spectrum was measured on as-grown, Na-swept, and air or hydrogen swept 5th overtone AT-cut blanks. Most of the acoustic loss runs were made using a automated transmission method with the resonator blank mounted in a gap-holder. The automated system which was previously described by Martin, Hwang, and Wilson[12] allows nearly "hands-off" operation over the 8 to 400 K temperature range. The measured crystal resistance was converted into loss, $Q^{-1}$, by determining the crystal inductance with an independent measurement of the room temperature Q using the log-decrement method. Sodium swept blanks were prepared using the process described by Doherty *et al.*[13]. The Al-Na center produces a strong anelastic loss peak at 53 K for 5 MHz crystals which can be used to determine the aluminum content. Based upon these and other in-house measurements of the 53 K loss peak in sodium swept samples we find that the Al-Na concentration, c, is given by

$$c = 0.044 DL_{53} \qquad (1)$$

where $DL_{53}$ is the height of the peak in units of $10^{-6}$. Eq. 1 should be used in place of the earlier result found by Martin[10]. The acoustic loss versus temperature spectra of unswept, Na-swept, and air swept

5 MHz 5th overtone blanks from growth run ME24-48 were measured as a function of radiation dose. An 18.8 krad/hr $^{60}Co$ gamma source was used as the radiation source.

## RESULTS AND DISCUSSION

Figure 1 presents, in bar graph form, the room temperature mechanical Q measured on unswept and air or hydrogen swept resonator blanks taken from the Z-growth region of the quartz stones grown for this project. For purposes of comparison, results for the blanks from Sawyer run number D14-45 are also shown in Fig. 1. The Q was measured using the log-decrement technique with the blank mounted in an evacuated gap holder. Since we often observe Q's of 100 million at low temperatures we assume that the measuring system is not significantly limiting the Q of the blanks at room temperature. The maximum mechanical Q that we observed was 2.8 million for swept X67, X78, and D14-45 5 MHz blanks. Warner[14] in his studies on natural quartz found that the maximum Q for a 15 mm dia 5 MHz blanks is about 3 million. Warner showed that the Q is inversely proportional to the frequency of operation. The ME24-48 samples were 5.6 MHz 5th overtone blanks and, therefore, swept ME24-48 probably has a higher mechanical Q then the 2.6 million value shown in Fig. 1. The Q of unswept MC1-7 was 1.3 million and its swept mate showed a Q of 2.4 million. While this is the most dramatic case observed in this study all swept samples showed an increase Q. The acoustic loss of a resonator is the sum of the intrinsic loss, the loss due to defects, the loss due to the mounting structure and electrodes, and the loss due to the damping caused by the surrounding atmosphere. These last two contributions are estimated to be less than $1x10^{-8}$ by our observations of the acoustic loss at low temperatures. The intrinsic loss is caused by interactions between the thermal phonons and the thickness shear AT-cut crystal vibration[15]. At room temperature, the intrinsic loss is nearly independent of temperature and is inversely proportional to the frequency of vibration. Thus, Warner's observation of a maximum Q of 3 million for an AT-cut crystal suggests that the intrinsic loss at 300 K and 5 MHz must be about 3 million. This phonon loss should not depend strongly upon the presence of defects in the crystal. Therefore, it seems reasonable to suppose that improvement in Q that we observe upon sweeping is caused by the modification of a defect (or defects) which have may have an anelastic loss peak near room temperature. We will see later that this is the case.

Table 1 gives the aluminum concentration for these samples as determined by the traditional analytical chemical methods of atomic absorption, AA, spectroscopy or by inductively coupled plasma source mass spectrometry, ICP/MS, and by three "physical" measurements on the solid quartz samples. These are: 1. Electron paramagnetic resonance, EPR[11], measured on a small sample taken from the Z-growth region of the stone. 2. The low temperature (80K) IR absorbance (optical density) of the 3367cm$^{-1}$ Al-OH center band measured on an air of hydrogen swept resonator blank. 3. The height of the 53 K Al-Na center acoustic loss peak measured in a sodium-swept resonator blank.

With the exception of the chemical analysis on MH11-19 which shows only 2.7 ppm aluminum while the physical methods give 9-12 ppm and possibly X67 where the chemical analysis gives 0.8 ppm aluminum and EPR gives 0.11 ppm, all of the different methods give aluminum contents which are in good agreement. The Al-OH band in ME24-48, X67, and X78 was too small to detect with our spectrophotometer. While X67 and X78 were grown on +X seeds and have a very low dislocation density[3] their aluminum content is the similar to that of ME24-48 which was grown on a conventional seed. These three samples which have the lowest aluminum concentration were grown from recrystallized quartz with the +X-growth region removed.

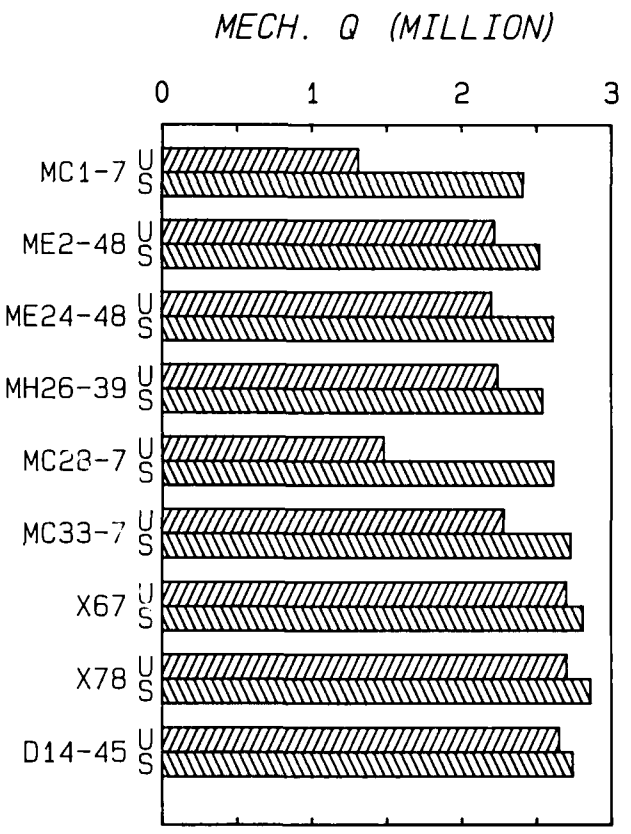



Figure 1. The room temperature mechanical Q of the unswept, U, and the air or hydrogen swept, S, 5 MHz 5th overtone AT-cut blanks fabricated from the growth runs listed are shown in bar graph form. Results for Sawyer Premium Q run D14-45 are also shown. In all cases, sweeping improved the mechanical Q.

Table 1. Aluminum content, C(ppma), by various methods.

| Sample | A(IR) | C Al-OH | $DL_{53}$ | C Al-Na | C EPR | C Chem |
|---|---|---|---|---|---|---|
| MC1-7 | 0.30 | 14 | na | - | - | 14 |
| ME2-48 | 0.057 | 2.6 | 80 | 3.6 | 3.6 | na |
| MH11-19 | 0.20 | 9.2 | 270 | 12 | 12 | 2.7 |
| ME24-48 | nd | nd | 4.4 | 0.2 | 0.19 | 0.9 |
| MH26-39 | 0.05 | 2.3 | 58 | 2.6 | 1.3 | 1.8 |
| MC28-7 | 0.13 | 6.0 | 135 | 6.0 | 5.2 | 3.3 |
| MC33-7 | 0.095 | 4.4 | - | - | 4.4 | na |
| X67 | nd | nd | - | - | 0.11 | 0.8 |
| X78 | nd | nd | 7 | 0.3 | 0.19 | 0.19 |
| D14-45 | 0.012 | 0.55 | 13 | 0.6 | 0.75 | |
| PQ-E | 0.21 | 9.7 | 200 | 9 | 8-10 | |

nd = not detected, na = not available.
The Al-Na loss peak, $DL_{53}$, is given in units of $10^{-8}$.
Chem analysis for MC1-7, ME2-48, and MH11-19 was by atomic absorption. The rest of the samples were analyzed using ICP/MS techniques.

Figure 2 shows the low temperature IR absorption spectra taken on an unswept and a hydrogen swept 5 MHz 5th overtone AT-cut blanks from run MC28-7. Since these are plano-convex samples we have not converted the absorbance values into absorption coefficients. All of the IR data was taken on these 5 MHz 5th blanks so internal comparisons can be readily made. The curve for the swept sample has been offset by 0.1 absorbance units. The $s_1$ OH growth-defect (as-grown OH) at 3350 $cm^{-1}$ is weak for this sample orientation. The $s_2$ through $s_4$ OH growth-defect bands at 3398, 3437, and 3581 $cm^{-1}$ are easily observed in the unswept spectrum. For the hydrogen swept sample, the 3367 and 3307 $cm^{-1}$ Al-OH $e_2$ and $e_1$ bands have come in. The 3367 $cm^{-1}$ band was used to determine the Al-OH content given in Table 1 above. We have never observed the two Al-OH bands in conventionally grown cultured quartz prior to sweeping or irradiation. It is interesting to note, that the OH growth-defect bands increase markedly when the sample was swept. For example, in the as received sample the 3581, $s_4$ has an absorbance of 0.05; this nearly doubled to 0.090 for the swept sample. From Fig. 1 we see that the room temperature mechanical Q of MC28-7 increased from 1.48 million to 2.54 million when the sample was swept. A similar result was observed for sample MC1-7 which also showed a large Q increase when swept. Figure 3 shows the low temperature IR absorption spectra taken on unswept and hydrogen swept samples from run X67. This sample which has a very low aluminum concentration has the lowest growth defect bands observed. No Al-OH (e) bands are observed in the curves for the hydrogen swept sample. Nor was any increase in the OH growth-defect bands observed. This sample has an unswept Q of 2.7 million which increased to 2.8 million when swept.

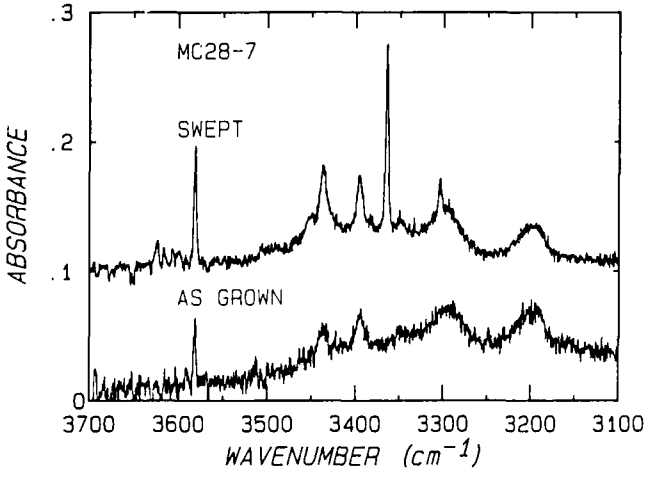



Figure 2. The low temperature IR absorption spectra for an as received unswept MC28-7 sample and for the same sample after hydrogen sweeping are shown. Sweeping produced the 3367 $cm^{-1}$ and 3306 $cm^{-1}$ bands. Sweeping also increased the strength of the growth-defect, s, bands and raised the mechanical Q from 1.48 to 2.61 million.

Figure 4 compares the acoustic loss versus temperature curves for as received (unswept), Na-swept, and air swept 5 MHz 5th AT-cut blanks from growth run ME24-48. Similar data were taken for the other growth runs. The small loss peak just above room temperature in the curve for the as-received sample is the only difference between its acoustic loss spectrum and that of the air swept sample. It is this loss peak that is limiting the room temperature Q of these samples. This

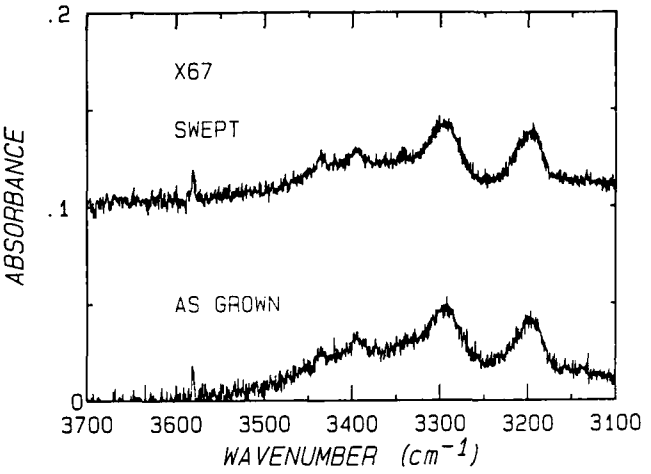



Figure 3. The low temperature IR absorption spectra for an as received unswept X67 sample and for the same sample after hydrogen sweeping are shown. The Al-OH bands too weak to observe in this 0.11 ppm Al sample. Sweeping raised the mechanical Q from 2.7 to 2.81 but any changes in the growth-defect bands were too small to observe.

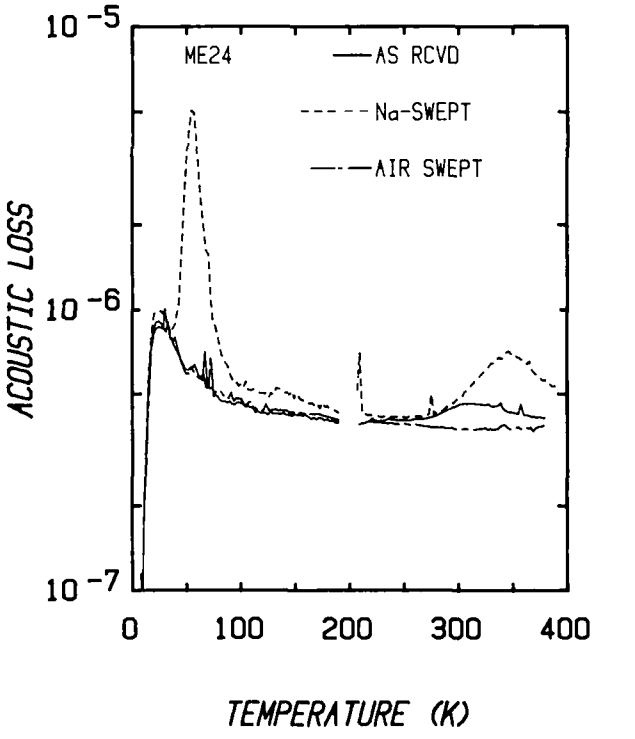



Figure 4. The acoustic loss versus temperature curves for as received, Na-swept, and air swept 5 MHz resonators blanks from run ME24-48 are shown. The as received spectrum shows a loss peak at about 305 K which limits the room temperature Q. This peak is not present in the air swept sample. In the Na-swept sample the room temperature peak has been shifted to about 335 K. The height of of the 53 K Al-Na peak in the Na-swept sample gives an aluminum concentration of 0.2 ppm.

peak which has a maximum at about 305 K was first reported by Koehler and Martin[18] who observed it Toyo Supreme Q quartz after a series of room temperature irradiations. They also reported a loss peak near 335 K in a sodium-swept resonator. This higher temperature peak is seen in Fig. 4 for our Na-swept blank. The Na-swept blank also shows, as expected, the 53 K Al-Na loss peak. The height of this 53 K peak was used to determine the aluminum content reported in Table 1 for the Na-swept samples. No 53 K Al-Na loss peaks were observed in the as-received unswept samples from any of the growth series. Thus, we conclude that the addition of the $Li_2CO_3$ to the mineralizer effectively excluded sodium from the point defects in these samples. The loss peak at 305 K is most likely caused by lithium trapped at an as-yet unknown defect. When the sample is sodium swept the lithium is replaced by sodium and the anelastic loss peak shifts to 335 K. When the sample is air or hydrogen swept the alkali is replaced with a hydrogen; the anelastic loss is removed; and the room temperature mechanical Q increases. Since the Q increase correlates somewhat with the increase in the OH growth-defect bands it is tempting to assume that the responsible defect is one of the "growth-defects" that has trapped a lithium during crystal growth rather than a hydrogen. Lopez, Hwang, and Martin[17] have determined the activation energies and relaxation times for the lithium and sodium related defects.

A series of irradiation tests were carried out on the unswept, Na-swept, and air swept resonator blanks from run ME24-48. Figure 5 shows the acoustic loss versus temperature curves for the unswept sample in the as received condition and after a 75 krad, and 1Mrad gamma irradiation. The irradiations were carried out at room temperature with the blank removed from the gap-holder. The radiation induced changes in the spectrum are small. There may be a small increase in the 305 K lithium related peak. The shifts seen around 150 K seem to be caused by remounting the blank in the gap-holder. The aluminum hole center produces a loss peak at 100 K[18] and two other peaks one at 135 K and one at 23 K are associated with its presence[17]. Based upon the results of Lopez, Hwang, and Martin[17] the maximum 100 K Al-hole loss in these 0.2 ppm Al sample would be $1.2 \times 10^{-7}$ which would be just barely observable. We should expect some room temperature stable aluminum hole centers in the unswept sample but apparently, too few are produced for detection with acoustic loss techniques.

Figure 6 shows the results for the air swept sample in the as received condition and after irradiations of 75 krad and 1 Mrad. No observable changes in the loss spectrum were produced by the irradiations. Very few room temperature stable Al-hole centers would be produced in a swept sample. Consequently, we should not expect to observe the Al-hole center related loss peaks in this very low aluminum content swept sample. The measured fractional frequency offsets for both the unswept and swept ME24-48 resonator blanks when irradiated to a does of 1 Mrad are given in Table 2 for both the upper and lower turnover temperatures. The positive offset for the upper turnover in unswept ME2-48 may be caused by a small increase in the 305 K loss peak. The small positive shift for the swept sample is not understood. While some of the offset may be caused by remounting the blanks in the gap holder the shifts for the swept sample are definitely smaller. Table 2 also shows the

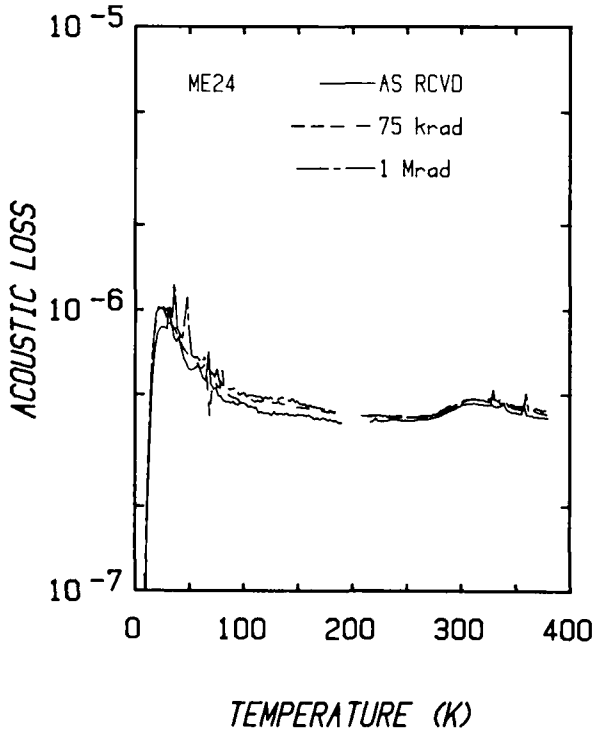



Figure 5. The acoustic loss curves for the unswept ME24-48 blank are shown in the as received condition and after two room temperature irradiations. There may be small radiation-induced increase in the 305 K peak. The increased loss around 150 K is probably caused by remounting the crystal in the gap holder.

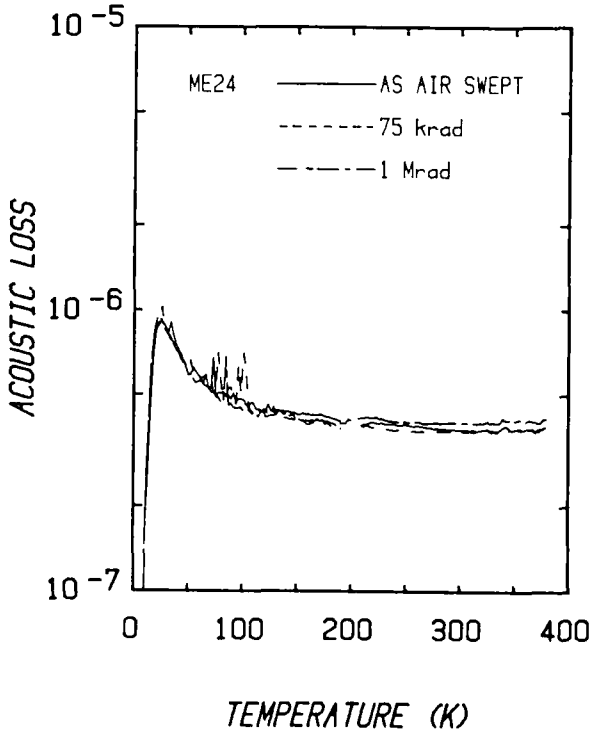



Figure 6. The acoustic loss versus temperature spectra for the air swept ME24-48 blank are shown in the as received condition and after two room temperature irradiations. No radiation-induced changes were observed.

fractional offsets for a pair of 10 MHz 3rd overtone finished crystals made from MH11-19 quartz. MH11-19 has about 12 ppm aluminum compared to the 0.2 ppm in ME24-28. For this higher aluminum content sample the radiation induced offsets are much larger. Again the swept sample shows smaller offsets. This result and the removal of the 305 K peak indicates that sweeping improves the performance of even very high purity quartz.

Table 2. Fractional frequency offsets in ppm at the upper and lower turnover temperatures caused by a 1 Mrad radiation dose.

| | ME24-48 | | MH11-19 | |
|---|---|---|---|---|
| | upper | lower | upper | lower |
| as grown | +1.55 | -0.78 | -3.20 | -3.48 |
| swept | +0.30 | +0.36 | -1.11 | -1.05 |

The sodium swept ME24-48 sample was given two relatively low dose irradiations. Figure 7 shows the acoustic loss spectra in the as-Na-swept condition and after irradiations to total doses of 14 krad and 38 krad. After 38 krad the the 53 K Al-Na peak is nearly gone; the small remaining peak corresponds to only about 13 ppb Al-Na centers compared to the original 200 ppb. Lopez, Hwang, and Martin[17] report that the Al-Na center shows an initial decay of the form $dL_{Na}/dt = -KL_{Na}\exp(-KD)$ where K ranges from 6 to 15 for samples with larger aluminum contents. For this Na-swept low aluminum sample the Al-Na center went out with a decay constant *of at least 40 which is much faster* than their results. No radiation-induced Al-OH centers were observed during our IR absorption tests on the ME24-48 samples nor were any detected in the air swept sample. A similar very rapid decay of the Al-Na center was also observed in a Na-swept X78 sample. Perhaps, this very fast decay of the Al-Na center (the Al-Li should go out the same way) is a measure of the initial conversion of the Al-Na into Al-hole centers. The 335 K sodium related peak increases rapidly with the irradiation and at 38 krad has nearly reached saturation. The defect responsible for this loss peak is an alkali trap; the lithium version is responsible for the 305 K loss peak in the as grown samples. The sodium version is 5 to 10 times stronger than the lithium analog. If we assume that the maximum height of the 335 K peak shown in Fig. 7 corresponds to all of the sodium from the Al-Na centers being trapped in these defects then we must have about 0.2 to 0.3 ppm of these defects. No paramagnetic defects are observed in quartz prior to irradiation. Therefore, it seems likely that the alkali is replacing a hydrogen; this behavior is consistent with the observed radiation induced decrease in the growth-defect bands where hydrogen is probably transferred to the aluminum site.

## CONCLUSIONS

By using cultured quartz with the +X-growth region removed and slow growth techniques cultured quartz with an aluminum concentration near 0.2 ppm was consistently grown. The "mined" nutrient gave better results than were obtained by using the full cultured quartz stone as the nutrient. The low aluminum samples also showed very low OH growth-defect infrared bands. No effects were observed that could be attibuted to differences between hydroxide and carbonate mineralizer.

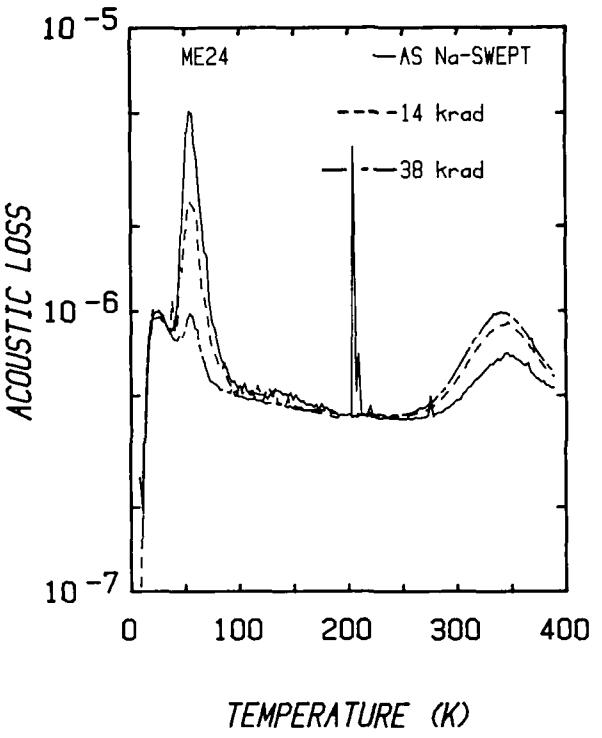



Figure 7. The acoustic loss curves for the Na-swept ME24-48 sample are shown in the as-Na-swept condition and after two low dose room temperature irradiations. *The Al-Na center decayed very rapidly.* The 335 K Na-related peak near 335 K increases with irradiation; at a dose of 38 krad it is very near saturation.

The very low aluminum samples had a room temperature mechanical Q of about 2.7 million. Sweeping improved the mechanical Q of these samples to 2.8 million. The room temperature mechanical Q in as-grown quartz is limited by a defect which traps lithium and causes an anelastic loss peak at 305 K. Air or hydrogen sweeping replaces the lithium by hydrogen and removes the loss. Sweeping raised the mechanical Q of unswept MC1-7 from 1.3 to 2.4 million. This sample was grown from natural quartz. The increase in mechanical Q upon sweeping is accompanied by an increase in the strength of the OH growth-defect IR bands. This result and their behavior during radiation suggest that the responsible defect may be one of the "growth-defects" that has trapped an alkali rather than a hydrogen during growth.

No radiation induced acoustic loss peaks were observed in the as-grown or swept low aluminum content samples. Irradiation caused the Al-Na peak in a 0.2 ppm Na-swept sample to go much more quickly than the reported observations on samples with higher aluminum contents. Air or hydrogen sweeping improved the radiation-induced frequency offsets even in the low aluminum samples. This result plus the observed increase in mechanical Q suggest that even the highest purity quartz should be swept.

## ACKNOWLEDGEMENTS

The work at Motorola and at Oklahoma State University was supported by the Solid State Sciences Division, Rome Air Development Center, Hanscom AFB, MA. The chemical analyses were performed at the Eagle Picher Research Laboratory, Miami, OK. C. Y. Chen and

D. Davis did the EPR measurements of aluminum. The authors would like to thank T. Walker of Motorola, Inc., M. Harris, J. O'Connor, J. A. Horrigan, E. Cormier of RADC for their assistance during this investigation. D. Hart at OSU did much of the hydrogen sweeping. L. E. Halliburton provided much helpful advice.

## REFERENCES

1. J. F. Balascio and A. F. Armington, Proc. 40th Ann. Symposium on Freq. Control, IEEE, 70 (1986).

2. A. F. Armington and J. F. Balascio, Proc. 39th Ann. Symposium on Freq. Control, IEEE, 230 (1985).

3. A. F. Armington, J. A. Horrigan, M. T. Harris, and J. F. Balascio, Proc. 41st Ann. Symposium on Freq. Control, IEEE, 213 (1987).

4. Baldwin Sawyer and D. R. Kinloch, Proc. 38th Ann. Symposium on Freq. Control, IEEE, 8 (1984).

5. S. P. Doherty, S. E. Morris, D. C. Andrews, and D. F. Croxall, Proc. 36th Ann. Symposium on Freq. Control, IEEE, 66 (1982).

6. J. J. Martin, IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control, in press.

7. J. C. Brice, Rev. Mod. Phys. *57*, 105 (1985).

8. J. J. Martin and A. F. Armington, J. Crystal Growth, *42*, 203 (1983).

9. J. C. King and D. R. Koehler, in *Precision Frequency Control*, E. A. Gerber and A. Ballato, eds., vol. 1, pp 147, Academic Press, 1985.

10. J. J. Martin, J. Appl. Phys., *56*, 2536 (1984).

11. L. E. Halliburton, N. Koumvakalis, M. E. Markes, and J. J. Martin, J. Appl. Phys., *57*, 3565 (1981).

12. J. J. Martin, Ho. B. Hwang, and T. M. Wilson, Proc. 40th Ann. Symposium on Freq. Control. IEEE, 32 (1986).

13. S. P. Doherty, J. J. Martin, A. F. Armington, and R. N. Brown, J. Appl. Phys., *51*, 4164 (1980).

14. A. W. Warner, Bell Syst. Tech. J., *40*, 1193 (1960).

15. D. B. Fraser, *Physical Acoustics*, W. P. Mason, ed. Vol. V., p 2536, Academic Press, NY, 1968.

16. D. R. Koehler and J. J. Martin, J. Appl. Phys., *57*, 5205 (1985).

17. Augusto R. Lopez, Ho. B. Hwang, and J. J. Martin, Proc. 42nd Ann. Symposium on Freq. Control, IEEE, in press.

18. J. C. King, Bell Syst. Tech. J., *38*, 573 (1950).

42nd Annual Frequency Control Symposium - 1988

# DOSE DEPENDENCE OF RADIATION-INDUCED DEFECT CHANGES IN QUARTZ

Herbert G. Lipson and Ferdinand Euler
Rome Air Development Center
Solid State Sciences Directorate
Hanscom AFB, MA 01731-5000

## Summary

Radiation-induced changes in OH, Al-Na, Al-OH, and Al-hole defect center concentrations have been determined from infrared, optical laser, and acoustic loss measurements. Premium-Q and High-Q quartz samples with aluminum concentrations ranging from 0.5 to 8 ppm, and grown-in OH distributions varying over a factor of six were evaluated. The samples were given successive exposure to $^{60}$Co gamma rays at room temperature using two widely different dose rates, 0.5 and 30-76 krad/min. The defect concentrations decay or saturate exponentially with dose d, involving the term exp(-kd). The rate constants k decrease with increasing dose rate and with increasing hydrogen concentration. For low dose rate and low OH concentration, OH decays with $k \approx 0.06$ krad$^{-1}$ while for high dose rate and high OH concentration, we find $k \approx 0.006$ krad$^{-1}$. The dose rate effect on the rate constants is attributed to insufficient time to form stable defects during irradiation. A similar dose rate dependence was found for Al-Na decay from acoustic loss measurements. The dose dependence of the Al-Na decay is the sum of two exponentials, and the higher of the two rate constants is comparable to that of the OH decay. The rate constants found for Al-OH production are generally lower than those for OH decay. This means that hydrogen is not transferred directly and exclusively from grown-in OH to Al-OH but some is trapped in intermediate sites. Simultaneous Al-hole production is observed even with abundant hydrogen in the crystal. Relative Al-hole and Al-OH production rates depend on the hydrogen-to-aluminum ratio.

## Introduction

Changes in the point-defect structure of alpha quartz are largely responsible for the steady state frequency offset of quartz oscillators which occurs with ionizing radiation.[1-3] The principal impurities introduced during quartz growth are substitutional aluminum, and interstitial hydrogen and alkali ions (M = Na,Li). There are at least eight types of point defects involved in the radiation process.[4] Three are introduced during growth: an aluminum ion, charge-compensated with an alkali metal, Al-M, a hydrogen bonded to an oxygen anion, OH, and hydrogen in other unspecified sites.[5] The following four defects are produced by irradiation, an aluminum hydrogen center, Al-OH [6,7] an aluminum hole center, Al-h, electron traps, and alkali traps. The remaining defect is a trapped alkali center with an associated acoustic loss peak which can be produced either during growth or by irradiation.[8]

Ionizing radiation generates electron-hole pairs, which mostly recombine. A fraction of the free electrons and holes are trapped. Holes can be trapped at Al-M centers, to the extent that electron traps are available. If these Al-M-hole centers are formed above 200 K, alkali ions can move to other sites leaving a stable Al-h center. At the same time the radiation releases hydrogen ions from grown-in OH and other hydrogen sites. These hydrogen ions, as well as alkali ions, can be trapped and form specific defects. The best known of these is Al-OH, formed when hydrogen replaces the hole at an Al-h center. This hydrogen ion may come directly from a grown-in source or by way of an intermediate trap.

In addition to actual frequency offset measurements on oscillators, several experimental techniques are available for non-destructive characterization of these defects in quartz.[9-16] The Al-Na center has two acoustic and two dielectric loss peaks, OH and Al-OH centers have low-temperature near infrared hydroxyl vibration peaks, and Al-h centers have characteristic electron-spin resonance (ESR) spectra and optical absorption bands. One of the absorption peaks of the Al-h center is nearly coincident with the 632.8 nm He-Ne laser line and can be conveniently used to determine relative Al-h distributions.[17]

Most of the previous work on the effects of ionizing radiation on the various defect centers has been at either rad or Mrad dose levels. Very few investigations have been made using radiation in the krad dose range or at different dose rates. Only recently has a study been made of the growth of Al-h centers during irradiation at different dose rates, but the dose dependence of other centers such as $OH^-$, $Al\text{-}OH^-$ and Al-Na was not determined.[18]

In this paper we present new results on the dose and dose-rate dependence of the principal centers involved in the irradiation process. This investigation covers cultured quartz samples with a range of grown-in OH and aluminum centers. We have used acoustic loss measurements to determine Al-Na decay, infrared scans to measure local variations in OH decay and Al-OH production, and an optical laser method to determine Al-h production. These studies have been made at two widely different dose rates. We observe large differences in OH decay and Al-OH production with dose rate. Curve fitting of these decay and production curves has been used to determine the rate constants involved. Direct evidence of Al-OH and Al-h defect competition has been obtained from simultaneous measurements of these defects with dose.

## Experimental Procedures

The quartz samples used for our spectroscopic studies were rectangular in shape and between 1 and 2 cm in size, with parallel x, y, and z faces. Table 1 lists the quartz type, origin and growth run designation, sample number, aluminum concentration range, the range of Al-OH band peak heights measured at 85 K, the range of OH 3581 cm$^{-1}$ band peak heights at 85 K, used as a relative measure of grown-in OH concentration, and the $^{60}$Co gamma ray dose ranges and rates used. The High-Q material was grown without the lithium salt additive used for Premium-Q and has a relatively high sodium content. Aluminum concentrations were determined from ESR measurements. Infrared transmissions were measured between 3100 and 3700 cm-1 with the focused 3 mm beam of a Nicolet 170SX Fourier spectrophotometer. The sample, mounted inside a Dewar cooled to 85 K, was scanned normal to the z-axis with the unpolarized infrared beam aligned along the crystal x- or y-axis. Beam positions are designated by the distance of the beam center from the sample z face furthest from the seed. Measurements were also made with the beam parallel to the crystal z-axis (E⊥z), scanning between the +z to -x or +y to -y faces. The peak absorptions of the strong narrow 3581 cm$^{-1}$ and 3366 cm$^{-1}$ bands are used to monitor relative changes in grown-in OH and

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

Table 1. Quartz sample characterization; dose and dose rate.

| quartz type | origin and growth run | sample identification | [Al] (ppm) | impurity band peak height ($cm^{-1}$) 3366 $cm^{-1}$ * E⊥x;y | 3366 $cm^{-1}$ * E⊥z | 3581 $cm^{-1}$ ** E⊥x;y | 3581 $cm^{-1}$ ** E⊥z | dose range (krad) | dose rate (krad/min) |
|---|---|---|---|---|---|---|---|---|---|
| Premium-Q | S BH-A | 1 | 3.7-7.6 | 0.43-0.56 | 0.92-0.95 | 0.28-0.42 | 0.58-0.62 | 10-150 | 0.5 |
| | | a | 11 | | | 0.24-0.38 | 0.59-0.62 | 200-1600 | 30; 76 |
| | S D14-45 | 40A | 0.5-1 | 0.08-0.13 | 0.19-0.21 | 0.07-0.12 | 0.15 | 10-150 | 0.5 |
| | | 43A | 1 | | | 0.07-0.11 | 0.15 | 200-6400 | 30; 76 |
| High-Q | S E42-21 | 8A | 5-8 | 0.23-0.39 | 0.61-0.65 | 0.07-0.13 | 0.15-0.17 | 10-150 | 0.5 |
| | | BBA19 | 6-8 | | | 0.08-0.12 | 0.15-0.16 | 200-3200 | 30; 76 |
| | | | | | | | | 1-150 | 0.45; 3.9(r) |
| experimental | R QA-38 | M(a) | 0.4-0.7 | 0.07-0.11 | 0.15 | 0.11-0.14 | 0.20-0.22 | 5-270 | 0.5(G) |

S Sawyer Research Products Inc. * dose: 100 krad (r) resonators
R Rome Air Development Center ** unirradiated (G) 30 ppm Ge

irradiation-produced Al-OH, respectively. Scans acro the z-face of the crystal show the full variation of OH and Al-OH in the growth direction, but with lower overall absorption than those made with the beam along the z-axis where the maximum polarized strength of each band is measured. More detailed descriptions of the infrared bands, measurement method, and polarization effects are given in our previous publications.[5,19]

The low-dose-rate samples were also scanned with a He-Ne 632.8 nm randomly polarized laser aligned along the x- or y-axes of the crystal. Absorption was determined for several beam positions that are representative of the 2.5 x 2.5 x 8 mm size ESR samples cut from an adjacent region of the crystal. In this way a direct comparison was made with ESR Al-h concentration values. Al concentration values were determined by ESR for adjacent sections of the same bar from which the infrared samples were cut. Since ESR samples could not be obtained from all of the regions covered by the infrared scans, local variations in aluminum are still possible beyond those shown in Table 1, as denoted by [Al]. The local variations in the 3366 $cm^{-1}$ Al-OH band peaks obtained at 100 krad included in Table 1, are in part due to the aluminum distribution. Samples were irradiated at room temperature in the RADC $^{60}Co$ source using successive doses of 200, 400, 800, 1600, 3200, and 6400 krad for the maximum dose rate. Up to an accumulated dose of 800 krad, this was 30 krad/min. Owing to a source change, irradiations up to an accumulated dose of 6400 krad were made at dose rates between 73 and 79 krad/min. Sample temperatures were not measured during irradiation, but based on previous measurements, the temperature rise does not exceed 11°C for doses up to 800 krad. For larger doses the temperature rises proportionally and may have reached 68°C at highest doses. Successive doses of 10, 15, or 20 krad at dose rates ranging from 0.47 to 0.55 krad were used for the low dose rate studies.

Several 5 MHz, 5th over .e AT-cut precision resonators were fabricated .m E42-21 quartz by Frequency Electronics, Inc. Two of these resonators received a series of $^{60}Co$ exposures at room temperature. Resonator Q was measured as function of temperature, initially and after each irradiation. A more detailed description of the Q measurement techniques was given in a recent publication.[20] It was previously found that the Q values changed with time when the resonator was kept at room temperature for several days or above 50°C for a few hours. Therefore, the complete Q(T) spectrum, with T increasing from 15 to 400 K, was measured twice after each irradiation. The first measurement, designated here as "quenched", was made with the resonator kept at room temperature no longer than 30 minutes between irradiation and measurement. The second measurement designated as annealed, was made after the resonator had been above 350 K for at least 40 minutes. The differences between the two sets of Q values are significant at the temperatures of radiation-sensitive acoustic-loss peaks. The two sets of data merge above 360 K. Resonator 1 was irradiated with dose rates of 0.45 ± 0.03 krad/min and resonator 2 with 3.9 ± 0.2 krad/min. The incremental doses ranged from 1 through 40 krad, and heating of the resonator during irradiation did not exceed a few degrees.

## Experimental Results

Figure 1 shows the absorption variation of the 3581 $cm^{-1}$ OH and 3366 $cm^{-1}$ Al-OH band peaks along the z-growth axis for the 0 to 100 krad dose range for sample D14-45. The range of absorption coefficients for OH, measured before irradiation, and for Al-OH, measured at an accumulated dose of 100 krad, have been included in Table 1 for all samples irradiated with a dose rate of 0.5 krad/min. Data obtained with the maximum dose rate of our source for D14-45, between 0 and 6400 krad, E42-21, between 0 and 3200 krad, and BH-A between 0 and 1600 krad, can be found in an earlier publication.[5] Figure 1 shows that the OH distribution of sample D14-45 increases sharply toward

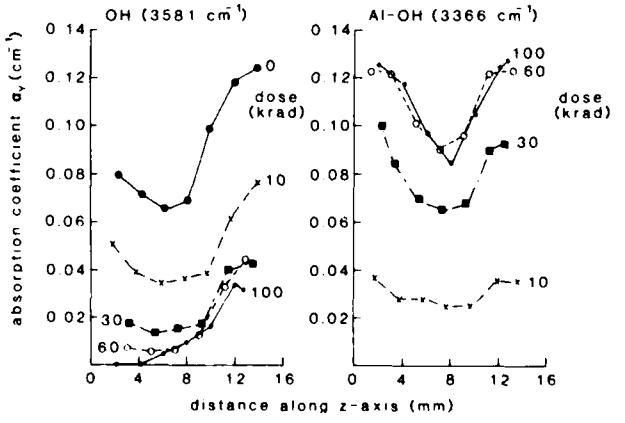



FIG. 1. Variation of impurity band absorption in quartz sample D14-45,40A, measured at 85 K (E⊥Y), after $^{60}Co$ gamma irradiation at room temperature; dose rate: 0.5 krad/min; [Al] = 0.5-1 ppm; z=0, furthest from seed.

the seed side of the crystal. The dip in Al-OH in the center of the sample results from variations in both hydrogen and aluminum concentrations across the crystal. Measurements on the high dose rate sample extending to higher dose levels indicate that, in this case, the aluminum decreases toward the seed. An increase in OH toward the seed was found for E42-21 and QA38. The OH concentration of BH-A is about three to five times larger than that of the other samples and peaks in the center. The Al-OH concentration was found to increase toward the seed side for E42-21, BH-A and QA-38, following the aluminum concentration.

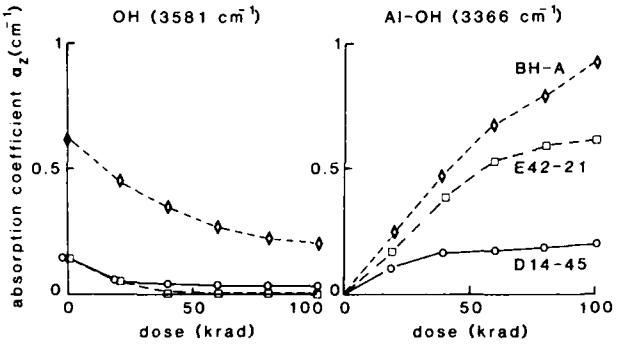



FIG. 2. Impurity band absorption of quartz samples measured at 85 K (E⊥Z), as function of dose, after $^{60}$Co irradiation at room temperature; dose rate: 0.5 krad/min.

The data shown in Figure 2, obtained with the beam aligned along the z-axis, essentially averages the effect of impurity variations introduced during growth. In this Figure we show the dose dependence of OH decay and Al-OH production for samples D14-45, E42-21 and BH-A. Above 50 krad OH depletes for E42-21 while residual OH is still observed at 100 krad for the other samples. Al-OH production approaches saturation at different rates for each sample. Although OH levels in D14-45 and E42-21 are nearly similar, the Al-OH level of E42-21 exceeds that of D14-45 by a factor of three. In fact, one third of the Al-OH absorption increase occurs above 40 krad after the OH is totally depleted. This is very strong evidence for the presence of a second source of grown-in hydrogen, with no known infrared signature.[5] The results for high and low dose rates are shown in Figures 3 and 4, log-log scale plots of absorption coefficient versus dose. The striking feature observed here is the strong dependence of OH decay and Al-OH formation an dose rate. In fact to produce the same change in OH or Al-OH the required dose is an order of magnitude larger at the high dose rate.

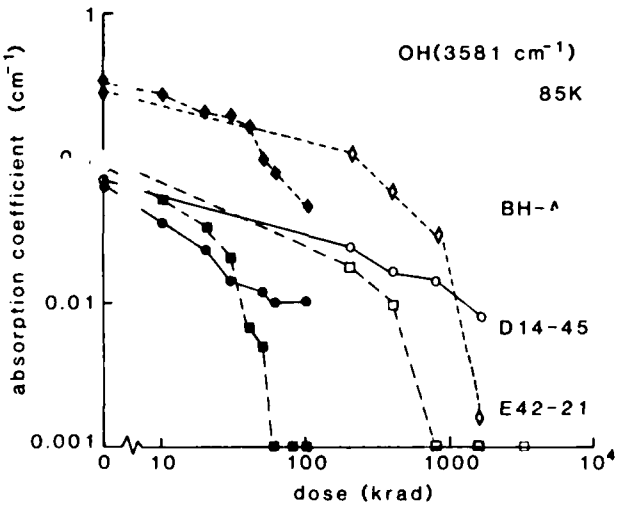



FIG. 3. Grown-in OH absorption after room temperature $^{60}$Co irradiation at different dose rates; full symbols: 0.5krad/min; open symbols: 30 krad/min below and 76 krad/min above 1000 krad; infrared measurements at 85 K, D14-45 and BH-A with E⊥Y, E42-21 with E⊥X.

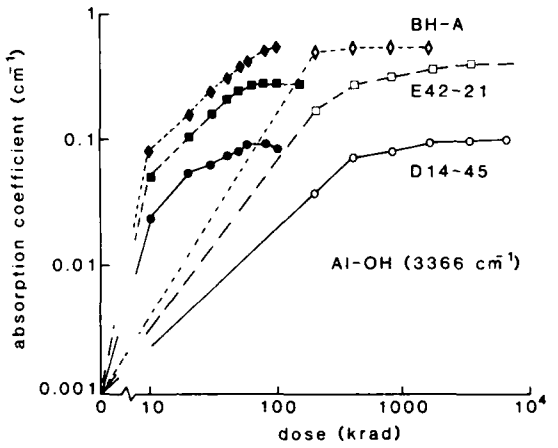



FIG. 4. Al-OH absorption after room temperature $^{60}$Co irradiation at different dose rates; full symbols: 0.5 krad/min; open symbols: 30 krad/min below and 76 krad/min above 1000 krad; infrared measurements at 85 K, D14-45 and BH-A with E⊥Y, E42-21 with E⊥X.

Figure 5 shows the absorption at 632.8 nm as a function of dose for two of the samples used in the low dose studies. Shown here are the spread and average of the data obtained with the beam positions close to the location of ESR samples. The dose dependence of E42-21 is corroborated by two ESR measurements after 50 and 90 krad accumulated dose, indicating 0.20 and 0.53 ppm of Al-hole centers respectively. This was not the case with BH-A, for which the absorption rises with dose at a much lower rate, even though ESR measurements showed a steep dose dependence close to that of E42-21. Sample D14-45 which had much lower aluminum concentration did not yield any reliable 632.8 nm absorption data.

Figure 6 shows acoustic loss spectra 1/Q(T) on a log scale measured with 14, 60 and 125 krad accumulated doses. For clarity, only the quenched spectra are shown. There are eight known major loss peaks affected by radiation.[1,8,9,21-25] In this paper, we discuss only the Al-Na α peak at 52 K. The dose dependence of the loss maximum for resonators irradiated with different dose rates is shown in Figure 7, for both the quenched and annealed states. In this semi-logarithmic plot, the decay curves show a change in slope in the vicinity of 50 krad, and hence, cannot be interpreted with a single exponential function. The figure clearly shows that the higher dose rate leads to a lower decay rate, similar to the decay of the infrared absorption of grown-in OH defects. There are also some differences between the quenched and annealed states. Up to 40 krad, these differences are masked by the

uncertainties in determining the base values of 1/Q, but at higher doses the losses for the annealed state are 5 to 10% higher than for the quenched state. Simultaneously, the Al-Na β peak at 135 K shows that the quenched state has higher losses, 30 to 50% above the annealed values. Apparently, radiation moves some of the $Na^+$ ions from Al-Na α to β sites. After irradiation these ions return to the α sites at temperatures above ≈280 K.

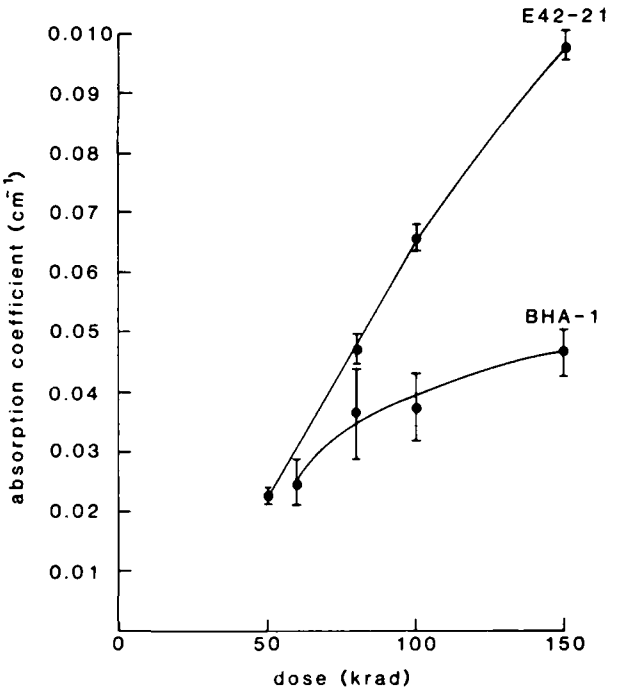



FIG. 5. Al-hole absorption at 632.8 nm, measured at room temperature as function of dose, after room temperature $^{60}Co$ irradiation; dose rate: 0.5 krad/min; E42-21 (E⊥X), BH-A (E⊥Y).

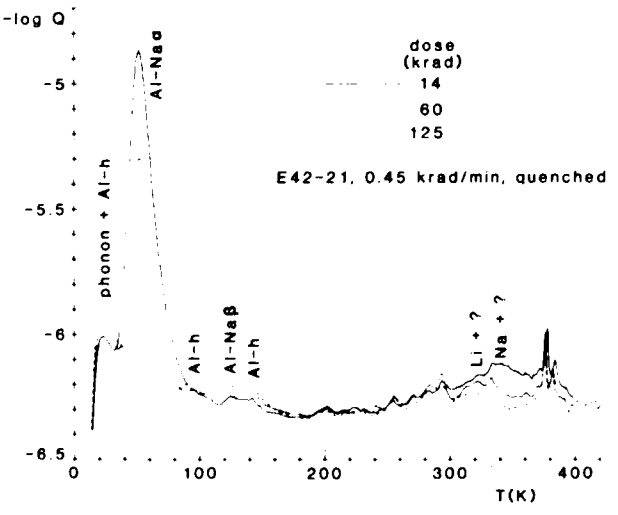



FIG. 6. Acoustic loss spectra 1/Q(T) for three selected doses of room temperature $^{60}Co$ irradiation.

## Analysis and Discussion

The simplex algorithm was utilized for fitting the dose dependence of the infrared absorption and resonator Q data with the general function

$$y = y_s + (y_o - y_s)\, e^{-kd} \qquad (1)$$

with dose d in krad, y either the measured absorption coefficient α or the Δ loss function

$$(1/Q)_{peak} - (1/Q)_{base} \qquad (2)$$

of the Al-Na α peak. Subscript o denotes the measured initial value (dose=0), s stands for the saturation value (dose= ∞); $y_s$ and the rate constant k are the two adjustable parameters.

The results for remanent OH, saturated Al-OH and rate constants for OH and Al-OH are shown in Table 2 for four different materials. Data was obtained for two different dose rates, normal and parallel beam alignment with respect to the z-axis, and several beam positions across the scanning axis. The materials are arranged from top to bottom in the order of increasing relative hydrogen content, as determined from the grown-in OH absorption coefficients before irradiation. A typical result, Figure 8, shows the fit of OH decay and Al-OH production for sample D14-45. For the two selected beam positions along the z-axis the saturation values and rate constants for Al-OH production are similar even though there is a difference in initial OH. The rate constant for OH decay is larger for the 12 mm position, where there is residual OH, than the 4 mm position, where OH is essentially depleted. The fitting error values are 0.002 $cm^{-1}$ for both cases of OH decay and 0.003 $cm^{-1}$ and 0.004 $cm^{-1}$ for Al-OH production. These errors are within measurement accuracy and are representative for our data, with the exception of that for BH-A (E⊥Y) which has a large absorption gradient along the z-axis. Higher fitting errors were also encountered with Al-OH data for E42-21.

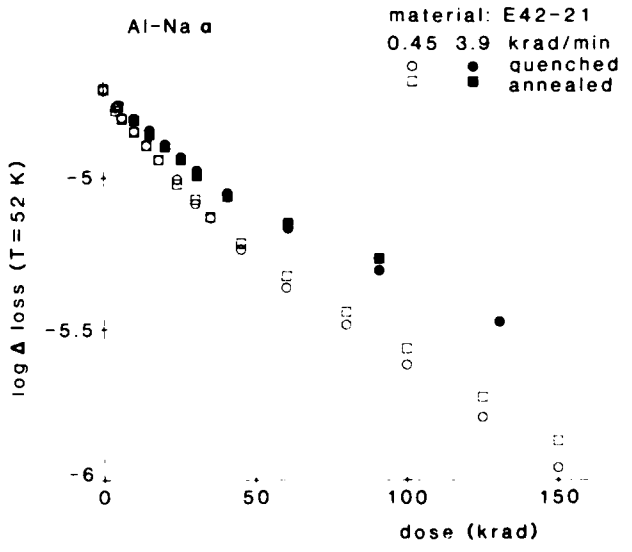



FIG. 7. Net acoustic loss, Δ loss = $(1/Q)_{peak}$ - $(1/Q)_{base}$ of the Al-Na α peak as a function of accumulated room-temperature gamma ray dose; resonator 1: 0.45 krad/min; resonator 2: 3.9 krad/min.

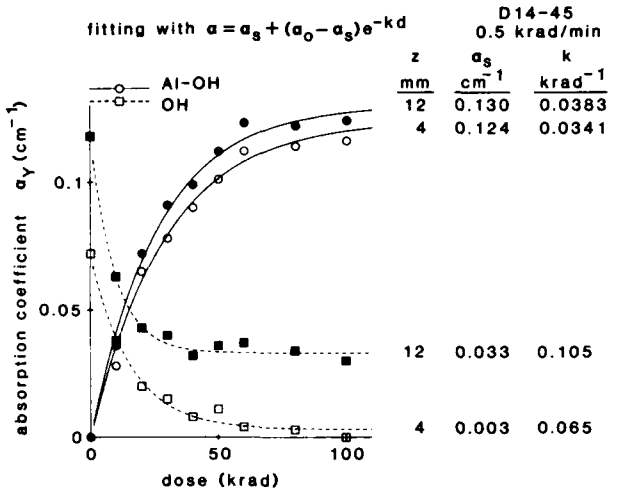



FIG. 8. Results of curve fitting the impurity band absorption data of Fig. 1; sample D14-45,40A, z-positions 4 and 12 mm.

Equation (1) is an appropriate model, and adequately fits all OH data. This means that the OH decay rate, as expected, is proportional to the prevailing OH concentration and is thus controlled by the supply of dissociable hydrogen, not by the available sites. The Al-OH production rate, proportional to aluminum sites not yet compensated by hydrogen, is controlled by the available sites, not by the hydrogen supply. Only the Al-OH dose dependence of E42-21, shown in Figure 9 cannot be adequately fitted with equation (1). These data have an inflection point in the region between 10 and 30 krad. The function

$$\alpha = \alpha_s (1 - e^{-kd})^2 \qquad (3)$$

yields a closer fit with the error reduced from 0.017 to 0.008 $cm^{-1}$. Its derivative, $2k\alpha_s e^{-kd}(1-e^{-kd})$, has a maximum at $d = k^{-1} \ln 2$, leading to an inflection point at that d value in equation (3). With $k = 0.049$ $krad^{-1}$ obtained for best fit, the inflection point is located at 14 krad. This means that there is a peak in the rate of Al-OH production. We suggest that the peak in this rate function results from a peak in the

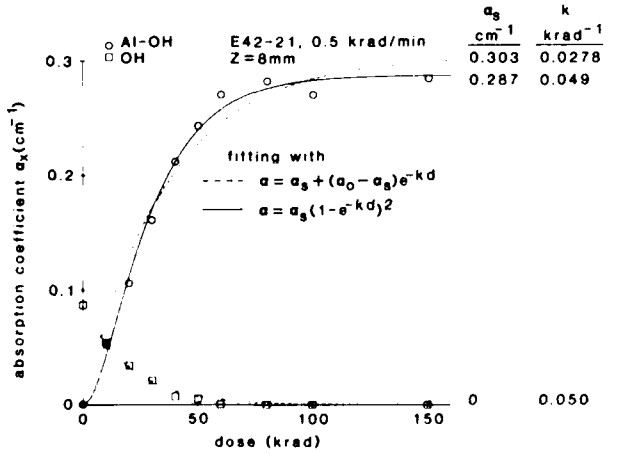



FIG. 9. Results of curve fitting the impurity band absorption data of sample E42-21,8A; z-position 8 mm.

Table 2. Dose dependence fitting with single exponentials.

| dose rate | * | growth run | ** | abs. coeff. ($cm^{-1}$) d = 0 OH | d = ∞ OH | d = ∞ Al-OH | rate constant ($krad^{-1}$) OH | Al-OH |
|---|---|---|---|---|---|---|---|---|
| low (a) | y | D14-45 | 4 | 0.072 | 0.003 | 0.124 | 0.065 | 0.034 |
| | | | 8 | 0.068 | 0.010 | 0.096 | 0.080 | 0.040 |
| | | | 12 | 0.118 | 0.033 | 0.130 | 0.105 | 0.038 |
| | x | E42-21 | 5 | 0.069 | 0 | 0.253 | 0.055 | 0.029(c) |
| | | | 8 | 0.087 | 0 | 0.303 | 0.050 | 0.028(c) |
| | | | 12 | 0.105 | 0 | 0.369 | 0.047 | 0.025(c) |
| | y | QA-38 | 4 | 0.116 | 0.035 | 0.081 | 0.028 | 0.021 |
| | | | 8 | 0.132 | 0.030 | 0.098 | 0.028 | 0.0193 |
| | | | 12 | 0.143 | 0.026 | 0.133 | 0.032 | 0.0166 |
| | y | BH-A | 4 | 0.296 | 0.056 | 0.544 | 0.0183 | 0.0154 |
| | | | 8 | 0.403 | 0 | 0.774 | 0.0070 | 0.0120 |
| | | | 12 | 0.331 | 0 | 0.769 | 0.020 | 0.0131 |
| | z | D14-45 | 8 | 0.145 | 0.029 | 0.195 | 0.075 | 0.045 |
| | z | E42-21 | 10 | 0.148 | 0 | 0.733 | 0.061 | 0.020(c) |
| | z | BH-A | 8 | 0.620 | 0.135 | 1.474 | 0.021 | 0.0097 |
| max (b) | y | D14-45 | 4 | 0.068 | 0.003 | 0.128 | 0.0051 | 0.0016 |
| | | | 8 | 0.065 | 0.011 | 0.100 | 0.0065 | 0.0027 |
| | | | 12 | 0.103 | 0.033 | 0.110 | 0.0074 | 0.0050 |
| | x | E42-21 | 4 | 0.087 | 0 | 0.280 | 0.0123 | 0.0027 |
| | | | 8 | 0.100 | 0 | 0.368 | 0.0079 | 0.0031(c) |
| | | | 10 | 0.113 | 0 | 0.406 | 0.0113 | 0.0034(c) |
| | y | BH-A | 4 | 0.272 | 0.065 | 0.405 | 0.0099 | 0.0082 |
| | | | 8 | 0.370 | 0.110 | 0.544 | 0.0056 | 0.0111 |
| | | | 12 | 0.300 | 0.023 | 0.553 | 0.0058 | 0.0113 |
| | z | D14-45 | 8 | 0.154 | 0.033 | 0.223 | 0.0069 | 0.0033 |
| | z | E42-21 | 8 | 0.156 | 0 | 0.565 | 0.0087 | 0.0032(c) |
| | z | BH-A | 4 | 0.620 | 0.152 | 1.598 | 0.0121 | 0.0042 |

* beam direction parallel to the x, y, or z axis as indicated.
** beam positions (mm) measured across the scanning axis, *normal* to the beam direction.
(a) 0.5 krad/min.
(b) source maximum:
30 krad/min for d < 1 Mrad; 76 krad/min for d > 1 Mrad.
(c) fitting error exceeds measurement error by factor two or more.

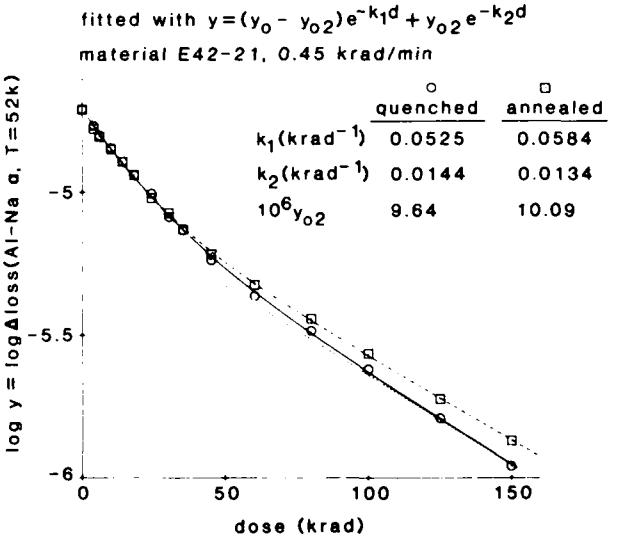



FIG. 10. Results of curve fitting the Al-Na α net acoustic loss from Fig. 7; resonator 1.

supply of hydrogen coming from trapping sites which first receive hydrogen from grown-in OH at lower doses and then release it to Al-OH with further irradiation.

The Al-h production curve for E42-21 has an inflection point between 60 and 100 krad, just above the 50 krad dose where OH depletes. Fitting with equation (3) yielded k in the order of 0.01 krad$^{-1}$. This value, however may not be reliable, since there is insufficient curvature at doses up to 150 krad to determine a valid saturation level. In this case, interpretation of the inflection point must involve simultaneous Al-h and Al-OH production competing for the same sites.

Figure 10 shows the fitting of the resonator data. The semi-logarithmic format demonstrates that fitting with a sum of two exponentials is appropriate for the Al-Na decay. In addition, fitting of log y vs dose was used since it emphasizes the more accurate lower loss values corresponding to higher Q values. The results for both resonators and both post-radiation states are shown in Table 3. The similarity of the normalized pre-exponentials implies that about half the Na goes into one type of site and half into another. Lopez et al[26] also found an appreciable second exponential decay term in a resonator fabricated from material with a high aluminum to hydrogen ratio, similar to that of E42-21. With the lower dose rate, the higher of the two rate constants, $k_1$ is similar to those found for the OH decay in E42-21. J.J. Martin[8] showed that an acoustic loss peak at 340 K is related to a radiation-induced defect center that contains Na. Our E42-21 quartz resonator shows that this peak, absent before irradiation, rises with dose and reaches saturation at 14 krad. It is not unreasonable to speculate that this defect receives Na mainly from Al-Na $\alpha$ by the process with the higher rate constant. The differences between quenched and annealed data are too small for evaluation.

Table 3. Al-Na $\alpha$ decay; dose dependence fitting with two exponentials

| resonator number | 1 | | 2 | |
|---|---|---|---|---|
| dose rate (krad/min) | 0.45 | | 3.9 | |
| measurement state * | q | a | q | a |
| rate constants (krad$^{-1}$) | | | | |
| $k_1$ | 0.0525 | 0.0584 | 0.0383 | 0.0483 |
| $k_2$ | 0.0144 | 0.0134 | 0.0082 | 0.0078 |
| normalized pre-exponentials | | | | |
| $(y_0-y_{02})/y_0$ | 0.51 | 0.48 | 0.51 | 0.47 |
| $y_{02}/y_0$ | 0.49 | 0.52 | 0.49 | 0.53 |

* q: quenched; a: annealed

Figure 11 shows the OH-decay rate constants of Table 2 plotted as function of measured initial OH absorption coefficients, $\alpha_o^{OH}$. In spite of the scatter in this data, we observe two trends. First at the low dose rates, k decreases with increasing initial hydrogen concentration, and second, k decreases with increasing dose rate.

Figure 12 is a plot of the rate constants for Al-OH production vs those for OH decay. It shows that Al-OH k values are generally lower than those for OH. Direct transfer of hydrogen from OH exclusively into Al-OH requires that the sum of the concentrations [OH] + [Al-OH] is constant, i.e. independent of dose. This, in turn, requires that the rate constants $k_{OH}$ and $k_{Al-OH}$ are equal, if both [OH] and [Al-OH] have the exponential dose dependence shown in equation (1). Since we generally find $k_{OH} > k_{Al-OH}$, the sum of the concentrations is not constant but follows the equation

$$[OH] + [Al\text{-}OH] = [OH]_o - [H_{trap}], \qquad (4)$$

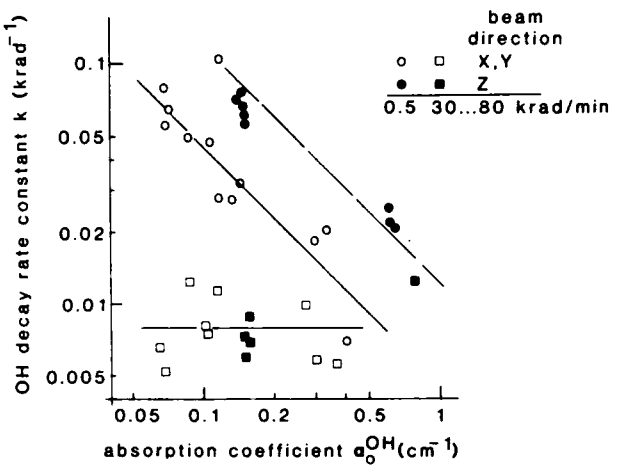



FIG. 11. OH decay rate constants as function of the initial OH absorption coefficients. The slanted trend lines are drawn for log-log slope -1, for E⊥X,Y and E⊥Z, for the low dose rate. The horizontal line represents the average of all high dose rate data.

where $[OH]_o$ denotes initial concentration and $H_{trap}$ the hydrogen trap postulated earlier. The concentration $[H_{trap}]$ behaves like a function of the type $\exp(-k_{Al-OH}d)-\exp(-k_{OH}d)$, which increases to a maximum as hydrogen coming from OH is trapped at low doses and decreases as hydrogen is released at higher doses to form Al-OH. This trap is present when $k_{OH} > k_{Al-OH}$. It can result in an inflection point in the Al-OH production curve, but only if the hydrogen supply is the limiting factor of that process. Figure 12 also shows that the data points lie essentially parallel to the solid line drawn for the equation $k_{OH} = k_{Al-OH}$. Therefore, the rate constants for the two processes are proportional to one another, and $k_{Al-OH}$ should also follow the trends shown for $k_{OH}$ in Figure 11.

It was already mentioned that Figure 11 shows that the rate constants of OH decay decrease with increasing dose rate. This is also observed here for

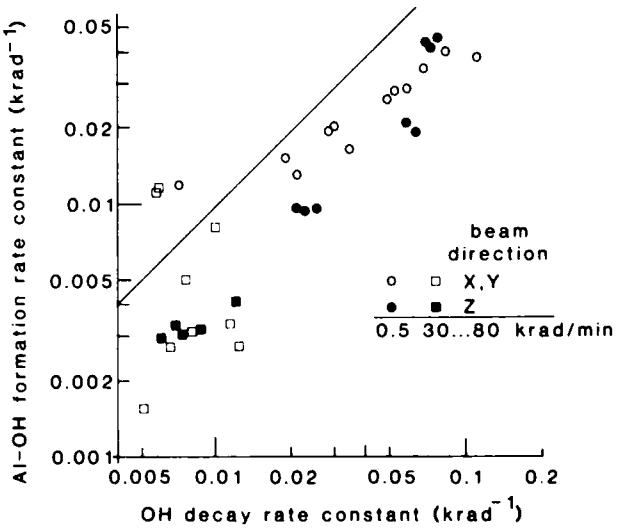



FIG. 12. Al-OH formation rate constants as function of OH decay rate constants. The slanted line indicates equal rate constants.

Al-Na decay and Al-OH production. The exponent in the power law $k=r^q$, r being the dose rate, can be used as an order-of-magnitude figure to describe the decrease of defect change rate constants k per increase of the dose rate. Both OH decay and Al-OH production show $q \approx -0.5$ for low and $\approx -0.3$ for high OH concentration. Examination of the Al-h production curves of ref. 18 shows that initial slopes decrease with increasing dose rate and an exponent $q \approx -0.5$ can be derived. For the Al-Na decay data in Table 3, one can derive q values ranging from -0.09 to -0.026. It should be pointed out that $q = -0.5$ is closely related to the time dependence $t^{0.5}$ known for one-dimensional diffusion. The time available for ions to diffuse in the quartz channels during irradiation is inversely proportional to the dose rate. If a rate constant depends on mobile ions finding stable trapping sites during irradiation, the value $q = -0.5$, can be expected for the process.

## Conclusion

We have determined the dose dependence of radiation-induced decay of grown-in OH and Al-Na and that of the production of Al-OH centers in four different types of quartz at two dose rates. The data were analyzed by curve fitting with exponential functions and the rate constants k for the three processes were determined. All rate constants decrease with increasing dose rate.

It was found that $k_{OH} > k_{Al\text{-}OH}$, requiring that part of the hydrogen transferred from OH to Al-OH, passes through a trapping site first. If the Al-OH production is hydrogen-supply limited, the trapped-hydrogen defect causes the appearance of an inflection point in the Al-OH production curve.

The need for sufficient time for ions involved in the radiation-induced processes to diffuse to trapping centers and finally to stable defects may be responsible for the dose rate dependence. Increased recombination rates at higher dose rates may also be a factor to be considered.

## Acknowledgments

The authors would like to acknowledge R. Andrews and L. Diaco for their competent technical assistance, S. Mittleman and J. Cappelli for performing the irradiation, P. Zecchini for his help with the computations and the manuscript, L. Halliburton and J. Martin of Oklahoma State University for ESR measurements, stimulating discussions and sharing their unpublished results.

## References

1. J. C. King and H. H. Sander, IEEE Trans. Nucl. Sci., NS-19, 23 (1972); Rad.Effects, 26, 203 (1975).

2. P. Pellegrini, F. Euler, A. Kahan, T.M. Flanagan and T. Wrobel, IEEE Trans. Nucl. Sci., NS-25, 1267 (1978).

3. F. Euler, H. G. Lipson, and P. A. Ligor, Proc. 34th Ann. Frequency Control Symp.; available from Electronic Industries Association, 2001 Eye St., Washington, DC 20006; 72 (1980).

4. L.E. Halliburton, J.J. Martin and D.R. Koehler in Precision Frequency Control, E.A. Gerber and A. Ballato, eds. Vol. I, p 32, Academic Press, 1985.

5. H. G. Lipson and A. Kahan, J. Appl. Phys., 58, 963 (1985).

6. A. Kats, Phillips Res. Reports, 17, 133 (1962).

7. R.N. Brown and A. Kahan, J. Phys. Chem. Solids 36, 467 (1975).

8. J. J. Martin, Ho B. Hwang, and H. Bahadur, Proc. 39th Ann. Frequency Control Symp.; available from IEEE, 445 Hoes Lane, Piscataway, NJ 08854; document no. 85CH2186-5; 266 (1985).

9. D. B. Fraser, Physical Acoustics, W. P. Mason, ed., New York: Academic Press, 1968, Vol. V, ch. 2, pages 59-110.

10. B. R. Capone, A. Kahan, R. N. Brown, and J. R. Buckmelter, IEEE Trans. Nucl. Sci., NS-17, (6), 217 (1970).

11. H. Jain and A. S. Nowick, J. Appl. Phys., 53, 485 (1982).

12. D. M. Dodd and D. B. Fraser, J. Phys. Chem. Solids, 26, 673 (1965).

13. H. G. Lipson, F. Euler, and A. F. Armington, Proc. 32nd Ann. Frequency Control Symp.; available from Electronics Industries Association, 2001 Eye St., Washington, DC 20006; 11 (1978).

14. M. E. Markes and L. E. Halliburton, J. Appl. Phys., 50, 8172 (1979).

15. K. Nassau and B. E. Prescott, Phys. stat. sol. (a), 29, 659 (1975).

16. B. K. Meyer, F. Lohse, J. M. Spaeth, and J. A. Weil, J.Phys.C: Solid State Phys., 17, L31(1984).

17. H. G. Lipson, Proc. 40th Ann. Frequency Control Symp.; available from IEEE, 445 Hoes Lane, Piscataway, NJ 08854; doc.no. 86CH2330-9; 63 (1986).

18. M. A. Mondragon, C. Y. Chen, and L. E. Halliburton, J. Appl. Phys. 63, 4937 (1988).

19. H. G. Lipson and A. Kahan, IEEE Trans. Nucl. Sci., NS-31, 1223 (1984).

20. 16 F. Euler and A. Kahan, Phys. Rev. B, 35, 4351 (1987).

21. D. B. Fraser, J. Appl. Phys., 35, 2913 (1964).

22. J. J. Martin, Proc. 38th Ann. Frequency Control Symp.; available from IEEE, 445 Hoes Lane, Piscataway, NJ 08854; doc.no. 84CH2062-8; 16 (1984).

23. D. S. Park and A. S. Nowick, Phys. stat.sol.(a), 26, 617 (1979).

24. J. J. Martin and S. P. Doherty, Proc. 34th Ann. Frequency Control Symp., 31 (1980), see ref. 3.

25. J. J. Martin, Ho B. Hwang, and T. M. Wilson, Proc. 40th Ann. Frequency Control Symp., 32 (1986), see ref. 17.

26. A.R. Lopex, Ho B. Hwang and J.J. Martin, Proc. 42nd Ann. Frequency Control Symp. (1988), this issue.

42nd Annual Frequency Control Symposium - 1988

# A PRODUCTION STUDY OF ACOUSTIC LOSS RELATED DEFECTS IN QUARTZ

Augusto R. Lopez, Ho B. Hwang, and J. J. Martin
Department of Physics
Oklahoma State University
Stillwater, OK 74078-0444

## ABSTRACT

Above 200 K, ionizing radiation converts the Al-Li and Al-Na centers into a mixture of Al-OH and Al-hole centers. Lithium is the majority alkali in as-grown cultured quartz; unfortunately, the Al-Li center does not have a spectroscopic signal. We have used the 53 K Al-Na center acoustic loss peak, the Al-OH infrared absorption, and the Al-hole center 100 K loss peak and its visible range optical absorption to track the production and destruction of these centers as functions of radiation dose in cultured quartz. Our results on a number of different sodium swept samples show that the Al-Na center goes out as exp(-KD) where K ranges between 6 and 40 and with the radiation dose, D, given in Mrad. IR absorption measurements of the Al-OH band on matching samples show that the Al-OH centers grow in as (1-exp(-K'D)) where K' approximately matches K, the decay constant of the Al-Na center. Only very small differences in the Al-OH production were observed between sodium and lithium containing samples. Thus, the decay of the aluminum alkali centers with irradiation seems to be controlled by the production of the Al-OH centers. For larger radiation doses, the decay of the Al-Na appears to go over to a slower decay. While the aluminum alkali and Al-OH destruction and production show this relatively simple behavior the production of stable Al-hole centers appears to be more complex. At saturation, the number of Al-hole centers produced in unswept quartz is governed by the availability of hydrogen which was originally trapped at the growth-defect sites. For the same aluminum content, relatively few Al-hole centers will be produced in quartz with strong as-grown OH bands. Loss peaks at 23 K and 135 K are also related to the presence of the Al-hole center. The 135 K peaks seems to consistently track with the 100 K peak. However, in some samples the 23 K peak is larger than the 100 K Al-hole center loss while in others it is smaller. Lithium and sodium related loss peaks at 305 K and 335 K respectively grow quickly at low doses. In some samples, these defects show only small decreases at larger doses while in others they go out completely.

## INTRODUCTION

Quartz controlled oscillators are used in a number of applications where radiation effects must be considered. When exposed to ionizing radiation, quartz oscillator crystals can exhibit both steady-state and transient frequency offsets and series resistance changes.[1-10] King and Koehler[1] have reviewed the radiation response of quartz with respect to frequency control applications. Norton, Cloeren, and Suter; and Suter and Maurer have recently reported low-dose frequency offset studies.[11,12] Flanagan, Leadon, and Shannon[13] have discussed mechanisms for low dose effects; and Palkuti and Truong[14] have described an automated radiation test facility. Radiation induced modification and production of defects which have anelastic loss peaks at temperatures below the operating temperature are thought to be the main cause of frequency offsets. The frequency offset, df/f, is then given by

$$df/f = -\Sigma dL_{mi} \qquad (1)$$

where $dL_{mi}$ is the change in the height of the loss peak for the ith defect. Thus, if a defect such as the Al-Na center which has a loss peak at 53 K is removed ($dL_{mi}$ is negative) a positive frequency offset is observed.

Cultured quartz contains a number of interesting point defects. Fraser[15] has discussed the early work on defect-related acoustic loss. Weil[16,17] has reviewed the paramagnetic defects which are observed after irradiation. A number of the point defects present n cultured quartz have been reviewed by Halliburton, Martin, and Koehler[18]. Quartz has a number of defects which trap hydrogen; often the hydrogen is bonded to one of the intrinsic oxygen atoms forming an infrared active OH molecule. Kats[18] carried out a classic study covering the OH related defects in both natural and cultured quartz. Cultured quartz has a set of OH-related defects which give rise to characteristic infrared bands at 3350, 3398,3437, and 3581 $cm^{-1}$. The defects responsible for these bands have not been identified. They are , however, the source of the hydrogen that is transferred to the aluminum site following irradiation at room temperature. In semiconductor terminology, aluminum substituting for silicon is an acceptor. Since quartz is grown in an environment that is rich in alkalis, sodium or lithium are trapped interstitially next to the aluminum. The sodium sits "off center" in the Z-axis channel; and the resulting Al-Na center causes the strong acoustic loss peak observed at 53 K in 5 MHz crystals. The Al-Na center also causes a much weaker loss peak at 135 K. Both of these acoustic losses have also been observed with dielectric loss techniques[20]. The lithium in the equivalent defect sits "on-center" and the Al-Li center shows neither acoustic nor dielectric loss[20]. Ionizing radiation creates free electrons and holes by excitation across the band gap. These free charge carriers migrate through the lattice during their short life time. The aluminum alkali center is then modified by the following process. One of the radiation-produced holes becomes momentarily trapped at the Al-Na (or Al-Li) site forming a Na-Al-hole center which is normally unstable. (This paramagnetic center can be observed at low temperatures by EPR techniques if a suitable electron trap such as Ge is present.) If the sample is above about 200 K during the irradiation, the Na (or Li) can diffuse away before the hole leaves or electron-hole recombination takes place. At temperatures below 200 K, the alkali is unable to escape before the hole recombines. If the alkali is subsequently trapped at another site and if no other processes take place then stable Al-hole centers are formed with the reaction going according to

$$(Al\text{-}Na) + hole \rightarrow (Na\text{-}Al\text{-}hole) \rightarrow Al\text{-}hole + Na \qquad (2)$$

However, at the same time hydrogen is released from the growth-defect (as-grown OH) sites and diffuses through the sample. When the hydrogen reaches an Al-hole center it replaces the hole forming the hydrogen equivalent to the aluminum-alkali center. The hydrogen prefers bonding to an adjacent oxygen so the center is IR active with bands at 3367 and 3306 $cm^{-1}$. We usually refer to this center as the Al-OH center.

CH2588-2/88/0000-176 $1.00 C 1988 IEEE

Equation 2 assumed that the irradiation was carried to completion and that no hydrogen was present. If hydrogen is available from the growth-defect sites and if only a partial irradiation is carried out a mixture of aluminum centers are present according to the equation

$$(Al\text{-}Na) \rightarrow x(Al\text{-}Na) + (1-x)[y(Al\text{-}OH) + (1-y)(Al\text{-}hole)]. \quad (3)$$

where x is the fraction of Al-Na remaining and y is the fraction of Al-OH centers. At saturation, x = 0, and we have y Al-OH centers and 1-y Al-hole centers. A similar equation would hold for the Al-Li center. Equation 2 is greatly simplified; it does not give the matching electron trapping, the alkali trapping, nor the source of the hydrogen. Acoustic loss measurements of the 53 K and 100 K loss peaks can be used to track the destruction of the Al-Na centers and the production of the Al-hole centers respectively. The Al-hole center can also be tracked by using EPR techniques and by visible band optical absorption measurements. The production of the Al-OH centers can be followed by using low temperature IR absorption measurements. Unfortunately, the Al-Li center does not have a corresponding spectroscopic signal. Hitt and Martin[22] found that while there are small differences between the release of lithium and sodium from the aluminum site their general overall behavior is the same.

Other point defects are also present in cultured quartz. Germanium can substitute for silicon in the quartz lattice and act as an electron trap. Weil[17] has given a complete discussion of the paramagnetic centers found in Ge-containing samples. Recently, Euler and Kahan[23] have reported an anelastic loss peak at 250 K (5MHz) caused by the Ge-Li center in germanium doped cultured quarts. This loss peak grows rapidly during the early stages of irradiation then decays at larger doses. Koehler and Martin[10] have reported a lithium-related loss peak at 305 K that has a sodium-related analog at 335 K. While its radiation response is similar to that of the Ge-Li center it does not seem to be connected to any known defect. Several types of oxygen vacancy centers have also been observed. However, no anelastic loss peaks have been observed that can be attributed to these vacancy centers.

Martin, Hwang, and Wilson[24] reported on the destruction of the Al-Na center and production of the Al-hole center 100 K acoustic loss peak following both electron and gamma irradiation in several cultured quartz samples. We report here an extension of their study to a number of different samples along with measurements of the Al-OH production in matching samples.

## EXPERIMENTAL PROCEDURE

Cultured quartz from a number of different sources was used for these experiments. With one exception, all of the crystals were taken from pure Z growth material. The exception was a R-growth stone, HA-A, which was the source of the high aluminum content material. The aluminum content of the quartz used in the study ranged from 2 ppm up to 60-70 ppm. Most of our measurements were made on Toyo Supreme Q quartz designated SQ-B and on Motorola grown material from autoclave run MH11-19. The Motorola quartz had about 12 ppm aluminum which was slightly more than the 8-10 ppm of the SQ-B samples. A few measurements were also made on crystals taken from a second Motorola bar, MH26-39, which had about 2 ppm aluminum. Low temperature IR scans were made on all of the quartz used in this study. The Toyo SQ-B quartz showed small growth-defect bands while the MH11-19 quartz which has a comparable aluminum content showed much larger OH bands. Frequency Electronics fabricated the 5 MHz 3rd overtone plano-convex AT-cut, BT-cut, and SC-cut blanks from the SQ-B sample. The rest of the blanks were made by Piezo Crystal. Most of these were plano-convex 5 MHz 5th overtone AT-cut blanks. 10 MHz 3rd overtone AT-cut and BT-cut crystals were made from the Motorola bar MH11-19. The blanks were sodium, lithium, or hydrogen swept using our standard process[25]. The blanks were then electroded and mounted in HC-40 holders using a silver-filled polyimide adhesive.

Most of the acoustic loss measurements were made using the automated transmission method system described by Martin, Hwang, and Wilson[24]. The manually tuned log-decrement method was only used on a few Na-swept samples. The finished crystals were mounted on the cold-head of a closed cycle helium refrigerator for measurements over the 9 - 300 K temperature range. They were then transferred to a separate "cryostat" that covered the 200 - 400 K range.

A $^{60}Co$ gamma source was used for the room temperature irradiations. A sample placed in the center position of the gamma source is currently irradiated at a rate of 18.8 kRad(Si) per hour. Low temperature irradiations were made "in-situ" with the refrigerator cold-head placed in-line with the 1.75 MeV electron beam from our Van de Graaff accelerator. The beam causes significant sample heating. This problem was overcome by chopping the beam so as to hold the sample temperature between 50 K and 90 K as indicated by the resonant frequency of the crystal. Successive room temperature irradiations and measurements were carried out until the samples reached saturation. A low temperature IR absorption measurements of the Al-OH center production were made on matching samples that were irradiated at the same time as the resonator crystals. Similar visible-uv absorption measurements of the Al-hole center production were also made on the high aluminum content sample.

## RESULTS AND DISCUSSION

Figure 1 shows the acoustic loss spectrum over the 9 to 400 K temperature range for a Na-swept 10 MHz 3rd overtone AT-cut crystal from bar MH11-19. The height of the low temperature peak for the as-Na-swept curve is consistent with the 12 ppm aluminum content of this sample. The temperature of peak is shifted slightly above the 53 K observed at 5 MHz because this is a 10 MHz crystal. The background loss is higher for the same reason. The as-Na-swept curve also shows a loss peak at 350 K; this is the same Na-related defect peak that we see at 335 K in 5 MHz crystals. As the sample was gamma irradiated the low temperature Al-Na peak decays to to less than 1% of its initial height at 1 Mrad. The 350 K peak increases with dose up to 0.21 Mrad then decreases slightly. IR measurements show that the Al-OH center production reached saturation at about 0.5 Mrad. The 100 K Al-hole center loss peak was not observed in this Na-swept MH11-19 sample nor was it observed in a similar but unswept sample. IR measurements indicate that most of the aluminum is converted into Al-OH centers. The 350 K peak and Al-hole production will be discussed later.

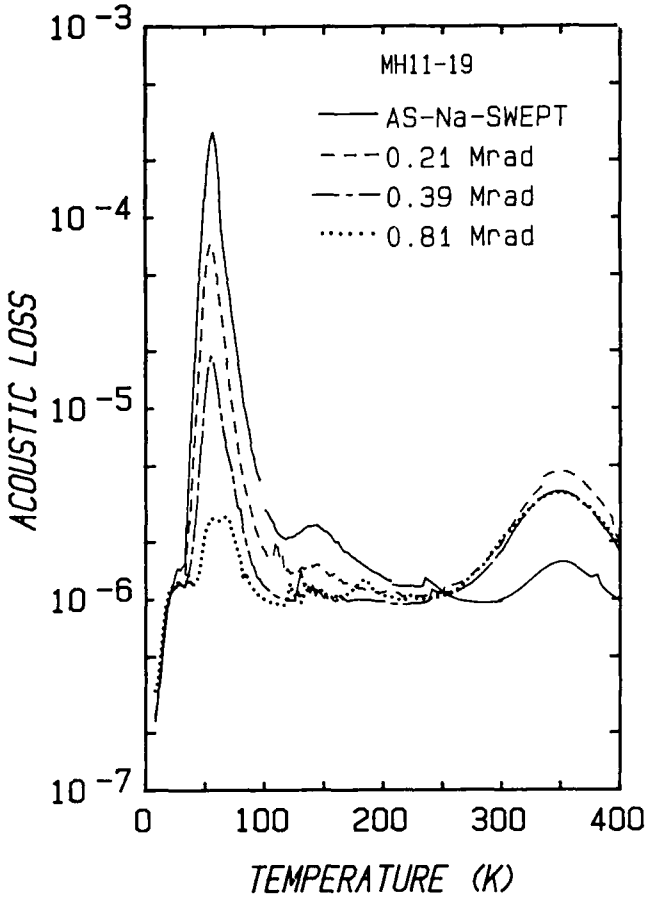



Figure 1. The reduction of the large Al-Na loss peak in a Na-swept 10 MHz 3rd overtone MH11-19 crystal is shown. The Na-related peak at 350 K grows with irradiation out to 0.21 Mrad then decreases slightly.

Figure 2 shows the height of the Al-Na loss peak versus radiation dose for this Na-swept MH11-19 crystal and for Na-swept samples taken from our SQ-B and MH26-39 bars. For all of the samples, measured the Al-Na center shows an initial exponential decay with irradiation dose of the form

$$dL_{Na}(D) = dL_{Nao}\exp(-KD) \qquad (4)$$

where $dL_{Nao}$ is the initial peak height, K is a constant, and D is the radiation dose. The decay of the Al-Na center in the Na-swept SQ-B sample also shows the initial exponential decay with a second exponential term showing up at the larger doses. This second slower decay is also seen in the curves for MH11-19 and MH26-39 but only after about 95 percent of the Al-Na centers have been destroyed. The exponential parameters for these samples are given in Table 1.

Table 1. Al-Na loss peak versus radiation dose parameters.

| Sample | Al-Na(ppm) | $dL_{no}(10^{-6})$ | K(Mrad$^{-1}$) | Comments |
|---|---|---|---|---|
| MH11-19 | 12 | 270 | 7.1 | single |
| SQ-B | 8-10 | 235 | 12.5 | 1st term |
| | | 47 | 2.5 | 2nd term |
| MH26-39 | 2.5 | 63 | 12.1 | single |

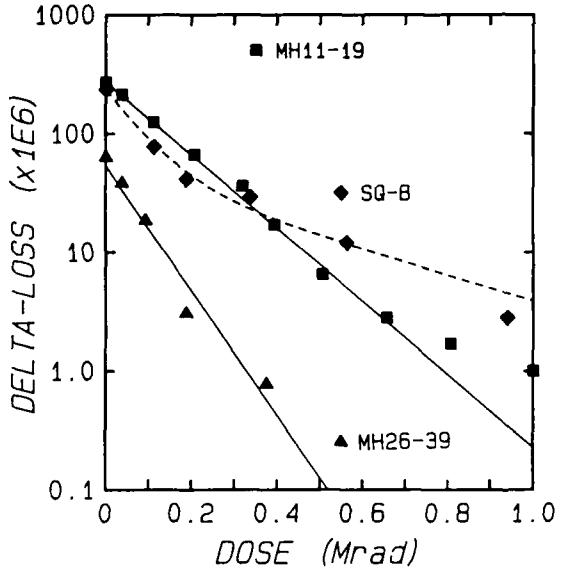



Figure 2. The Al-Na center as measured by the height of the 53 K loss peak is seen to follow an initial exponential decay with radiation dose. As the number of centers is reduced to decay appears to go over to a second slower term. This result is most clearly seen in the curves for the SQ-B sample.

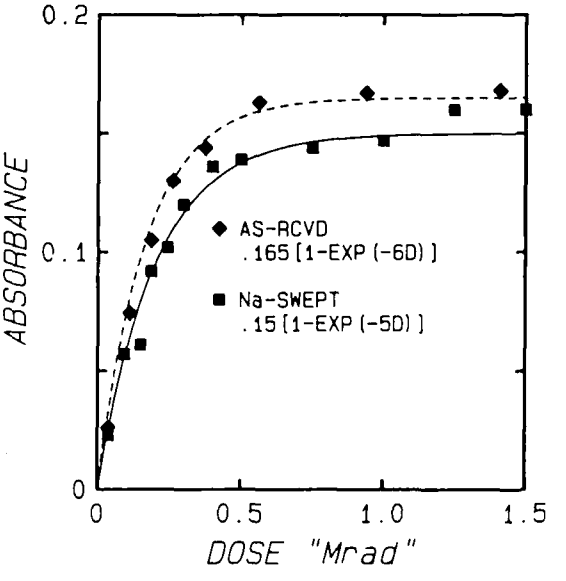



Figure 3. The production of the Al-OH center in an as received MH11-19 and after the sample was Na-swept. Since lithium is the dominant alkali in the unswept sample the curves compare the conversion of Al-Li and Al-Na into Al-OH centers. Eq. 5 describes the lines through the points.

As described in Eq. 3, irradiation at room temperature converts the Al-Na (and Al-Li) centers into a mixture of Al-OH and Al-hole centers. The production of the Al-OH center was measured by tracking the 3367 $cm^{-1}$ absorption band as a function of radiation dose in matching samples. Figure 3 shows the production of the Al-OH center versus radiation dose in an unswept sample from bar MH11-19 and in the same sample after it was Na-swept. We have consistently found that while there may be small dips or "wiggles" in the production curves the general features are described by

$$A(D) = Am[1 - \exp(-K'D)] \qquad (5)$$

where Am is the saturated height (absorbance) of the 3367 $cm^{-1}$ band, K' is a constant and D is the dose in Mrad. No Al-Na 53 K loss peak was observed in the unswept MH11-19 resonator; consequently, lithium must be the alkali ion providing the charge compensation for the aluminum. The differences between the production curves for the unswept and Na-swept samples are very small. This is consistent with the results presented by Hitt and Martin[22] for the release of the alkali from the aluminum site. Table 2 gives the production parameters for a number of samples.

Table 2. Production parameters for the Al-OH center.

| Sample | Al(ppm) | Ao | K'(Mrad$^{-1}$) | comments |
|---|---|---|---|---|
| HA-A | 60-70 | 0.42 | 4 | Li-swept |
| MH11-19 | 12 | 0.165 | 6 | unswept |
| MH11-19 | 12 | 0.15 | 5 | Na-swept |
| SQ-B | 8-10 | 0.7cm$^{-1}$ | 9 | Na-swept* |
| MH26-39 | 2.5 | 0.025 | 15 | unswept |

*The SQ-B data is in terms of the absorption coefficient.

While the constant, K, for the Al-Na (or Al-Li) decay does not perfectly match the growth constant, K', for the Al-OH center they are remarkably close. The two constants also show a trend towards larger values for the low aluminum samples which also have a lower OH content. The early stage decay of the Al-Na centers seems to be governed by hydrogen coming from the growth-defect sites and annihilating the Al-hole center initially produced by the irradiation. If the process works in this manner then room temperature stable Al-hole centers observed after the irradiation should grow at a slower rate. We have carried out an optical absorption experiment on a high aluminum sample, HA-A, where this seems to be the case. A Li-swept 5 MHz 5th overtone blank from bar HA-A was chosen for this test. Low temperature IR and room temperature uv-vis optical absorption scans were used to track the Al-OH and Al-hole center production as a function of gamma dose. Figure 4 shows the IR spectra for the sample after a 5.66 Mrad irradiation and after a 500°C anneal which restored it to the as-Li-swept condition. For comparison, the spectrum for a similar H-swept sample is also shown. The growth-defect centers do not contain enough hydrogen to convert more than about 25 % of the aluminum sites into Al-OH centers. The remaining aluminum was converted into Al-hole centers which cause the smoky coloration seen in irradiated quartz. Figure 5 shows the room temperature uv-vis optical spectra for the Li-swept HA-A sample after the 5.66 Mrad irradiation and after the subsequent anneal. The Al-hole center causes the broad band which peaks near 500 nm.

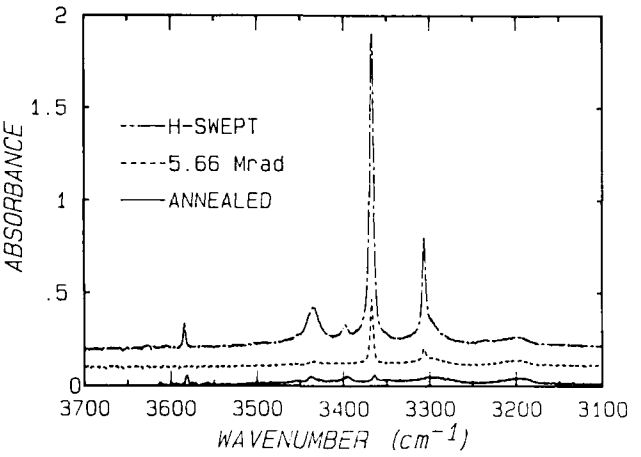



Figure 4. The low temperature IR spectra for the Li-swept HA-A optical sample are shown after the 5.66 Mrad irradiation and after a subsequent anneal. A H-swept sample is also shown.

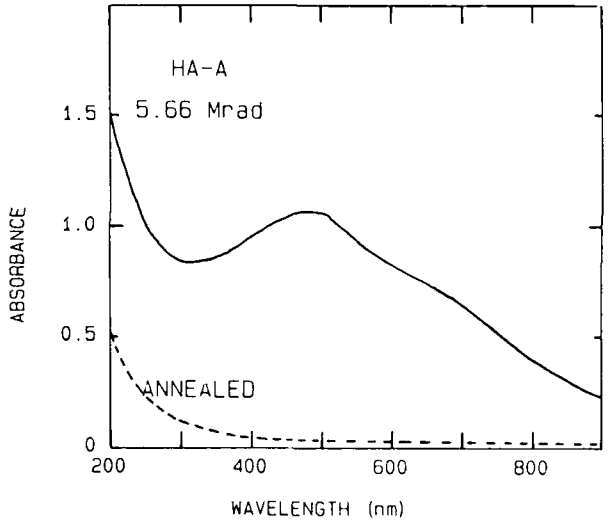



Figure 5. The uv-vis spectra for the HA-A optical sample are shown in the irradiated and annealed condition. The irradiated sample appeared very dark to the naked eye.

The 500 nm Al-hole band and the 3367 $cm^{-1}$ Al-OH band were used to track the production of the two centers. Figure 6 shows the growth curves for both centers. The concentration of Al-OH center shown in Fig. 6 versus radiation dose was estimated by comparing the height of the 3366 $cm^{-1}$ band in the Li-swept sample with that in the H-swept sample and assuming that the H-swept sample had 70 ppm Al-OH centers. The Al-hole center concentration was estimated by assuming that at saturation the aluminum centers were hole compensated. For this sample, the conversion of the Al-Li defects into Al-OH proceeds much faster than the production of stable Al-hole centers. The estimated decay of the Al-Li is also shown in Fig. 6. The 60-70 ppm aluminum content of HA-A quartz is so large that measurements of the 53 K peak in a Na-swept sample are not practical.

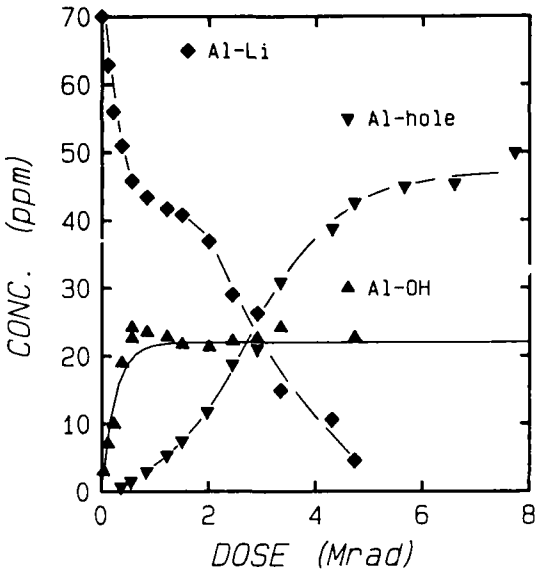



Figure 6. The production of Al-OH and Al-hole centers in HA-A quartz is shown as a function of radiation dose.

The effect of irradiation on the aluminum-alkali site seems to proceed as follows:

1. A radiation induced hole combines with the Al-Na.
2. The sodium drifts away from the combination center forming Al-hole centers.
3. Hydrogen released from the growth-defects annihilates the Al-hole.
4. Step 3 continues until the hydrogen source is exhausted.
5. The remaining aluminum sites stay as room temperature stable Al-hole centers.

As described above, the process is over simplified since the details of the release of the hydrogen from the growth-defect sites are omitted, electron and hole trapping are ignored as well as the effects of other traps. Paige[26] has discussed the kinetics of Al-hole formation in natural quartz; however, he did not consider the transfer of hydrogen to the aluminum site. Euler and Kahan[23] have extended Paige's model to include electron and alkali trapping by a substitutional germanium atom. Neither model seems to be directly applicable to the present situation. McKeever[27] in his review of thermoluminescence discusses a model for trap filling originally due to Maxia[28,29]. They found that the filling of an electron trap could be described as

$$n = N[1 - \exp(-ft/2N)] \qquad (6)$$

where n is the number of filled traps, f is the rate of electron-hole generation (number per unit volume per unit time), and N is the total number density of traps. The quantity ft is proportional to the radiation dose. Our problem is analogous to this electron trapping in that the radiation-induced Al-hole center acts as a trap for the hydrogen released from the growth-defect sites. The dose dependence in Eq. 6 matches our observations. In addition, the constant K increases in samples with low hydrogen content (small N) such as MH26-39 where K=15 as predicted by Eq. 6. The growth of the Al-OH should match the decay of the Al-Na until the available hydrogen is depleted then the decay would be governed by the electron and alkali trapping associated with the stable Al-hole centers.

King[30] first reported a radiation-induced acoustic loss peak at 100 K which he attributed to the Al-hole center. More recently, King and Sander[9] reported that a low temperature irradiation (T=77K) also produced this 100 K peak in swept quartz. This result is consistent with the production characteristics of the Al-hole center. King's as-irradiated curve shows a shoulder or second peak at about 135 K. Martin[21,31] reported that irradiation at room temperature produced a new peak at 23 K in addition to the peaks at 100 K and 135 K. An annealing study showed that all three peaks decayed at the same temperature as the Al-hole center. Figure 7 shows the acoustic loss spectra for an unswept MH11-19 10 MHz 3rd overtone AT-crystal. This sample is a mate to the Na-swept results shown in Fig. 1. The solid curve shows that the Al-Na center loss is completely absent in the as-received spectrum and, therefore, we assume that lithium is the only alkali associated with the aluminum. Irradiation at room temperature produced the additional loss peak near 328 K. This is the lithium version of the defect responsible for the 350 K peak in the Na-swept sample. There may also be a very weak radiation induced peak near 180 K. The room temperature irradiation did not produce observable 23 K and 100 K loss peaks in this sample. This is consistent with the production of very few Al-hole centers because of the availability of hydrogen. The sample was subsequently irradiated at 80 K, cooled to 8 K, then the acoustic loss spectrum shown by the starred curve in Fig. 7 was obtained. The low temperature irradiation produced the additional peak at 23 K and the larger peak at 100 K. Both of these peaks are associated with the presence of the Al-hole center. Both loss peaks disappeared when the sample was warmed to room temperature and then remeasured.

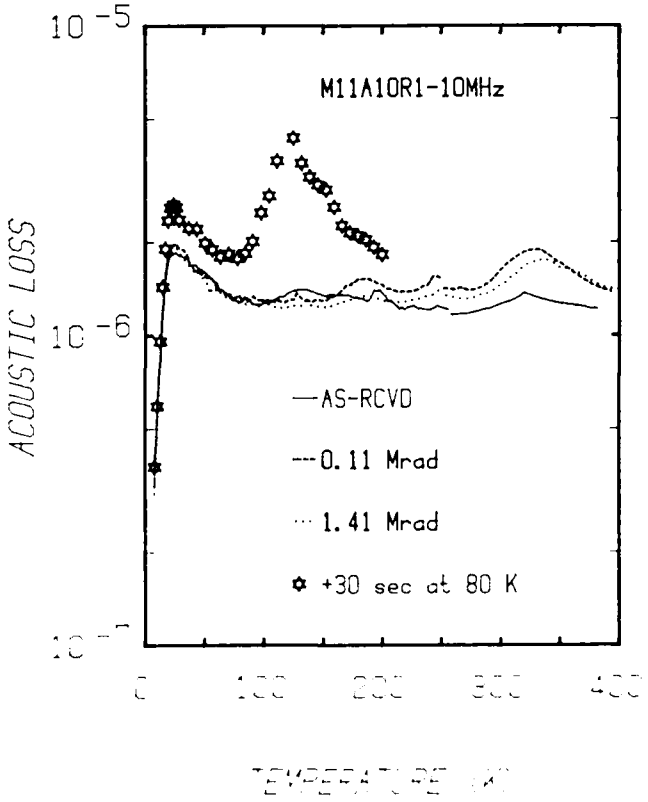



Figure 7. The acoustic loss spectra for unswept crystal MH11-19 are shown in the as-received condition, after two room temperature irradiations, and after a low temperature irradiation. The low temperature irradiation produced the loss peaks at 23 K and 100 K.

A 5 MHz 5th overtone AT-cut resonator was prepared from the 60-70 ppm aluminum content sample, HA-A. This crystal was a mate to the optical sample discussed above. Figure 8 shows the low temperature acoustic loss spectra for this crystal in the as-Li-swept condition and after a 4.73 Mrad irradiation at room temperature. The radiation produced strong loss peaks at 23 K, 100 K and 135 K. The lines through the points represent the calculated additional loss due to the defects assuming that it can be described by the usual anelastic relaxation expression[15].

$$dL = D\,\omega\tau/(1 + \omega^2\tau^2) \qquad (7)$$

where dL is the additional loss, D is the strength of the interaction, $\omega$ is the angular frequency, and $\tau = \tau_o \exp(E/kT)$. $\tau_o$ contains the "jump time" and an entropy factor for the defect, and E is the barrier height. The calculation shows that the 135 K peak is clearly resolved. Table 3 gives the loss peak parameters used for these Al-hole related defects.

The parameters given in Table 3 are only slightly changed from the values reported by Martin[21]. The higher temperature peaks are adequately described by Eq. 7 while the low temperature side of the 23 K peak is not matched very well. Neither moving the peak temperature nor changing the activation energy improved the fit. Ling, Lim, and Nowick[32] reported that irradiation produced a low temperature dielectric loss peak in both PQ-E material and in natural quartz. The shape of their peak varied from sample to sample.

Symmetry arguments show that a given defect can have only two anelastic relaxation peaks in a trigonal crystal such as quartz[33]. Yet all three peaks of the peaks listed above show a connection with the presence of the Al-hole center. Figure 9 shows that the heights of all the three peaks increase linearly with the 500 nm Al-hole center optical absorption for the Li-swept HA-A sample. In PQ-E and SQ-B samples we found that the 23 K peak was larger than the 100 K[21] yet in both the room temperature irradiated HA-A sample and in the low temperature irradiated MH11-19 sample the opposite held. While the Al-hole center almost certainly is responsible for the 100 K peak the cause or the others is uncertain. Our results on HA-A can be used to estimate the concentration dependence of the 100 K loss peak. We find that the height of the 100 K peak is given approximately by

$$dL(10^{-6}) = 0.6C(ppm) \qquad (8)$$

where C is the aluminum hole center concentration in ppm. The Al-hole loss peak at about 100 K shown in Fig. 7 for sample MH11-19 may not be fully saturated according to Eq. 8.

A sodium trap is responsible for the 350 K loss peak shown in Fig. 1 for the 10 MHz Na-swept MH11-19; the same trap but in this case with lithium instead of sodium causes the 328 K peak shown in Fig. 7 for the unswept 10 MHz crystal. The same defects produce loss peaks at 335 K and 305 K in 5 MHz crystals. The production curves shown in Fig. 9 are similar to our earlier observations of the same loss peaks in SQ-B and PQ-E samples. In all cases, the lithium-related loss peak is much weaker than the sodium version. Apparently a few of the traps hold an alkali for the as-Na-swept or as-grown condition. As additional alkalis are released from the aluminum site they are trapped in this defect and the loss grows quickly with radiation dose. As the radiation continues the defect is saturated and then usually shows a decline for larger doses. The 10 MHz 3rd overtone AT-cut MH11-19 crystals also worked very well on their 5th overtone. The anelastic loss expression given in Eq. 7 is

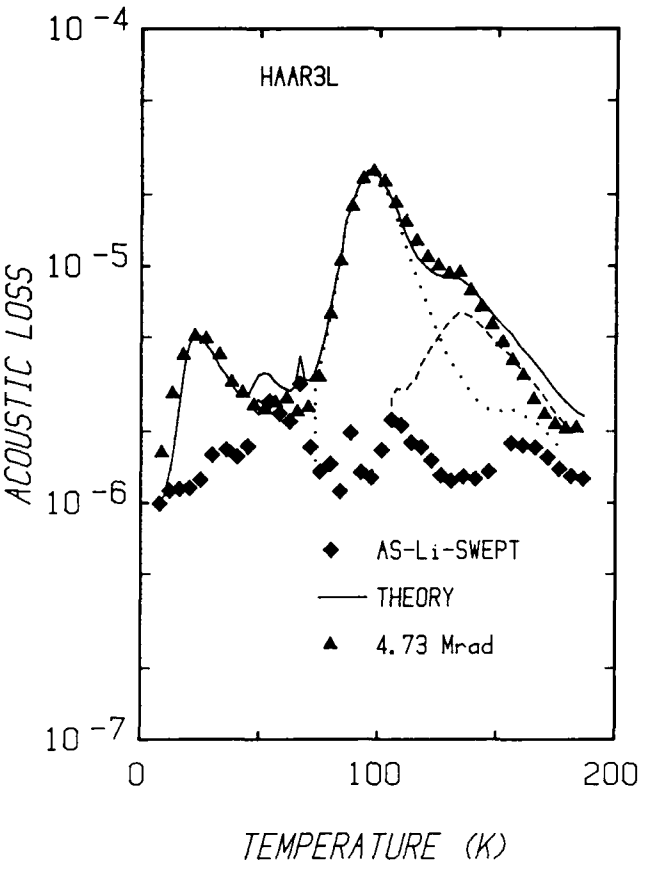



Figure 8. The low temperature acoustic loss spectra for the high aluminum content sample is shown in the as-Li-swept condition and after a 4.66 Mrad irradiation.

Table 3. Maximum temperature at 5 MHz, activation energies, and relaxation times found for the radiation induced loss peaks.

| Defect | $T_{max}$(K) | E(meV) | $\tau_o$(psec) | Method |
|---|---|---|---|---|
| Al-hole-related | 23 | 8 | 565 | fit to loss |
| " | 100 | 90 | 0.61 | " |
| " | 135 | 120 | 1.1 | " |
| Lithium | 305 | 350 | 0.51 | fit to loss |
| | 305 | 335 | 0.15 | 1000/T |
| Sodium | 335 | 350 | 0.15 | fit to loss |
| | 335 | 410 | 0.023 | 1000/T |

maximum when $\omega\tau = 1$. Thus, a plot of frequency versus 1000/T can be used to find the activation energy, E, and $\tau_o$ for the loss peak. The 16.7 MHz 5th overtone data taken on the Na-swept and unswept MH11-19 crystals was combined with the 10 MHz data on the same crystals and our earlier 5 MHz data on other samples to get frequency versus 1000/T plots shown in Fig. 10. The activation energies and $\tau_o$ obtained from Fig. 10 for the sodium and lithium-related defects are given in Table 3. The values obtained by fitting a curve to the loss versus temperature data are also given in Table 3. The activation energies obtained by the two methods are in reasonable agreement. However, the $\tau_o$ values seem short and the shortest values are "non-physical." The relaxation times are very sensitive to the choice of activation energy.

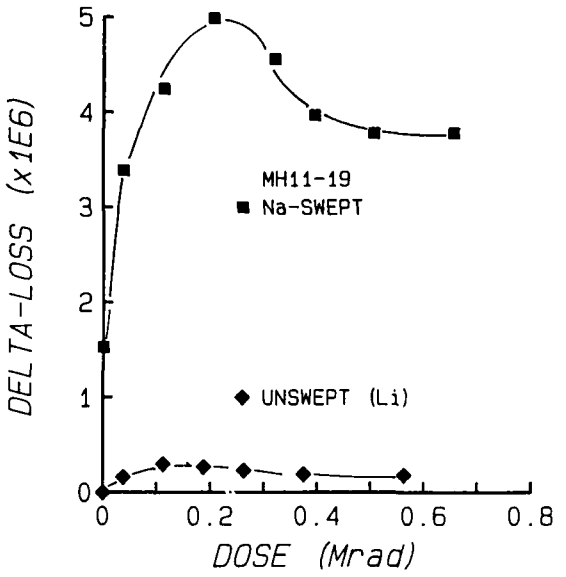



Figure 9. The production curves for the Li-related and Na-related loss peaks in MH11-19 quartz are shown. The Li-related peak is at 328 K while the Na-related peak is at 353 K in these 10 MHz crystals.

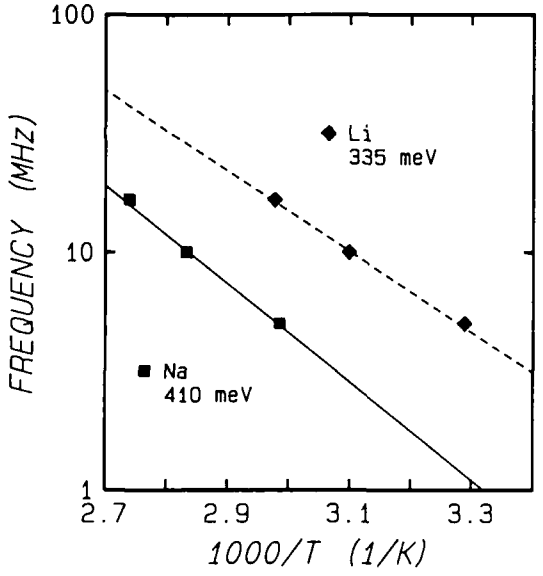



Figure 10. Frequency versus 1000/T curves are plotted for the lithium and sodium related high temperature loss peaks.

Unfortunately, we have not been able to identify the defect responsible for the lithium and sodium loss peaks. Euler and Kahan[23] have found a strong correlation between a lithium related peak that they observed at 250 K (5 MHz) in germanium-doped quartz and the radiation induced behavior of various germanium EPR centers. While we have observed this peak in natural quartz we have never seen it in conventional cultured quartz. Thus, it seems likely that a germanium center is not involved. One possible candidate is the growth-defect centers; when hydrogen leaves this center it should be able to trap an alkali. We have also observed an increase in the growth-defect infrared OH bands upon hydrogen sweeping. This result suggests that some of the growth-defects may initially contain alkalis which are replaced with hydrogen during sweeping. This increase is often accompanied by an increase in the room temperature mechanical Q. Sodium sweeping could introduce sodium into the growth-defects and cause the observed 335 K (5 MHz) peak which we find present prior to irradiation. The radiation behavior of the defect responsible for the 305 K (Li) and 335 K (Na) peaks is not understood. In some samples such as PQ-E the peaks grow rapidly and then decay with radiation dose. While in others, such as MH11-19 their initial growth is the same but they show only a small decrease for large doses.

## CONCLUSIONS

The general features of the destruction of the aluminum-alkali center and the production of the Al-OH center by ionizing radiation follows relatively simple kinetics. The initial decay of the Al-Na (or Al-Li) is exponential in radiation dose and the decay is matched by the growth of the Al-OH center. The Al-OH center grows as 1-exp(-K'D) and this growth evidently controls the initial Al-Na decay. In samples where the hydrogen content is equal to or greater than the aluminum content relatively few room temperature stable Al-hole centers are produced. However, when the hydrogen content is substantially less than the aluminum content a significa fraction of the aluminum is converted into stable Al-hole centers by irradiation. This conversion shows up as a second slower term in the Al-Na decay curve.

Irradiation at room temperature fills alkali traps which cause loss peaks at 305 K for lithium and 335 K for sodium in 5 MHz crystals. A few of these traps are populated in the as-grown condition; however, their population grows quickly with irradiation. In some samples, they go out completely for large radiation doses while in others they show only a small decay. The traps responsible for these loss peaks have not yet been identified.

## ACKNOWLEDGEMENTS

This work was supported by Sandia National Laboratories and by the Solid State Sciences Division, Rome Air Develpment Command, USAF. The authors thank J. F. Balascio of Motorola for supplying the MH series quartz stones. The authors thank A. F. Armington, A. Kahan, F. Euler, L. E. Halliburton, J. C. King, D. R. Koehler, H. Lipson, S. W. S. McKeever, and A. S. Nowick for a number of helpful discussions.

## REFERENCES

1. J. C. King and D. R. Koehler, in *Precision Frequency Control*, E. A. Gerber and A. Ballato, eds., vol. 1. pp. 147, Academic Press, 1985.

2. B. R. Capone, A. Kahan, R. N. Brown, and J. R. Buckmelter, IEEE TRans. Nucl. Sci., *NS-17*, 217 (1970).

3. R. A. Poll and S. L. Ridgway, IEEE Trans. Nucl. Sci., *NS-13*, 130 (1965).

4. T. J. Young, D. R. Koehler, and R. A. Adams, Proc. 32nd Ann. Freq. Control Symposium, EIA, 34 (1976).

5. A. Kahan, F. Euler, H. Lipson, C. Chen and L. E. Halliburton, Proc. 41st Ann. Freq. Control Symposium, IEEE, 216 (1987).

6. P. Pelligrini, F. Euler, A. Kahan, T. M. Flanagan, and T. F. Wrobel, IEEE Trans. Nucl. Sci. *NS-25*, 1267 (1976).

7. J. C. King and H. H. Sander, Radiat. Eff. *26*, 203 (1975).

8. J. C. King and H. H. Sander, IEEE Trans. Nucl. Sci., *NS-19*, 23 (1972).

9. J. C. King and H. H. Sander, IEEE Trans. Nucl. Sci., *NS-20*, 117 (1973).

10. D. R. Koehler and J. J. Martin, J. Appl. Phys. *57*, 5205 (1985).

11. J. R. Norton, J. M. Cloeren, and J. J. Suter, Proc. 38th Ann. Symposium on Freq. Control, IEEE, 63 (1984).

12. J. Suter and R. Maurer, Proc. 40th Ann. Symposium on Freq. Control, IEEE, 134 (1986).

13. T. Flanagan, R. Leadon, and D. Shannon, Proc. 40th Ann. Symposium on Freq. Control, IEEE, 127 (1986).

14. Leslie J. Palkuti and Quang T. Truong, Proc. 38th Ann. Symposium on Freq. Control, IEEE, 55 (1984).

15. D. B. Fraser, *Physical Acoustics*, W. P. Mason, ed. Vol.V. p 2536, Academic Press, NY, 1968.

16. J. Weil, Radiat. Eff. *26*, 261 (1975).

17. J. A. Weil, Phys. Chem. Minerals, *10*, 149 (1984).

18. L. E. Halliburton, J. J. Martin, and D. R. Koehler, in *Precision Frequency Control*, E. A. Gerber and A. Ballato, eds. vol. 1, pp. 1, Academic Press, 1985.

19. A. Kats, Phillips Res. Rpts., *17*, 133 (1962).

20. S. Park and A. S. Nowick, Phys. Stat. Sol. A, *26*, 617 (1979).

21. J. J. Martin, J. Appl. Phys., *56*, 2536 (1984).

22. K. B. Hitt and J. J. Martin, J. Appl. Phys., *54*, 5030 (1983).

23. F. Euler and A. Kahan, Phys. Rev. B, *35*, 4351 (1987).

24. J. J. Martin, Ho B. Hwang, and T. M. Wilson, Proc. 40th Ann. Symposium on Freq. Control, IEEE, 32 (1986).

25. Augusto R. Lopez, J. D. West, and J. J. Martin, Mat. Res. Soc. Symp. Proc., *60*, 451 (1986).

26. E. G. S. Paige, Phil Mag. (Ser. 8), *2*, 864 (1957).

27. S. W. S. McKeever, *Thermoluminescence of Solids*, pp. 118, Cambridge University Press, 1985.

28. V. Maxia, Lett. Nuov. Cim., *4*, 89 (1979).

29. V. Maxia, Phys. Rev. B, *21*, 749 (1980).

30. J. C. King, Bell Syst. Tech. J. *38*, 573 (1959).

31. J. J. Martin, Proc. 38th Ann. Symposium on Freq. Control, IEEE, 18 (1984).

32. S. Ling, B. S. Lim, and A. S. Nowick, Mat. Res. Soc. Symp. Proc.,*60*, 443 (1986).

33. A. S. Nowick and B. S. Berry, *Anelastic Relaxation in Crystalline Solids*, Academic Press, NY, 1972.

# THE THERMOLUMINESCENCE (TSL) OF LITHIUM- AND SODIUM-SWEPT QUARTZ CRYSTALS

A. Halperin and S. Katz
Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

## Summary

The TSL in Li-containing and in Na-swept quartz is described. The glow curves change drastically with the pretreatment of the sample. Thus a double x-irradiation at low temperature and warming to about 250K between the two irradiations produces very strong new TSL peaks at 190K for Li-containing samples, and at 202K for Na-containing ones. Other treatments e.g. prolonged preirradiation at RT cause the 190K and 202K to disappear when other TSL peaks are formed and dominate the glow peaks. The effects are discussed and models for the defects involved and for the processes responsible for the observed effects are suggested. The low temperature phosphorescence accompanying the TSL is also briefly described.

## Introduction

Quartz crystals generally contain aluminum impurities substitutional for silicon in the lattice. Charge compensation is then provided by interstitial monovalent ions, e.g. lithium or sodium ions. The thermoluminescence (TSL) of x-irradiated quartz was found to exhibit a variety of glow peaks. Some of these TSL peaks depended strongly in intensity on the procedure of the x-irradiation and on the history of the crystal treatment, and some were found to be related to the charge compensating monovalent ions. Thus, with lithium, the glow curve obtained under specified procedures of a double irradiation showed a very strong peak near 190K,[1,2] which was shown to be related to the $[SiO_4/Li]^0$ center in quartz[2,3]. Other TSL peaks in quartz were also found to be strongly affected by the applied irradiation procedure, and by the thermal and irradiation history of the crystal. These included peaks at 136, 161, 172 and 181K in lithium containing crystals.[4]

In sodium-swept crystals a peak analogous to the 190K peak was observed at 202K[5] and TSL peaks at 138, 160 and 1789K seem to be analogous to the TSL at nearly the same temperatures in the Li-containing crystals. There are, however, remarkable differences between the Li- and Na-containing samples. These include some TSL peaks and the low temperature (~20K) phosphorescence which are much stronger in the Na-swept samples compared to the Li-swept ones.

In the present work we compare the TSL and the low temperature phosphorescence of Li- and Na-containing quartz stressing the differences between the two. Suggestions are also made on the nature of the centers involved in the emission of the various TSL peaks and the low temperature phosphoresence.

## Experimental

Samples were cut from Y-plates of a Sawyer synthetic electronic grade quartz crystal. Each sample was 11 x 8 x 2 $mm^3$ in dimensions with the main faces perpendicular to the Y-axis, and was cut either from the +x-growth section (X-sample) or from pure +z-growth section(Z-sample). The aluminum content in the +X-samples was about 30-40 ppm per silicon atom, and that in the +Z-samples was lower by about one order of magnitude. The unswept samples contained practically only lithium as charge compensator. The lithium was replaced by sodium by electrical sweeping at 530°C between two graphite electrodes pasted with NaCl salt. An electric field of 400V/cm was applied along the c-axis of the sample and the sweeping was continued until a charge of 4-5 coul/$cm^3$ was passed through the sample. This was found to replace practically all the lithium by sodium.

Samples were polished to nearly optical flatness, then cleaned and annealed for about 1 hour at 530°C, thus restoring the crystal to nearly virgin conditions. An Air-Products Displex double-stage refrigeration system enabled measurements over the temperature range of about 10-370K. The crystal was provided with Spectrosil fused silica windows, a thin (0.3mm) aluminum window for the x-raying, a heater, a platinum sensor for the precision temperature controller and a Chromel-gold + 0.07% iron thermocouple. More details on the experimental arrangements were described elsewhere[6].

The heating rate used when measuring the total undispersed TSL emission was 10°/min. TSL emission spectra were scanned in about 10 sec covering the range 300-700 nm. The heating rate in the latter measurements was 2.5°/min.

## Results

Figure 1 shows glow curves obtained for nearly virgin lithium containing quartz samples cut from the +x-growth zone (Curve X) and from the +z-growth zone (Curve Z). The two curves do not differ much from each other below 150K. The behavior is, however,

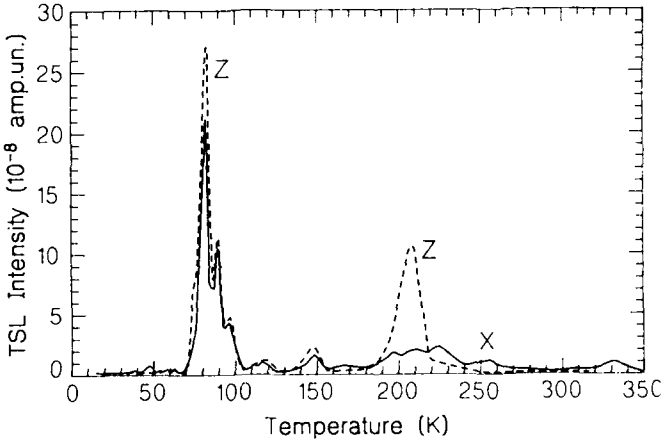



Figure 1. Glow curves for a nearly virgin Li-containing quartz crystal. Z - a z-growth sample; X - a +x-growth zone sample. Each curve obtained after 30 min x-irradiation at 10K.

different in the range 150-300K, where the glow curve of the X-sample shows a variety of peaks, while the curve of the Z-sample shows mainly one peak at 206K. The curves in Figure 1 were obtained by "regular" treatment, which means by cooling the crystal

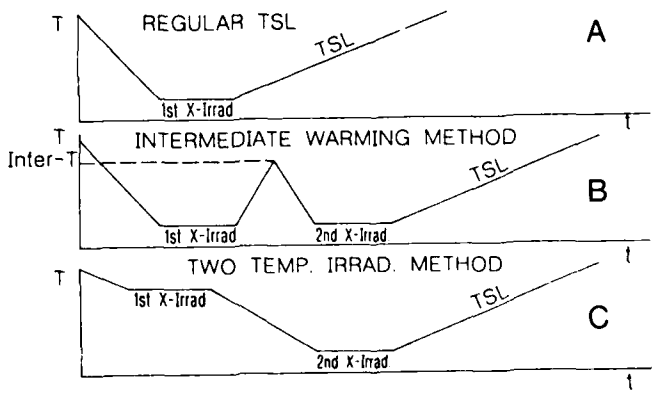



Figure 2. A schematic diagram of the various procedures in the TSL measurements.

CH2588-2/88/0000-184 $1.00 C 1988 IEEE

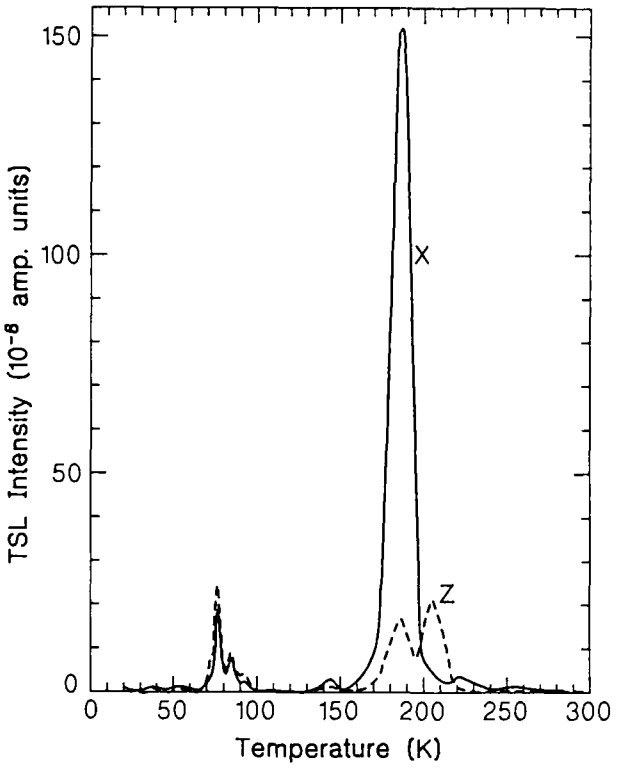



Figure 3. Glow curves for a Li-containing Z-sample (Curve Z) and a X-sample (Curve X) obtained by 30 min of x-irradiation at 10K followed by warming to 220K and a second 30 min x-irradiation at 10K.

to a low temperature (10-20K in the present work), and x-irradiation at this low temperature after which the TSL was measured during heating (Figure 2A). However, completely different glow curves were obtained after double irradiation excitations. These included the methods of "Intermediate warming" (Figure 2B) and the "Two temperature irradiation (Figure 2C). The resulting TSL is described below:

(a) Formation of the glow peaks related to the $[SiO_4/M]^0$ centers (M=Li, or Na). Figure 3 shows for example the glow curves for an X-sample (Curve X) and a Z-sample (Curve Z) of a Li-containing crystal obtained after a first x-irradiation at 10K, followed by intermediate warming to 220K and a second irradiation at 10K after which the glow curve was taken during warming (Intermediate warming method, Figure 2B). The dominating TSL in Curve X is now a peak near 190K which was absent in the regular TSL (Figure 1). Under certain conditions of the double irradiation the 190K peak was stronger by two orders of magnitude or more compared to any other peak in the glow curve. This 190K peak was found to be directly related to the $[SiO_4/Li]^0$ center[2,3]. The Z-sample also shows the 190K peak (Figure 3, Curve Z) but weaker by about one order of magnitude compared to that in Curve X, which is due to the lower concentration of the aluminium (and lithium) impurities in the Z-sample. More details on the behavior of the 190K TSL peak were described elsewhere[2-6]. We shall only stress here that almost all the TSL peaks in the Li-containing samples in the range 150-300K including the one at 190K emitted a spectral band peaking at 3.26 eV(380nm). Emission spectra for some of the glow peaks are shown in Figure 4.

After electrical sweeping of sodium into the crystal the 190K peak was replaced by one at 202K. This is shown in Figure 5, in which Curve 1 gives the glow curve in the Na-swept, x-growth zone sample as obtained by a single x-irradiation (5 min. at 20K). Curve 2 gives the same but after two x-irradiations (each for 5 min at 20K) with an intermediate warming to 260K between the two irradiations.

There are two remarkable differences between Curve 1 and Curve 2: a TSL peak at 202K which was absent in Curve 1 is now dominating the glow curve (It is shown on a reduced, x0.02 scale).

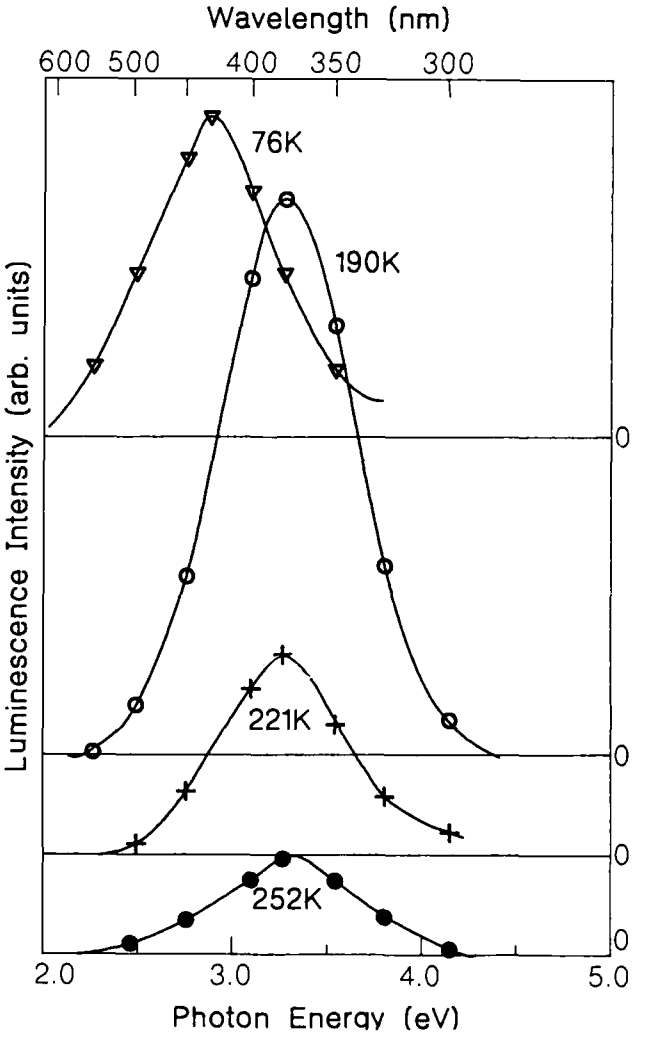



Figure 4. Emission spectra for some TSL peaks of a Li containing X sample. Peak temperatures are indicated in the figure.

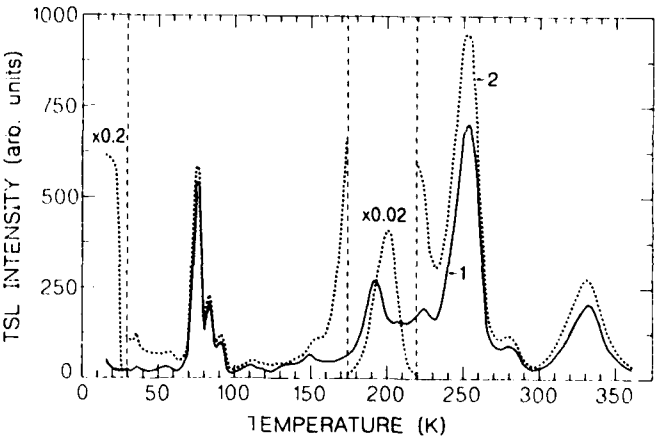



Figure 5. Glow curves for a Na-swept, quartz crystal. 1- After a single 5 min. x-irradiation at 20 K. 2- After a double x-irradiation (5 min each at 20K) and warming to 260K between the two irradiations.

In addition, Curve 2 shows a very strong phosphorescence at 20K (given on a x0.2 scale) which falls abruptly to almost zero near 25K. This phosphorescence is formed along with the formation of the 202K peak and seems to be closely related to it.

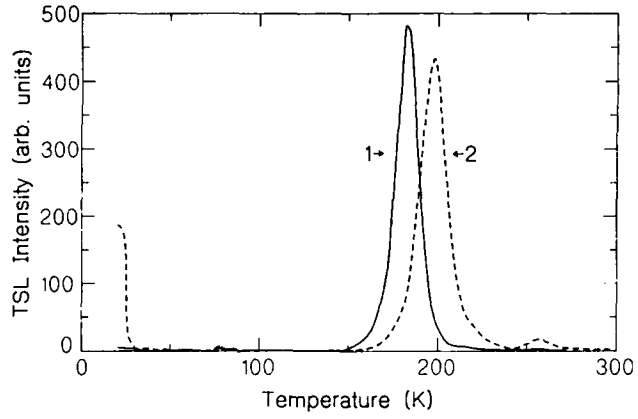



Figure 6. Comparison of the glow curves obtained by the "Intermediate warming method" (see text). 1 - for the Li containing X-sample; 2 - Same for the Na-swept sample.

Figure 6 shows for comparison the glow curves obtained by the Intermediate warming method for the Li-containing sample (Curve 1) and for the Na-swept sample (Curve 2). The 190K peak is shown to dominate and all the other glow peaks can hardly be seen on the given ordinate scale. The same is true for the Na-swept sample (Curve 2), where the 202K peak is almosts equal in intensity to the 190K peak in Curve 1. Remarkable in Curve 2 is also the high intensity of the phosphorescence at 20K, which on the given scale

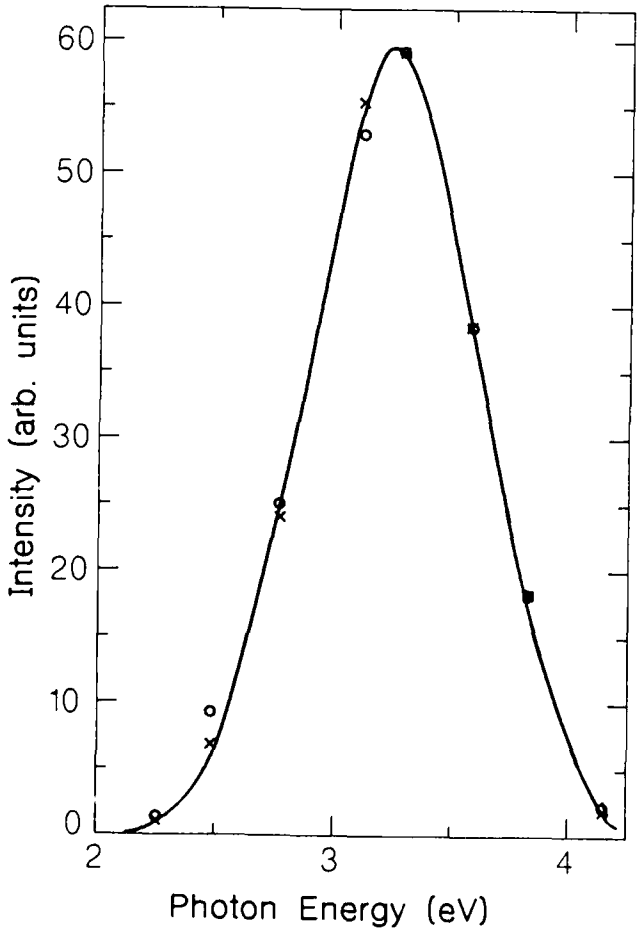



Figure 7. Emission spectrum for a Na-swept sample. x - 202K TSL peak; 0 - phosphorescence at 20K (normalized at maximum).

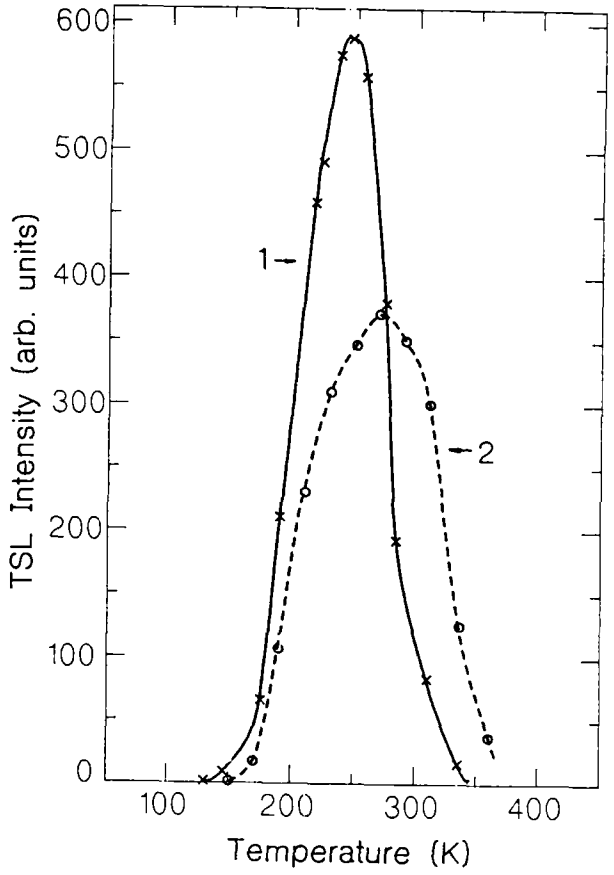



Figure 8. Formation curves for the Li containing X-sample (1); and same after Na-sweeping (2). Both curves obtained by the "Intermediate warming method".

can hardly be seen in Curve 1. Both the 20K phosphorescence and the 202K glow peak emit the 3.26 ev(380nm) band as shown in Figure 7, where the points on the curve are for the emission at the 202K peak (x) and for the 20K phosphorescence (o), normalized at the maximum. We note that the emission is just the same as observed for the 190K peak of the Li-containing sample (Figure 4).

Figure 8 gives the "formation curves" for the 190K peak of the Li-containing sample (Curve 1) and for the 202K peak of the Na-swept one. The curves give the intensity of the relevant peak as function of the intermediate temperature reached between the two x-irradiations. One notices that the maximum intensity of the 190K peak is obtained with an intermediate warming to about 230K, and that for the 202K peak appears at 270K.

The remarkable analogy in behavior between the 190K and the 202K peaks illustrated above suggests that the 202K peak is related to the $[SiO_4/Na]^0$ center in the same way as the 190K peak is realted to the $[SiO_4/Li]^0$ center. The following scheme describes the processes involved in the formation of the $[SiO_4/M]^0$ centers (M = Li or Na):

1st step - free holes made available by the first low temperature x-irradiation get captured at the $[AlO_4/M]^0$ centers:

$$[AlO_4/M]^0 + h^+ \rightarrow [AlO_4/M,h]^+$$

2nd step - Warming to an intermediate temperature of 200-300K supplies enough thermal energy to the $M^+$ ions, which leave the aluminum centers and migrate away along the x-axis channels in the quartz lattice to a regular $SiO_4$ site in the lattice:

$$[AlO_4/M,h]^+ \text{ warming} \rightarrow [AlO_4/h]^0 + [SiO_4/M]^+$$

3rd step - Free electrons created during the second low temperature x-irradiation get captured by the $[SiO_4/M]^+$ centers:

$$[SiO_4/M]^+ + e^- \rightarrow [SiO_4/M]^0$$

4th step - On warming to 190K (for Li) or to 202K (for Na) electrons at the $[SiO_4/M]^0$ centers are released to the conduction band and on recombination with holes at the luminescence centers emit the 190 or 202K peaks.

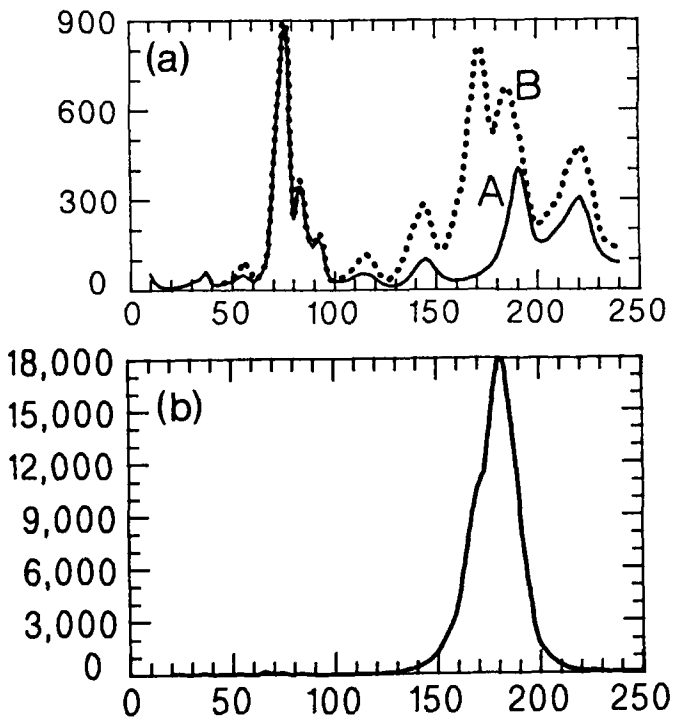



Figure 9. (a) - A glow curve for a nearly virgin Li containing X-sample (A), and the same after a double x-irradiation, at 20K each with warming to 360K between the two irradiations. (b) - as (a) but after 7 hours of RT x-irradiation followed by 5 min. of x-irradiation at 20K.

(b) Other TSL peaks formed in Li- or Na-containing quartz by special procedures of irradiation and of thermal treatment. TSL peaks at 136, 161, 172 and 181K in Li-containing samples were mentioned above as affected strongly by various treatments. In fact they are formed to some extent by just the same method as the 190K peak, which means, by a double low temperature irradiation with an intermediate warming. However, they can not be observed under these conditions in the presence of the enormously strong 190K peak. One has therefore to eliminate the 190K peak. This was achieved by choosing the temperature of 360K for the intermediate warming between the two x-irradiations (at 20K). The results are shown in Figure 9(a), where Curve A gives the glow curve for the nearly virgin crystal and Curve B gives the same after the double irradiation (and warming to 360K between the irradiations). One can see the formed 172K and a weaker 181K peak. Much higher intensities, especially of the 181K peak can be obtained by a different treatment, namely a prolonged x-irradiation at RT followed by a second x-irradiation at 20K (essentially as in Figure 2C). The results of 7 hours of RT x-irradiation followed by 5 min. of x-irradiation at 20K is shown in Figure 9(b). An extremely intense peak appears now at 181K (compare scales in Figures 9(a) and 9(b)). The TSL at 136K can hardly be seen on this ordinate scale and the 161K peak is seen as a shoulder to the one at 181K. More details on these peaks were given elsewhere[5].

The analogous effect in a Na-containing X-sample is shown in Figure 10. In this figure Curve 1 gives the "regular" glow curve for the nearly virgin sample and Curve 2 gives the same but after exposure of the sample to 5 hours of x-irradiation at room temperature. One can see that peaks in the range 90-200K were strongly enhanced by room temperature irradiation, with the peak at 178K as the strongest. Of interest is that the TSL at 93K which was not affected by room temperature irradiation in the Li-containing

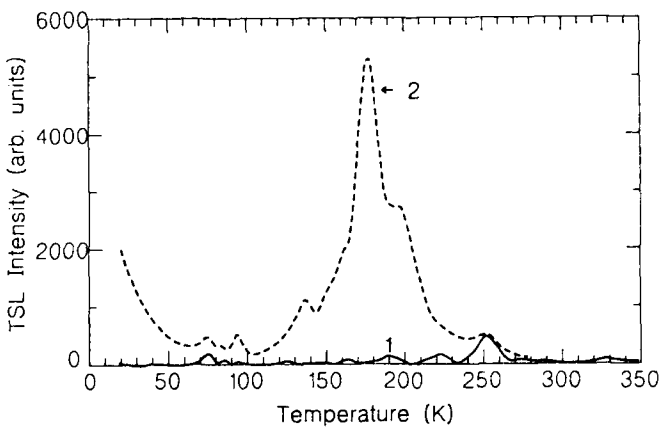



Figure 10. Glow curves for a Na-swept X-sample.
1 - After a single x-irradiation (5 min. at 20K). 2 - After 5 hours of RT x-irradiation followed by 5 min. of x-irradiation at 20K.

sample (see Figure 9(a) and (b)) was enhanced by an order of magnitude by the same treatment in the Na-containing sample. The low temperature phosphorescence was also enhanced by the treatment. This, however, is different from the phosphorescence formed along with the 202K TSL peak (Figure 5); it does not fall abruptly to zero at 25K, but decays monotonously and can still be observed even above 100K. This phosphorescence is formed and grows in intensity with the 178K TSL peak.

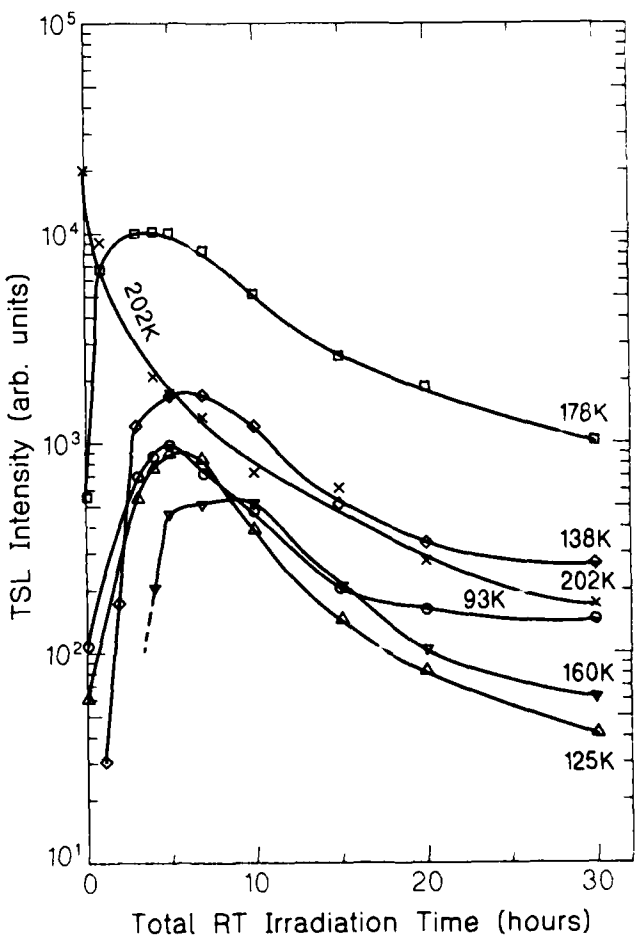



Figure 11. Intensity of some TSL peaks of a Na-swept X sample as function of the time of preirradiation at RT.

The growth in intensity of some of the TSL peaks with the time of the room temperature x-irradiation in the Na-containing sample is shown in Figure 11. The 202K peak is seen to fall down sharply when the other peaks grow up to exposures of some 5-6 hours. At longer exposures almost all the TSL peaks come down in intensity in a way similar to that observed in Li-containing samples[5].

## Discussion and Conclusions

The TSL in the Li-containing quartz crystals in the range 50-110K was found to change very little with the various thermal and irradiation treatments. The peaks at 76 and 83K were found to emit at 2.88 nm and those at 93 and 109K emitted at 2.70 ev (460 nm).[6] It was concluded that at these peaks holes are released from various trapping levels and recombine with electrons still trapped in the relevant centers. This notion gained support from measurements of thermally stimulated currents[7]. The Na-swept crystals exhibited just the same TSL peaks in the above temperature range; the peaks appeared at exactly the same temperatures and most of them remained almost unaffected by the various treatments. The emission, at least of the 76 and 83K peaks was the same 2.88 eV band. We conclude therefore that here again holes are released and recombine with electrons still trapped in the centers. Somewhat surprising is the behavor of the 93K peak. The peak at this temperature was only weakly affected by the various treatments in the Li-containing samples (see Figure 9). Its intensity however, increased by an order of magnitude or more after preirradiation at RT in the Na-swept samples (Figures 10 and 11). The reason for this difference in behavior is still unclear.

The full similarity in formation and behavior of the Li-related 190K peak and the Na-related 202K peak suggests strongly that the latter peak is related to the $[SiO_4Na]^0$ center just in the same way as the 190K peak is related to the $[SiO_4/Li]^0$ center; in both the emission is obtained when electrons in the above centers are released and recombine at luminescence centers. The fact that both the TSL peaks emit the same band, namely the 380 nm band, implies that in both cases recombination takes place at the same luminescence center. It is not simple to identify this recombination center. The identical emission in both cases indicates that the recombination center does not contain the relevant monovalent ion. McKeever et al [8], when investigating the TSL emission in quartz at 100°C, suggested that the 380 nm emission is not obtained by recombination at $[AlO_4]^0$ centers (aluminum-hole centers), but rather at some other unidentified hole trap. The 190 and 202K peaks also emit at 380 nm. Our measurements however, show that after the emission of these peaks the crystal is restored to its original state with the $M^+$ ionsback at the aluminum impurities. It seems therefore very likely that the hole traps at which recombination takes place in our case is the $[AlO_4]^0$ center. Still, the possibility that restoration takes place in a cascade of more than one step can not be ruled out.

The TSL peaks at 136, 161, 172 and 181K (further referred to as the "181K group") in the Li-containing samples are formed essentially by a double irradiation procedure though under conditions different from those needed for the formation of the 190K peak. The most effective method for the formation of the "181K group" was by prolonged preirradiation at RT, by which practically all the Li+ ions get trapped at certain defects, and so are not available any more for the formation of the $[SiO_4/Li]^0$ centers.[2] The reduction in the intensity of the 190K peak related to the $[SiO_4/Li]^0$ centers, was found to be accompanied by growth in intensity of the "181K group" (see Ref. 4, Figure 2). This fact and the similarity in the way of formation of this group and the 190K peak suggest that the peaks in the "181K group" are related to $[X_i,Li]^0$ centers, where $X_i$ are the unidentified defect centers at which the Li ions get trapped during the RT x-irradiation. The fact that the emission of the "181K group" is at 380nm, implies that here again electrons are released and recombine at the same hole centers as in the case of the 190K peak.

The Na-swept samples exhibited a group of TSL peaks formed in a way similar to that of the "181K group". In the Na-containing samples there are at least 5 peaks strongly affected by prolonged RT x-irradiation. These are the peaks at 94, 125, 138, 160 and 178K (see Figure 11) with the 178 K peak as the strongest in the group. We shall refer to this group as the "178K group". All the peaks in this group grow with the time of RT irradiation on the account of the 202K peak which decreases with the RT irradiation (Figure 11). The strong 178K peak seems to be analogous to the 181K which was the strongest in the "181K group". The present results, however, do not enable a clear one to one fitting between the other peaks in the two groups. Still, the similarity in formation and behavior suggests that the processes involved are similar; here the peaks in the "178K group" are emitted when electrons released from $[X_i, Na]^0$ centers recombine with holes still trapped at the luminescence centers. The identical emission band (380 nm) suggests that recombination takes place at the same recombination centers as in the case of the Li-containing samples.

The peaks in the "178K group" do not seem to be shifted towards higher temperatures compared to those in the "181K group". Things were different with the pair of 190 and 202K peaks, where the Na-related peak appeared at a temperature 12K higher compared to its Li-analogue. The lack of a considerable shift between the two groups of the other peaks means that the binding energies of the Li-ions at the $X_i$ defects are nearly equal to those of the Na-ions. Identification of the various $X_i$ defects should provide an explanation for this behavior.

Radiation is well known to affect the frequency and the Q-factor of quartz resonators. An investigation combining the radiation affected TSL peaks and the effects on the frequency and quality of quartz resonators may provide a better understanding of the effects of radiation on resonators. It may also help to provide a better control of the radiation effects on quartz resonators.

The low temperature phosphorescence was also found to emit the 380nm band, both in Li and Na containing samples. More than that, the formation curves for the low temperature phosphorescence were similar to those for the 190 and 202K TSL peaks for Li and Na containing samples respectively. In addition, under the conditions at which the "181K" and "178K" groups were formed the low temperature phosphorescence was found to be formed along with the related group of TSL. A full discussion of the low temperature phosporescence is beyond the scope of the present work. We only give here a few points:

1. The phosphorescence of the Na-containing samples is much stronger compared to that in the Li-containing ones.
2. In the Na-containing samples the phosporescence related to the 202K peak dropped sharply to zero near 25K, while that related to the "178K group" showed only a slow decay with temperature (practically with time) up to at least 100K.
3. Both with Li and with Na the phosphorescence emitted at 380 nm, and so the emission must be due to conduction band electrons recombining with holes at the same luminescence center as in the case of the TSL emission.
4. The exact model for the process and the transitions involved in the emission of the low temperature phosphorescence has still to be worked out.

## References


1. A. Halperin and S. Katz; J. Lum. 31 &32, 129 (1984).
2. A. Halperin, M.G. Jani and L.E. Halliburton; Phys. Rev. B34, 5702 (1986).
3. M.G. Jani; L.E. Halliburton and A. Halperin; Phys. Rev. Lett. 56, 1392 (1986).
4. A. Halperin and S. Katz; J. Phys. Chem. Solids (in press).
5. A. Halperin and S. Katz; Solid State Comm. 63, 697 (1987).
6. S. Katz and A. Halperin; J. Lum. 39, 137 (1988).
7. S. Katz and A. Halperin; Phys. Rev. B36, 6646 (1987).
8. S.W.S. McKeever, C.Y. Chen and L.E. Halliburton; Nucl. Tracks, 10, 489 (1985).

42nd Annual Frequency Control Symposium - 1988

# NEW TECHNOLOGY FOR DETECTION AND REMOVAL OF SURFACE CONTAMINATION INVOLVING PARTICULATES OR WATER/ORGANIC MATERIALS

Professor. S.A. Hoenig, Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ 85721

## Introduction

Contamination control is critical in the manufacturing and packaging of quartz oscillators used for frequency control. As part of the production process it is necessary to detect the presence of contaminants that can affect the operation of the device. If contamination is present it must be removed without damaging the resonator, system or any auxiliary assemblies.

## Contamination Detection

Contaminants may be divided into two catagories:

### 1) Condensed Organic/Water Vapors

These materials tend to form layers on the surface. In the past various types of water break tests were used for detection, but it has been demonstrated that systems of this type may yield very poor results with hydrophylic organics.[1] In any case the test involves a human operator and, as such, is difficult to do in production operations. Auger and/or infrared (IR) systems can be used to detect organics, but the Auger unit is costly and requires a high vacuum. IR can be done in an air environment but requires a transparent substrate.

We have investigated the application of ultraviolet (UV) light induced photoemission for monitoring surface contamination. It is well known that photoemission is very sensitive to surface condition and contamination.[2] The apparatus can operate in air at one atmosphere. In this case the electrons are immediately attached to oxygen or water molecules to form negative molecular ions. It is this ion current that is detected in our apparatus.

In figure 1 we show a system developed to test for residual contamination in liquids (eg. water, organics or acids). A small sample (10 to 50 microliters) is placed on a suitable surface and allowed to evaporate. If evaporation leaves no residue, the photoelectric emission from the surface will be unchanged from the value observed before the liquid was deposited.

Some results with various water samples are shown in figure 2. Results with acids are presented in figure 3. Data with alcohol and water residues are shown in figure 4. In other tests we have established that the sensitivity of the technique for liquid contaminants is at the PPB level.

The same unit can be used to follow the build up of organic/water contamination on a clean surface exposed to a laboratory environment, figure 5. If we assume that the thickness of the surface layer, after 75 minutes, is about 100 Angstroms it is clear that contamination at the Angstrom or submonolayer level is observable.

The same system can be used to monitor the clean up of surfaces with UV light. This clean-up technology was first reported by Vig[3], but users have had problems knowing when the clean-up process was complete. In figure 6 we show how the photoelectric system can be used for process monitoring. It would seem that a device of this type might well be an add-on to the commercial UV cleaning units.

### 2) Particulates

In the semiconductor industry particle detection is usually done with optical scattering systems based on a He/Ne laser system. The technology has been discussed in some detail and is quite effective for particles larger than 0.3 micrometers.[4] The test surface must be smooth and suitable for insertion into the laser scattering system. This makes it difficult to test components that have been hard wired into a support structure.

We have investigated the use of UV for this application with the idea that small particles will generate a strong photoemission current, until they charge up to the point that electron emission ceases. There is evidence that micron/submicron particles will behave differently from the bulk in terms of having a smaller photoelectric work function.[5] If this is the case we might expect that the charged particles will generate a nonuniform electrostatic field that can be detected by suitable apparatus scanning over the surface.

In figure 7 we show some results with small particles on a various surfaces. Our program was limited by funding considerations, but we were able to show that 10 micron particles could be detected quite easily. Detection of submicron particles appears to be quite feasible. The apparatus is simple and might well be suitable for on line examination of quartz crystals during various parts of a processing cycle.

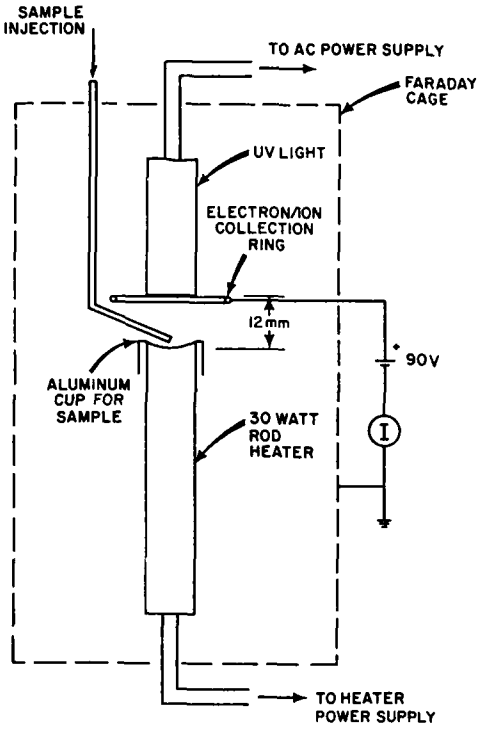
## Cleaning Technologies

### 1) Dry Ice Snow

It is well known that blowing gases over a surface will not remove small (eg 1 micron) particles. Liquids are more effective, provided that high pressures or brushes are used to move the particles. Both of these techniques can damage delicate surfaces and we felt that other systems were needed.

The dry ice technology is shown schematically in figure 8. Liquid $CO_2$ is drawn up into a tube and allowed to expand to form dry ice flakes. The flakes can be large and soft, for cleaning gold surfaces, or small and hard for more aggressive cleaning (eg. sandblasting).[6] Some results with this system are shown in figures 9, 10 and 11. The technology is in regular use by a number of observatories (for telescope mirrors) and

CH2588-2/88/0000-189 $1.00 C 1988 IEEE




ULTRAVIOLET LIGHT AND HEATER SYSTEM FOR DETECTION OF CONTAMINANTS IN LIQUIDS

Figure 1

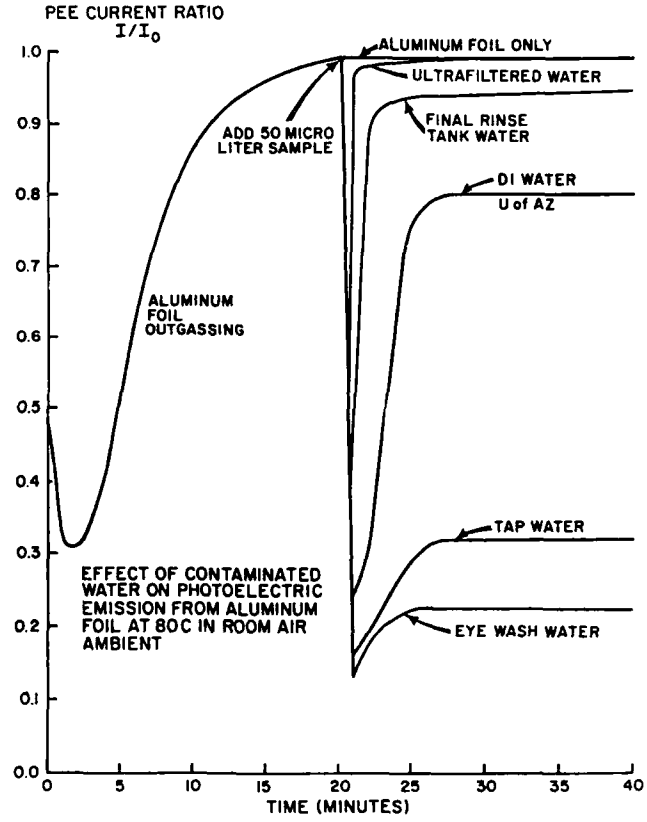
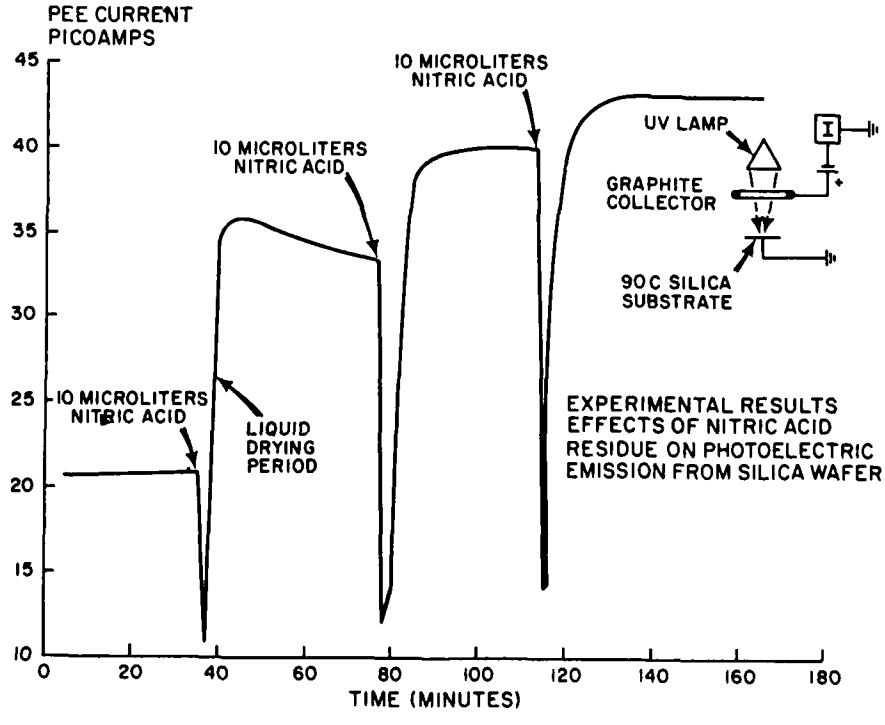



Figure 2

Figure 3



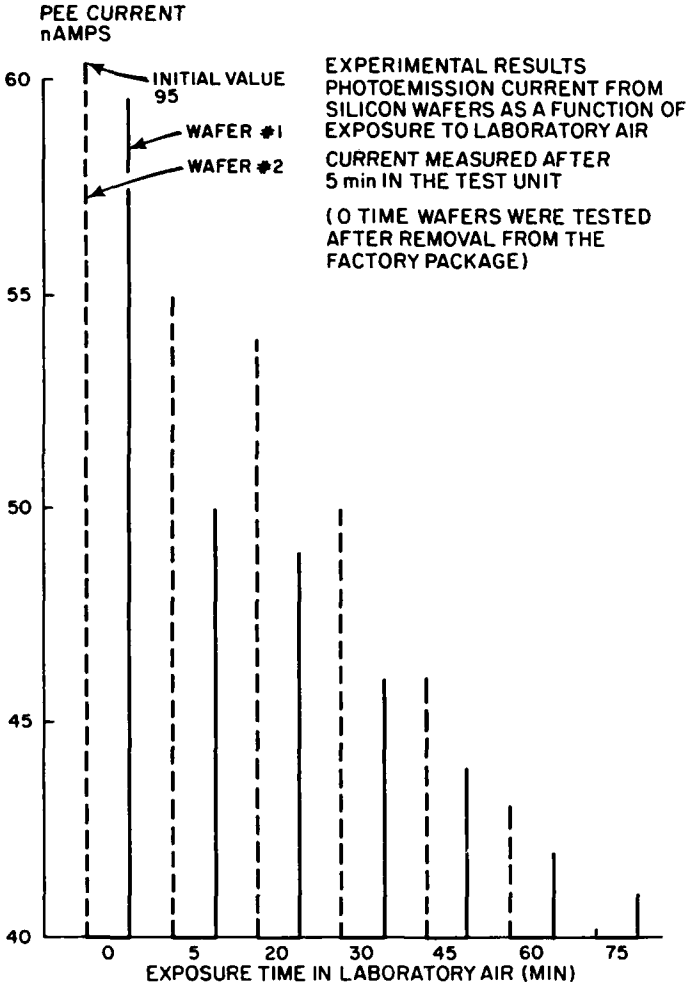



Figure 5

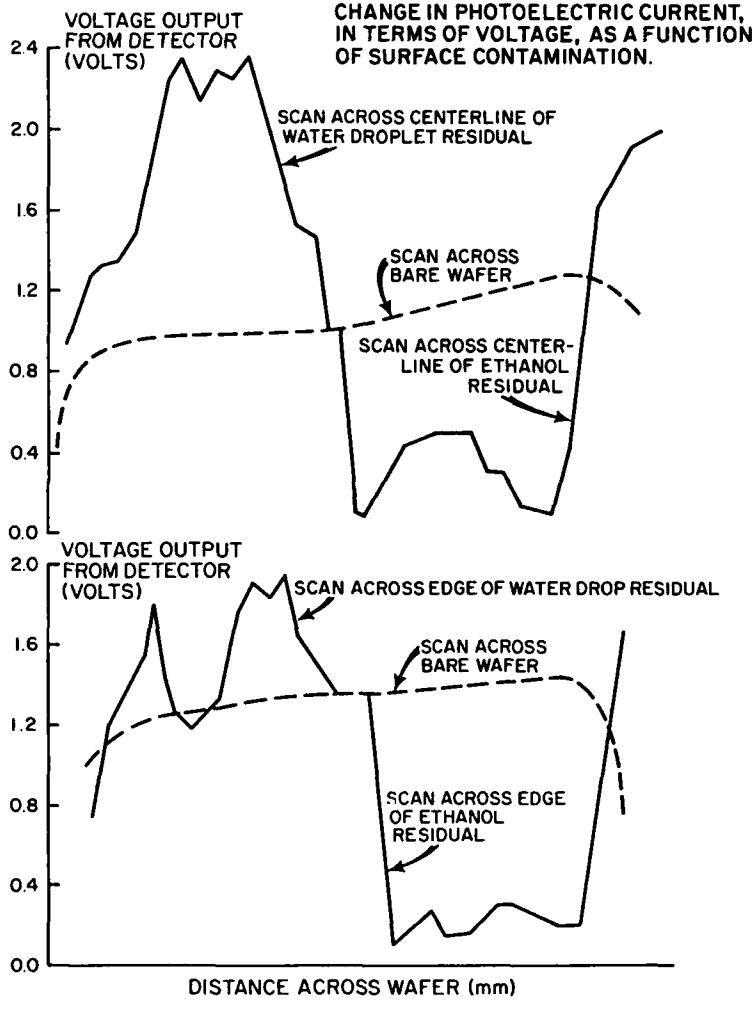
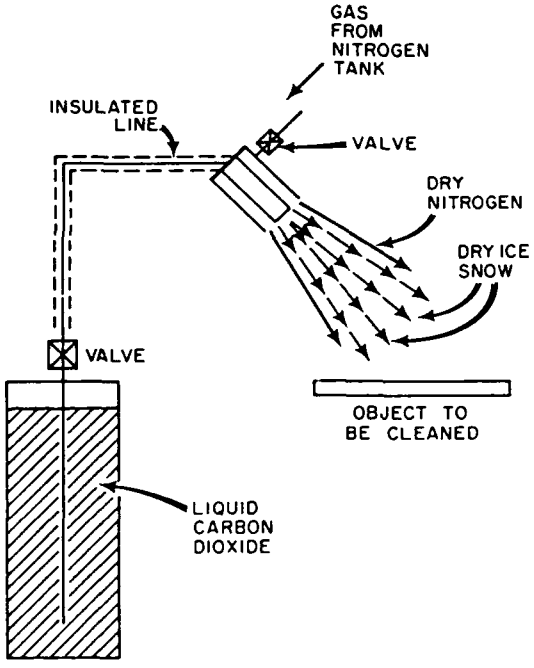



Figure 4




SCHEMATIC DRAWING DRY ICE SNOW CLEANING TECHNOLOGY (SCALE NONE)

Figure 8

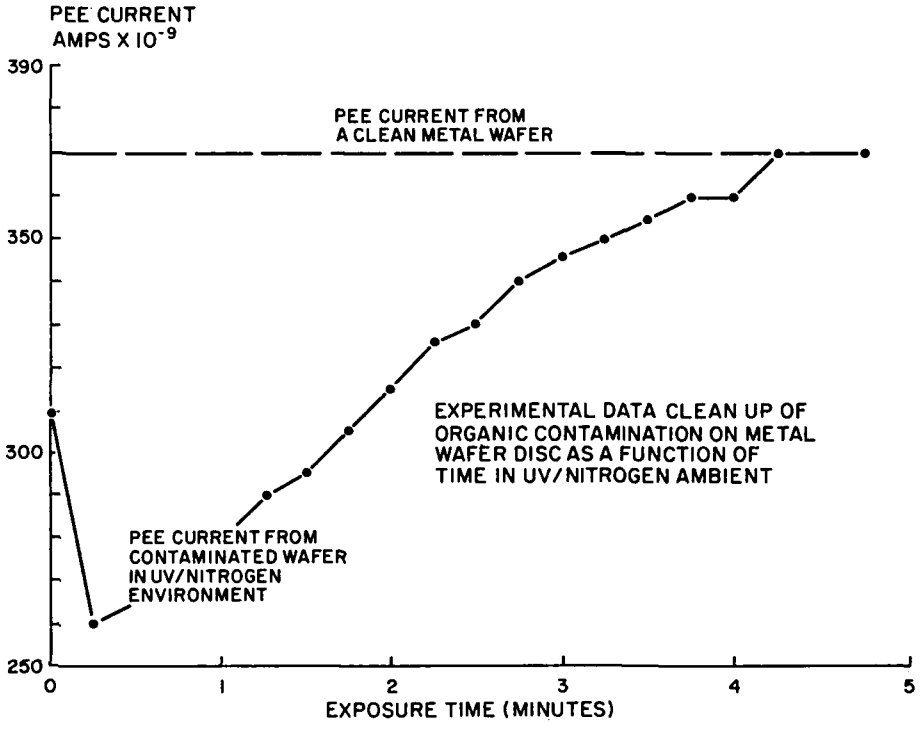
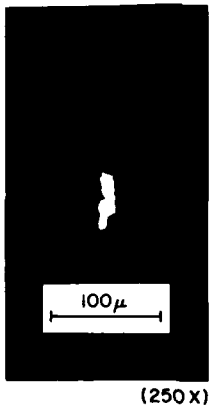


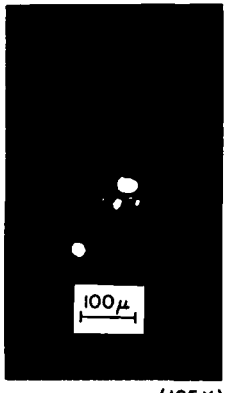

Figure 6




(250X)

SINGLE ALUMINA PARTICLE

(~15X50μ SIZE)




(125X)

ALUMINA PARTICLES

(4 → 5-40μ DIA. +15 → <5μ DIA.)

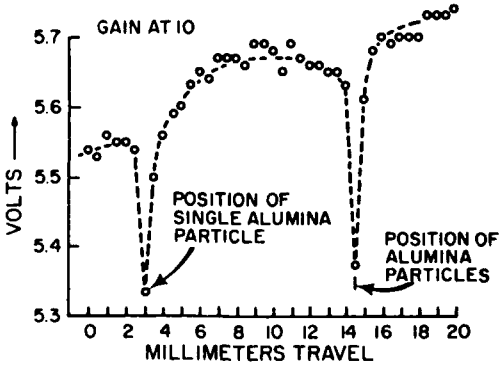



PHOTO ELECTRON EMISSION (PEE) RESPONSE CURVE (FROM SCANNING OVER THE ALUMINA PARTICLES AT 3mm/minute)

Figure 7

DETECTION OF ALUMINA PARTICLES

PARTICULATE REMOVAL WITH BLOWING
NITROGEN VERSUS DRY ICE SNOW

Figure 9




REMOVAL OF FINGERPRINT CONTAMINATION
FROM A 75 mm D SILICON WAFER WITH DRY
ICE SNOW. DATA TAKEN WITH TENCOR CO.
SURF SCAN MODEL 100. TOTAL CLEANING
TIME WAS 15 SECONDS




INTEGRATED CIRCUIT
CONTAMINATED WITH
SILICON DUST 50X



INTEGRATED CIRCUIT
AFTER 2 SECOND CLEAN-
ING WITH DRY ICE SNOW
50X

observatories (for telescope mirrors) and semiconductor/disc drive manufacturers for general cleaning.

In most cases the dry ice flakes are allowed to slide almost parallel to the surface for dust removal. If the flakes are allowed to impact at a very sharp angle to the surface there is some local melting and light oils (eg. fingerprints) can be removed, figure 12. Liquid $CO_2$ is well known to be an excellent solvent for organic oils, but has almost no effect on silicone oil or grease.

The dry ice cleaning apparatus is being manufactured in Tucson on a semicommercial basis. In most cases it is best to discuss the problems over the telephone before deciding to order the equipment. Potential users should contact the author (602)621-1617 to discuss possible applications.

2) Polishing With Dry Ice

Conventional polishing technologies have not changed appreciably in the last 5000 years. The Egyptians polished glass and metal with jeweler's rouge and other mineral powders in water or oil media. While these systems are effective, in terms of generating smooth surfaces, there is always the problem of cleaning the system whenever the polishing material must be changed. If residual particles are forced into the polished surface, they can serve as optical scattering sites or locations for the inception of stress cracking. If the particles break free of the surface, they can increase the contamination level inside the device, after packaging.

We have begun investigating the application of dry ice as a polishing media for polishing ceramics, metals and polycrystalline materials. The operation is done at -80 C to preclude rapid loss of the dry ice. When polishing is complete the system is allowed to warm up to room temperature and the dry ice disappears as $CO_2$ gas, leaving no residual material behind.

Dry ice offers several advantages here in that there is no liquid media to complicate the operation.[7] It is possible to use polar and/or non polar liquids (eg. alcohols) in conjunction with the dry ice polishing operation. This allows the operator to apply electrical potentials to change the hardness and the polishing characteristics of the substrate. Effects of this type are well known but have not yet been commercially applied in the optical industry.[7]

3) Cleaning With Ozone, Water and Ultrasonics

Conventional cleaning with solvents is becoming more difficult as federal agencies restrict the use of solvents and make it harder to dispose of them after they have been used. The problems of solvent cleaning, recycling, and disposal have been discussed in a recent book.[8] Solvent cleaning is not a simple process.

We have been investigating the use of water/ozone solutions with the addition of ultrasonics to speed up the reaction. The ability of ozone to destroy bacteria and break down organics has been reported.[9] The ultrasonic generator, shown in figure 13, simply speeds up the process.

In figure 14 we show a PVC tube that has been coated with a fluorescent dye. The system after partial cleaning is shown in figure 14. Cleaning of stainless steel tubing is shown in figure 15. Removal of residual solder flux presents no problems and the ozone/water/ultrasonic system may have important applications in packaging. The cleaning process is quite rapid, and technologies to preclude human exposure to ozone have been developed. There have been some studies of a small scale hypodermic needle system that would make it possible to use the ozone/water/ultrasonic system in very restricted areas. This would permit effective cleaning inside blind holes or in wired packages, without the hazard of damage to adjacent elements.

4) Electrets For Surface Cleaning

Electrets are usually made up of non conducting polymers that have been "poled" by heating to a temperature close to the melting point in a strong electrostatic field gradient. This effectively aligns the polar molecules in the polymer. As a result the material has a strong, permanent electrostatic charge.

Some applications of electrets involve their use as dust collectors, figure 16. In other cases the sheet material has been made into brushes for cleaning wafers, figures 17 and 18. Cleaning is quite effective and the pressures are low enough to preclude scratching. After use the electret brushes are usually discarded. Some of the electret brushes currently available are shown in figure 19.

Electrets can be used for dust monitoring, figure 20. Recently an electret monitor for radon has been offered for sale.[10] Readers are invited to call the author to discuss possible applications for electrets in their operations.

5) Gloves And Personnel Protection Systems

Gloves are often used to protect delicate equipment against oils and or particulates generated by human hands. Readers should be aware that gloves can be a source of problems. In figures 21, 22 and 23, we show some particle generation date from various types of gloves and data on particles produced by "new" and "old" latex gloves.

We have been working with the University of Arizona Medical School on other types of skin/product protection systems based on hydrogel. In figure 24 we show some results on particle generation by human fingers with and without a coating of hydrogel. In figure 25 we show laser scans of wafers that have been touched by human fingers with and without a coating of hydrogel. The hydrogel effectively precludes the transfer of skin oils and flakes to the wafer surface.

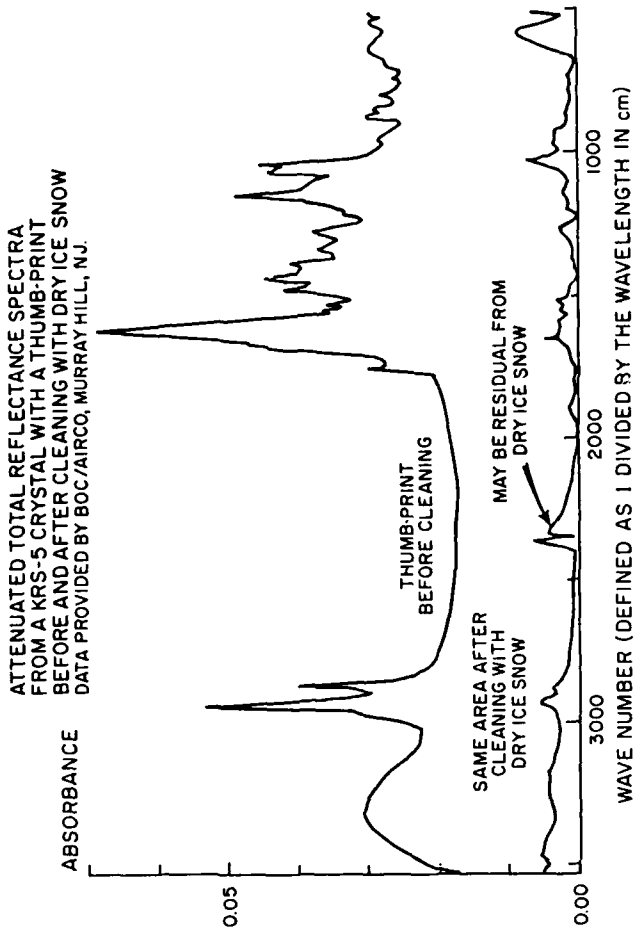
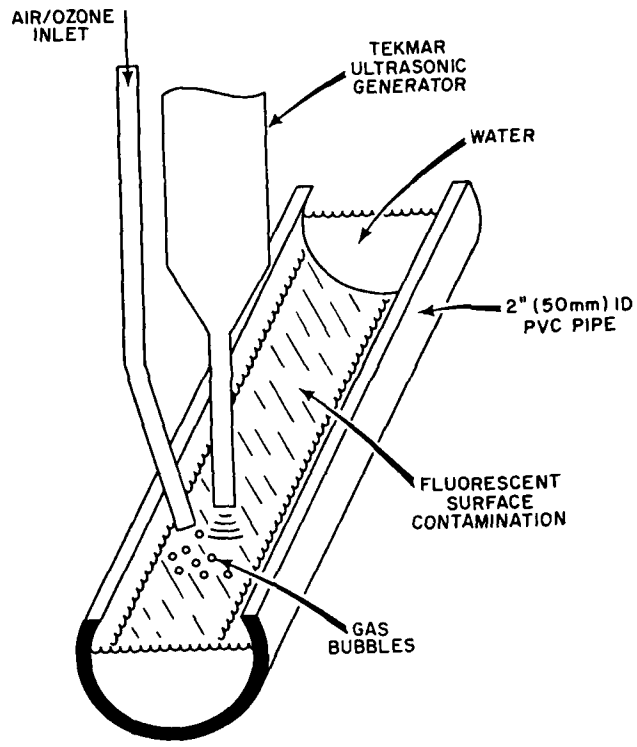



Figure 12

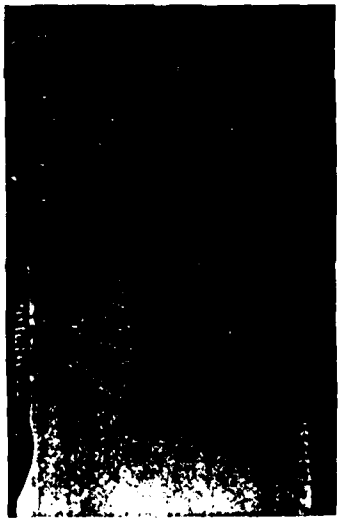



SCHEMATIC DRAWING TEST ASSEMBLY FOR
OZONE/WATER/ULTRASONIC CLEANING

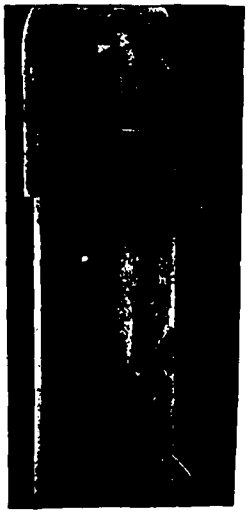
Figure 13

Figure 14

ULTRASONIC SYSTEM IN ACTION ON PVC PIPE COATED WITH MOTOR OIL AND FLUORESCENT DYE



PVC PIPE AFTER CLEANING (CLOSEUP)



PVC PIPE AFTER CLEANING (OVERALL VIEW)

OZONE/WATER/ULTRASONIC SYSTEM FOR REMOVAL OF ORGANIC CONTAMINATION FROM SURFACES

STAINLESS STEEL TUBE AND ULTRASONIC CLEANING HEAD

STAINLESS TUBE AFTER CLEANING

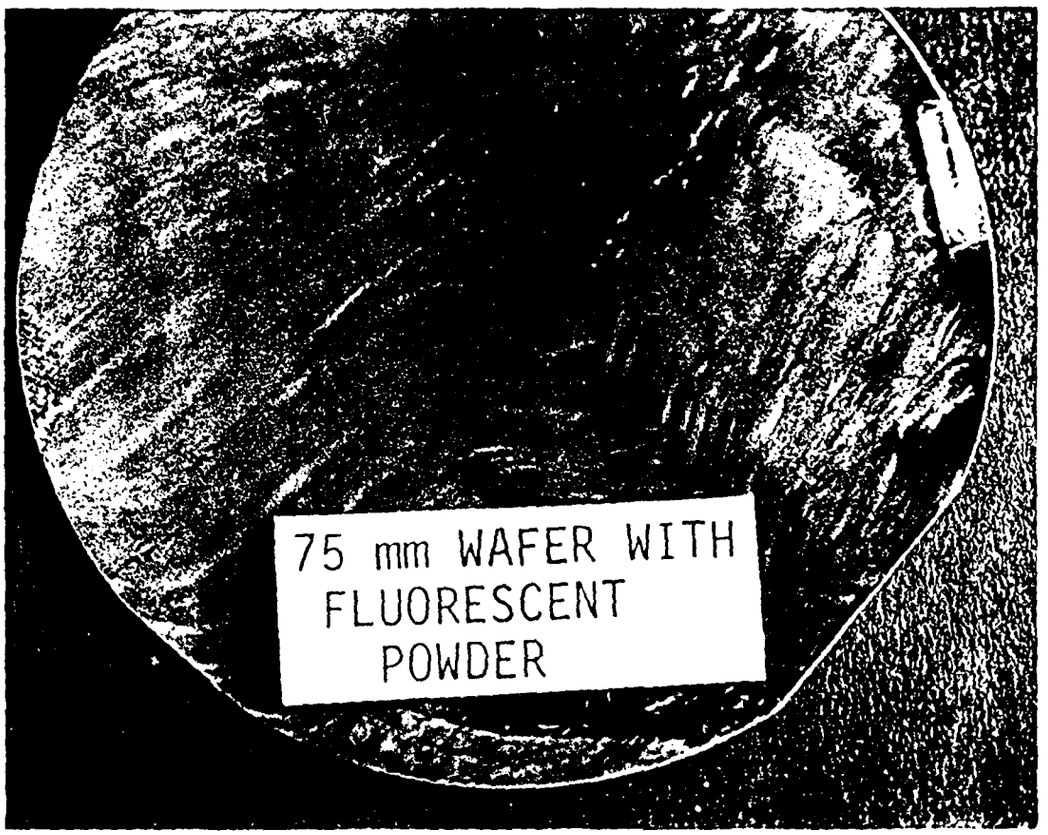
STAINLESS TUBE WITH GROOVE AFTER CLEANING

REMOVAL OF ORGANIC CONTAMINATION FROM STAINLESS STEEL WITH OZONE, WATER AND ULTRASONIC ENERGY

Figure 15




Figure 16


75 mm WAFER WITH
FLUORESCENT
POWDER

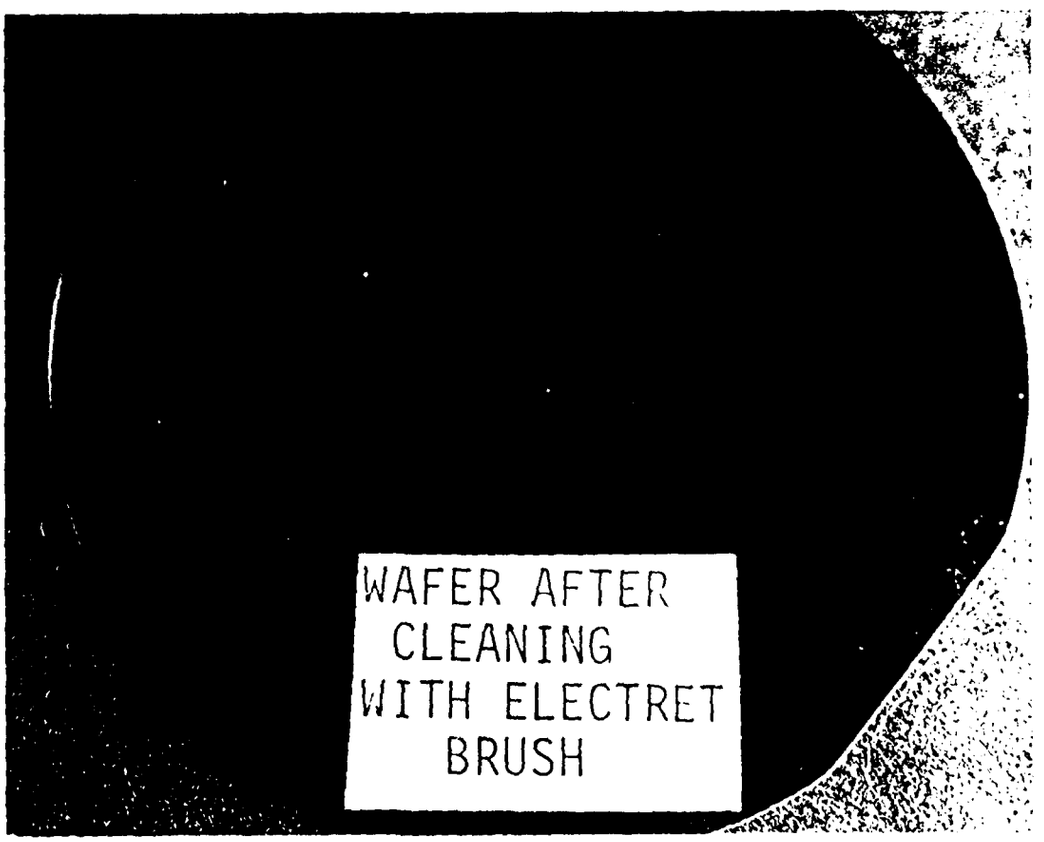
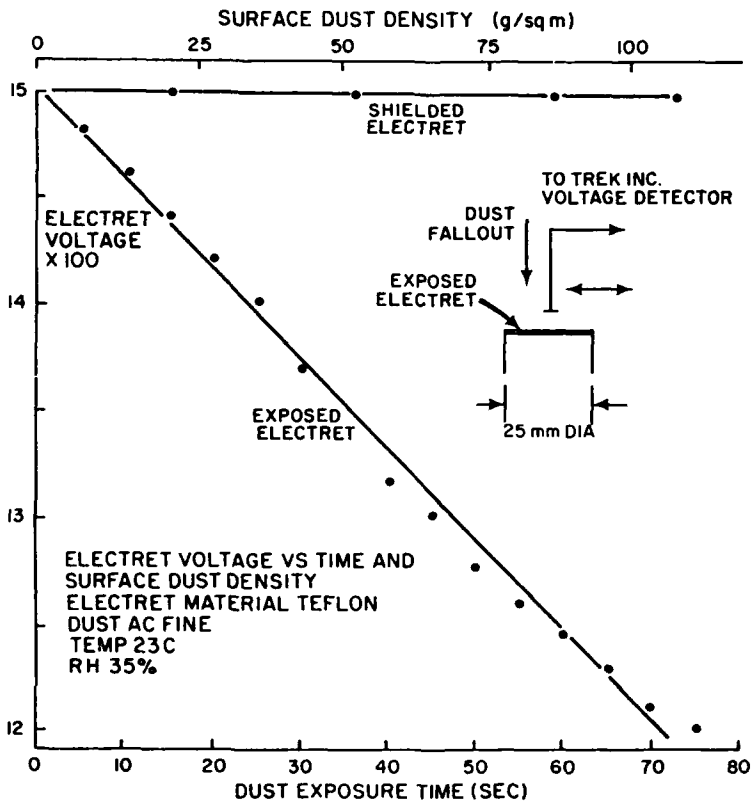
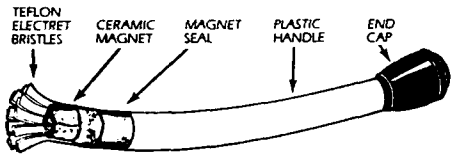

WAFER AFTER
CLEANING
WITH ELECTRET
BRUSH

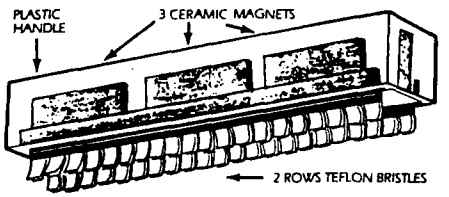
Figure 20



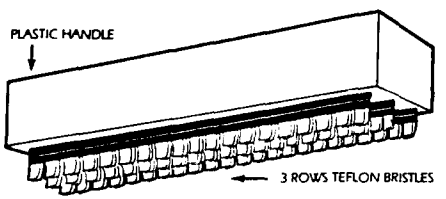

**K-ELECTRET**
**ELECTRET BRUSH WITH MAGNET**
**87-0092 SMALL AREA BRUSH**
**6 INCHES LONG**

For use in small areas, the permanently charged surfaces of the Teflon Electret attracts minute dust particles, while the magnets attract magnetic contamination.



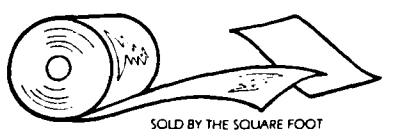

**K-ELECTRET**
**ELECTRET BRUSH WITH MAGNETS**
**87-0091**
**7 INCHES LONG**

The permanently charged surfaces of the Teflon Electret attracts minute dust particles, while the magnets attract magnetic contamination.




**K-ELECTRET**
**ELECTRET BRUSH**
***87-0093***
**7 INCHES LONG**

The permanently charged surfaces of the Teflon Electret attracts minute dust for micro-contamination control




**K-ELECTRET**
**ELECTRET SHEETS**
***87-0090-2 (2 mil)***
***87-0090-5 (5 mil)***

Available by the square foot in a 2 mil or 5 mil thickness Electret sheets of permanently charged Teflon. One side positive the other side negative

**K-ELECTRET**
4690 W. Wild Horse Driv
Tucson, AZ 85741
Telephone (602) 744-237

Figure 19

Figure 21

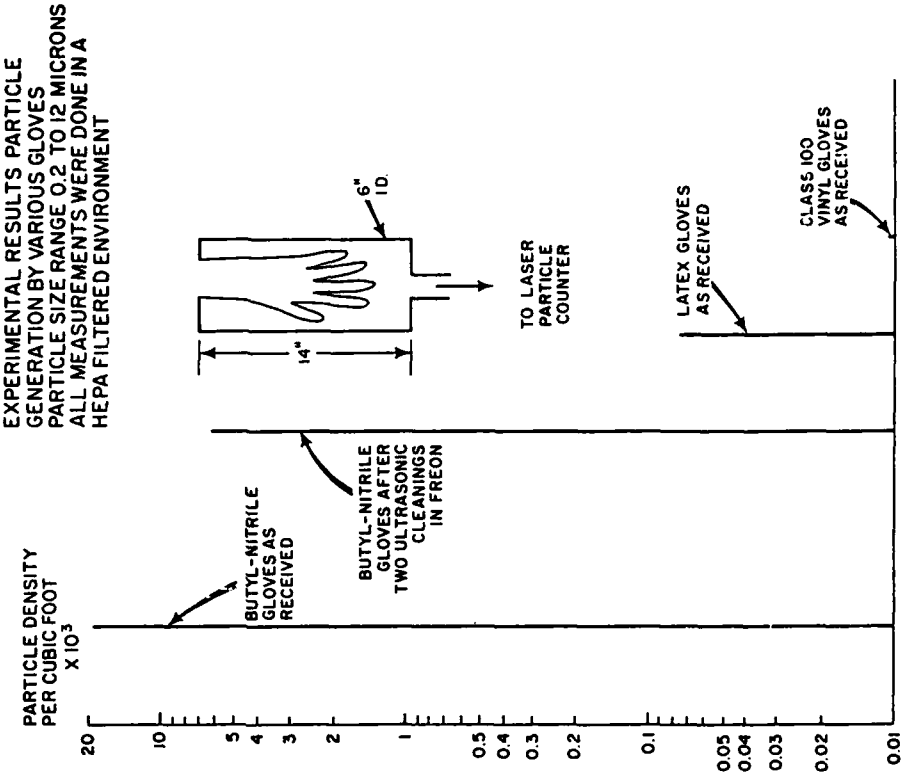



Figure 22

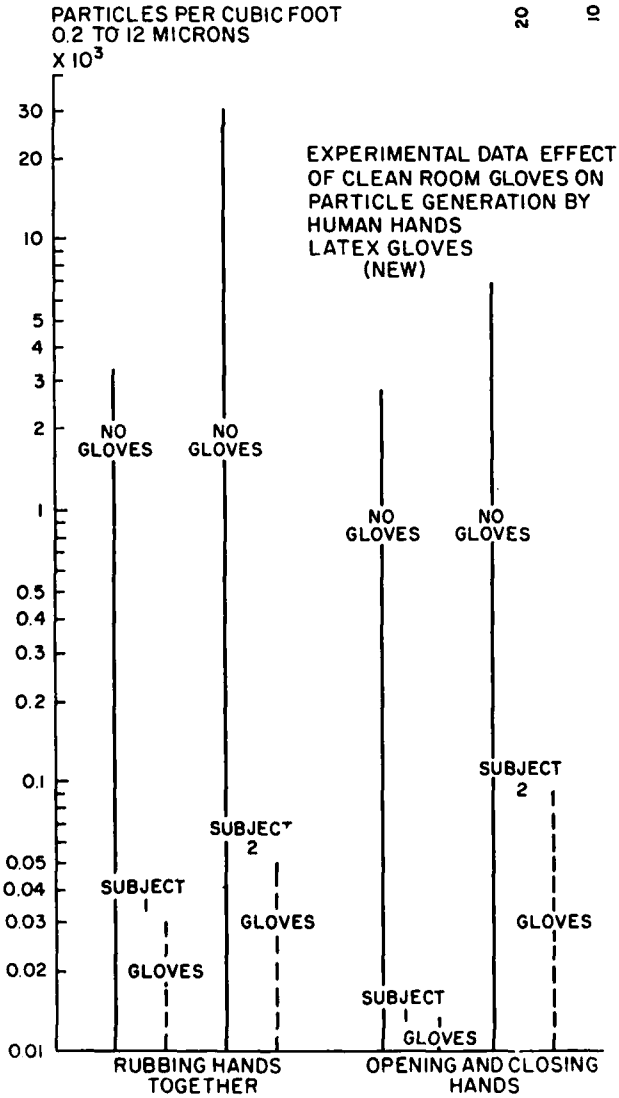
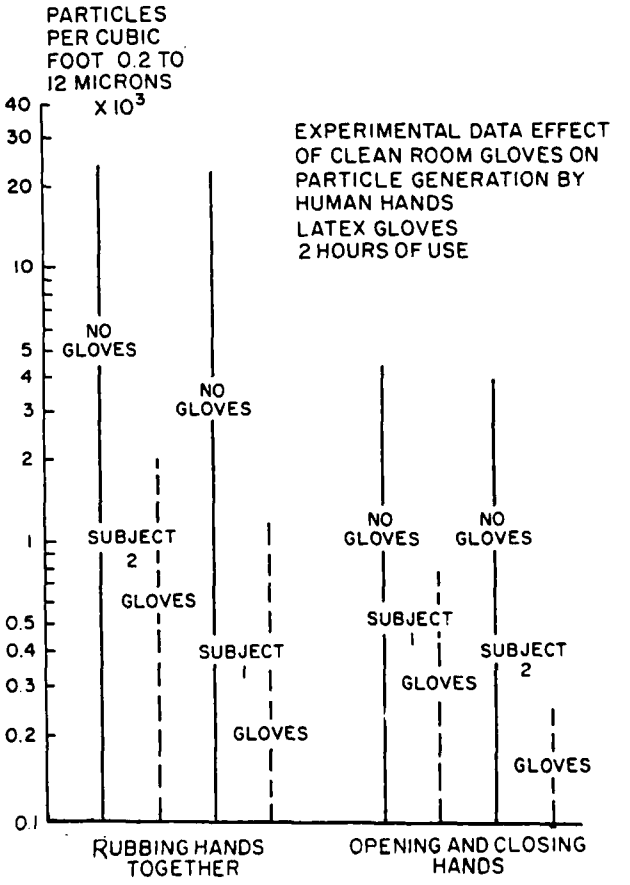



Figure 23



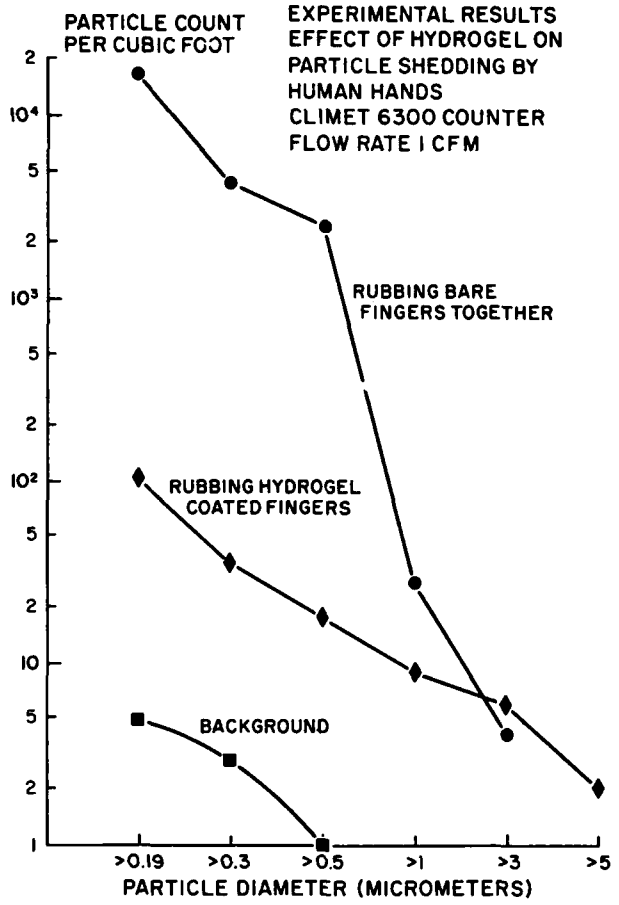
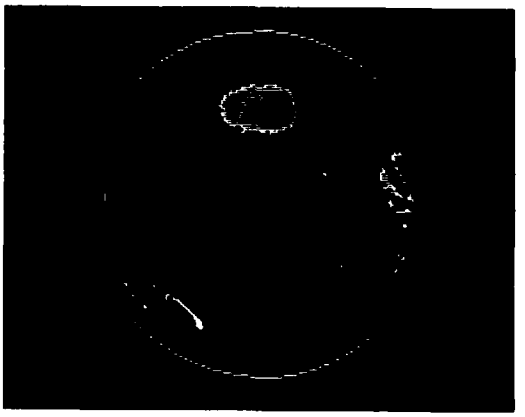
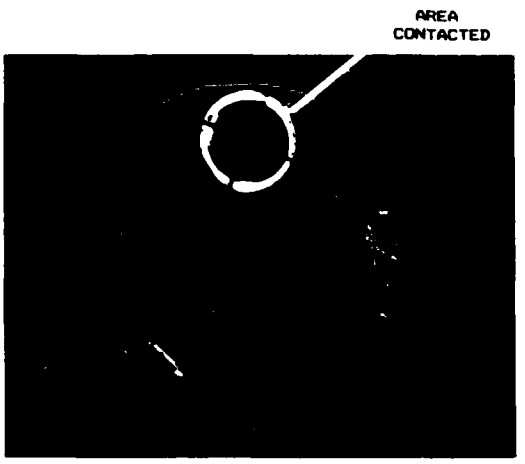



Figure 24



SURF SCAN (TM) MODEL 100 IMAGE OF A
75 MM WAFER WITH A FINGERPRINT (ALL
PARTICLES 1 MICRON AND ABOVE WERE IMAGED




SURF SCAN IMAGE OF 75 MM WAFER AFTER
TOUCHING WITH A FINGER COATED WITH
HYDROGEL. AREA CONTACTED HAS BEEN CIRCLED

Figure 25

Hydrogel can be removed from human skin by rinsing with warm water and or alcohol solutions. Hydrogel coatings have been shown to provide 100% protection against bacterial contamination. Tests with virus particles are under way. The liquid can be poured over a clean surface to provide a removable protective coating while a product is in storage. If it is poured over a contaminated surface, allowed to dry, and then removed there is a significant cleaning action, figure 26. We see this product as being far more effective than the usual strippable coatings.

The hydrogel material is composed of non toxic ingredients that have been individually approved for cosmetic use. The product is still experimental, but small quantities are available for test purposes to users who will agree to provide a report on their findings. Interested readers should contact the author for more information.

References


1) Amer. Soc. Testing and Materials. "Standard Method of Test for Hydrophobic Surface Films by the Atomizer Test", 1969 ASTM Book of Standards, pg. 21-78.

2) Hughes, A.L., L.A.Dubridge. "Photoelectric Phenomena". New York, McGraw Hill, 1932, Ch. XIII

3) Vig. J.R., J.W.LeBus. "UV Ozone Cleaning Of Surfaces". IEEE Transactions on Parts, Hybrids and Packaging, PHP-12, pg. 365 (1976)

4) Lilienfeld, P. "Optical Detection of Particle Contamination on Surfaces: A Review", Aerosol Sci. and Tech., Vol. 5, (1986), Pg. 145-65

5) Schmidt-Ott, A. "Enormous Yield of Photoelectrons from Small Particles", Phys. Rev. Lett., Vol. 45, No. 15, (1980), pg. 1283

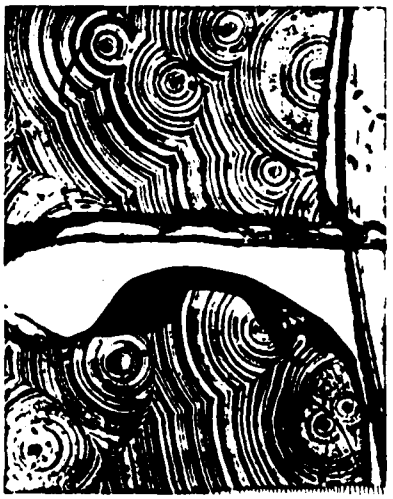
6) Fong. C.C. "Taking the Sand out of Blasting", Amer. Machinist, Vol. 11, Issue 21, (1974), pg. 67.

7) Westwood, A.R.C., R.M. Latanision. "Environment-Sensitive Machining Behavior of Nonmetals", Nat. Bur. Standards Sp. Tech. Pub. 348 (Ed. S.J. Schneider, Jr, R.W. Rice), May 1972, pg. 141.

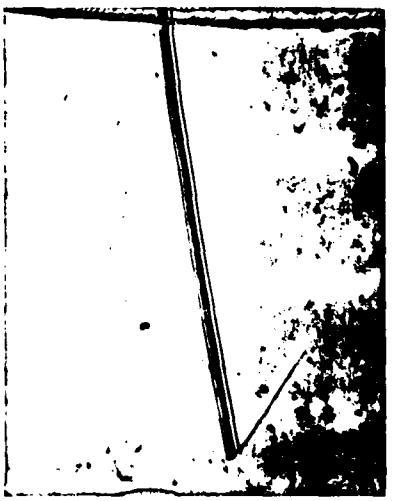
8) Ellis, B.N. "Cleaning and Contamination of Electronics Components and Assemblies". Electrochemical Publications Ltd. 8 Barns Street, Ayr, Scotland (1986)

9) Nebel, C. W.W. Nezgod. "Purification of Deionized Water by Oxidation with Ozone", Solid State Tech. Vol. 27, No. 10, (1984), pg. 185-92

10) Rad-Elec E-Perm (Radon Detector) RAD Electric, Inc., PO Box 1126, Frederick, MD 21701

HYDROGEL COATED OVER PARTICULATE CONTAMINATION



HYDROGEL COATED AREA AFTER RINSING WITH 40C TAP WATER MAGNIFICATION 50 X

UNCOATED AREA AFTER RINSING WITH 40 C TAP WATER MAGNIFICATION 50 X

EFFECT OF HYDROGEL AS A WATER SOLUBLE STRIPPABLE COATING TO REMOVE CONTAMINATION FROM SEMICONDUCTOR WAFERS

Figure 26

42nd Annual Frequency Control Symposium - 1988

# MEASUREMENT OF PLANO-CONVEX SC QUARTZ BLANKS USING LATERAL FIELD EXCITATION

A.W. Warner
Frequency Electronics, Mitchel Field, NY

## SUMMARY

Precision measurements of quartz crystal blanks before the electrodes are applied is advantageous in the production of crystal units excited by a **lateral** field. This can best be done by use of a flat, lateral field jig. The precision is further enhanced when chemical polishing techniques[1] are used, which preserves the flatness and orientation of the lapped or ground quartz plates. The same jig may also be used to identify unwanted or satellite resonant modes.

## C-MODE FREQUENCY MEASUREMENT

The crystal design used in this study is shown in Figure 1. It is a 2.5 MHz third overtone SC unit, with a resistance of 160 ohms and a Q of 5 million. Obviously it is a lateral field design, that is, the electrical field is across the crystal, not through it. The position and the shape of the electrodes guarantee an essentially single frequency of operation, as described in our 1985 Symposium talk[2]. The lack of a central electrode, while advantageous, limits the adjustment of frequency by the usual change in electrode thickness. In fact, the difference in frequency between plated and unplated quartz blank is only 150 Hz. It is therefore desirable to closely control the frequency of the unplated quartz blank so that subsequent frequency adjustment can be kept to a minimum.

By making the quartz blank plano-convex, and providing a flat measurement jig with the same electrode configuration as the crystal, some remarkable measurements can be made. The jig shown in Figure 2 has the crystal in place and a scale to indicate the field orientation. Figure 3 shows the jig alone. Figure 4 is a chart from our previous work at 5 MHz and shows how the impedance changes with lateral field orientation. The little symbols along the bottom indicate field direction. Zero degrees is a field along Z'. The lowest resistance for the C-mode occurs at about 110 degrees as shown.

In rotating the quartz blank on the measurement jig to find the highest coupling, we also will find the lowest frequency. This is illustrated in Figure 5. The reason for this is that the higher coupling results in a larger frequency difference between resonance and anti-resonance, and this difference becomes apparent as a lowering of the series resonant frequency. This is, of course, not new, having been used by Piezo-Technology and by LABCOM among others.

There is a rather sharp rise in activity, that is, a rise in the meter reading of the oscillator used, at the lowest frequency. The trick is to move the crystal blank about on the jig until this condition is realized, i.e., highest output, lowest frequency. After measurement, the crystal blank can then be further etched or polished to desired frequency. While the crystal is on the jig, it is very easy to check on the cutting shop to see that the -x flat is correct and that the contour is centered and on the correct side.

As an example of what can be accomplished and also of the sensitivity of the method, two sets of blanks were prepared. One set of four was polished by chemical etching[3] and the other set of four was mechanically polished. These were measured and then plated by vacuum deposition of gold. There were two platings, each consisting of half chemically polished and half mechanically polished.

The chemically polished units changed on plating by -150 Hz ± 2 Hz and the mechanically polished changed by -159 Hz ± 3 Hz as shown in Figure 6. The better performance of the chemically polished blanks, with a variation of 4 Hz, is attributed to the fact that chemically polishing does not change the flatness or the orientation of a blank as much as mechanical polishing does. The group difference of 9 Hz is attributed to the slight contouring effect of the mechanical polishing technique.

If the turn-over temperature is known, as from x-ray data, and the room temperature is carefully monitored, finished crystal units needing no subsequent frequency adjustment are possible. In practice, up to 50 Hz of adjustment is provided for.

## MODE IDENTIFICATION AND MODE PATTERNS

In manipulating the blank on the measuring jig, (Figure 3) it was noticed that other frequencies could be excited by changing the field direction and position. The electrode shape was designed originally to selectively excite the C-mode alone, excluding all satellite modes, and it follows that we should be able to selectively excite other modes as well by using the appropriate field position.

Figure 7 shows an old Bell Labs photo, probably by W.J. Spencer, illustrating mode shapes made visible by x-ray topography. This is an AT cut with a normal field. The main resonance at the left is 3.200 MHz and

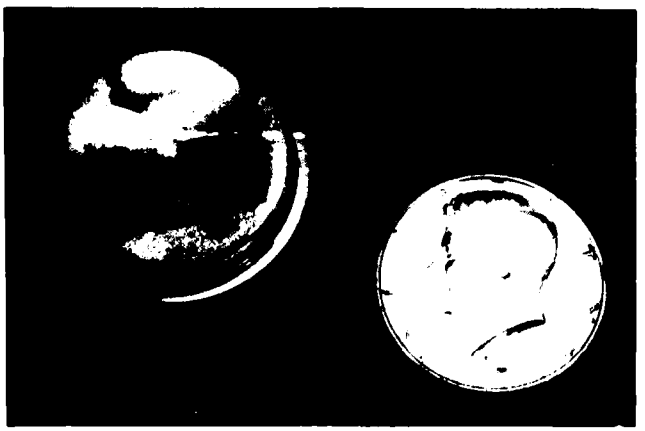


Figure 1. 2.5 MHz Third Overtone SC Crystal Unit

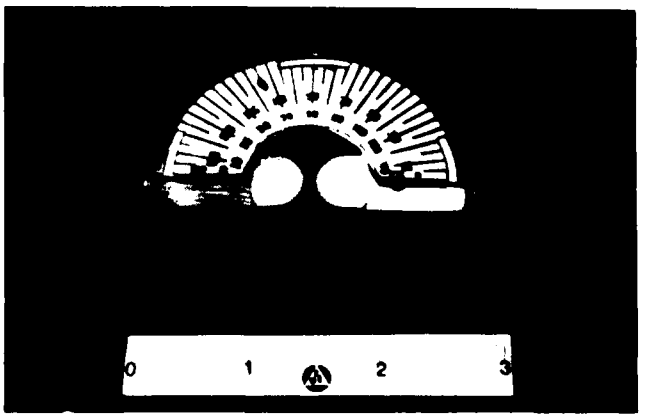



Figure 2. Jig for Lateral Field Blank Measurement With Scale and Crystal

CH2588-2/88/0000-202 $1.00 C 1988 IEEE

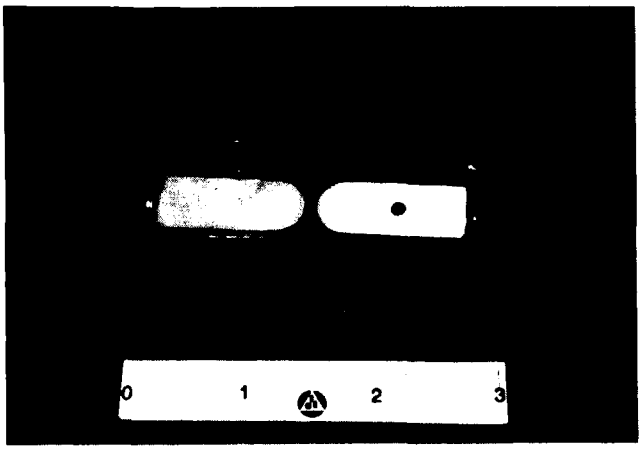



Figure 3. Jig for Lateral Field Blank Measurement Alone

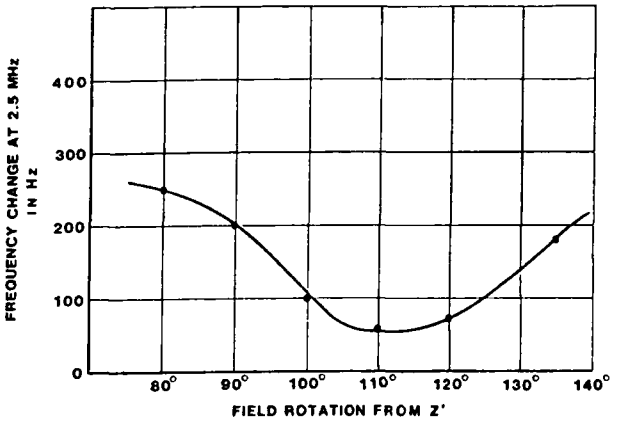



Figure 5. Crystal Blank Frequency vs. Field Rotation

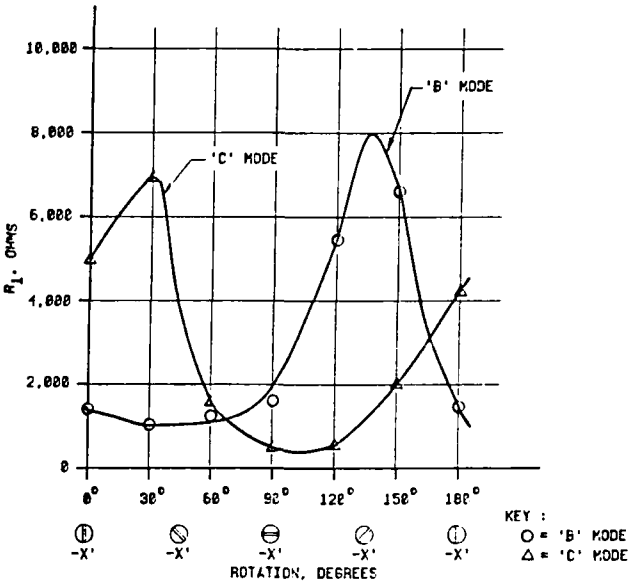



Figure 4. Effect of Gap Rotation on 'B' and 'C' Mode Resistance

| FREQUENCY, Hz | | | FREQUENCY, Hz | | |
|---|---|---|---|---|---|
| BLANK | PLATED | Δf | BLANK | PLATED | Δf |
| 2,499,982 | 2,499,820 | -162 | 2,500,135 | 2,499,985 | -150 |
| 2,500,002 | 2,499,846 | -156 | 2,500,087 | 2,499,935 | -152 |
| 2,500,118 | 2,499,959 | -159 | 2,500,109 | 2,499,959 | -150 |
| 2,500,079 | 2,499,920 | -159 | 2,500,104 | 2,499,956 | -148 |
| Δf= -159 ± 3 | | | Δf= -150 ± 2 | | |
| MECHANICAL POLISH | | | CHEMICAL POLISH | | |

Figure 6. Frequency Control in Plating

there are 12 satellite modes up to 3.8 MHz. Even more odd-ball modes exist in SC plates, since both X and Y fields can be present. With the lateral field jig these can be selectively excited.

Several field directions have been explored using the lateral field jig and SC cut blanks. The most prolific direction was along X' direction, that is, by sliding the crystal blank along the jig. Series of modes were also found along the C-mode direction + 20 degrees away, and the B-mode direction, - 10 degrees away. Along the X' axis, as shown in Figure 8, there was an almost linear series of resonances, 42 KHz apart, each consisting of two to eight active spots, that is, each resonance consisted of an odd or even number of phase reversals. The center of the plate can be either a mode or a loop. The position of the field, which is always placed at the outermost active spot, is shown along the ordinate of Figure 8, expressed as mm from the center of a 25 mm blank. The one exception, at 84 KHz above the main response, is a four loop mode, with the loops forming a square. This resonance could be excited along the Z' axis as well as the X' axis. Of course, when any mode is excited, it may be probed to find the active areas. The possibilities of finding new frequencies seem to be endless and the explanation will be left to others. Any one crystal blank can have as many as 21 distinct frequencies within 10 percent of the principal frequency.

Figure 9 is an artist's conception, showing the blank moved to the right so that the field is at the outermost active area of one of the modes of vibration. If the blank were rotated 20 degrees clockwise, the C-mode series would be excited by moving along that axis and so forth. The possible uses for this technique might include a) interfering mode identification or analysis, b) multiple frequency operation, or c) by knowing the shape of unwanted modes, to better design electrodes for single frequency operation. Several sets of electrodes could be applied to the same crystal, and a number of frequencies switched in or out, using the same crystal unit.

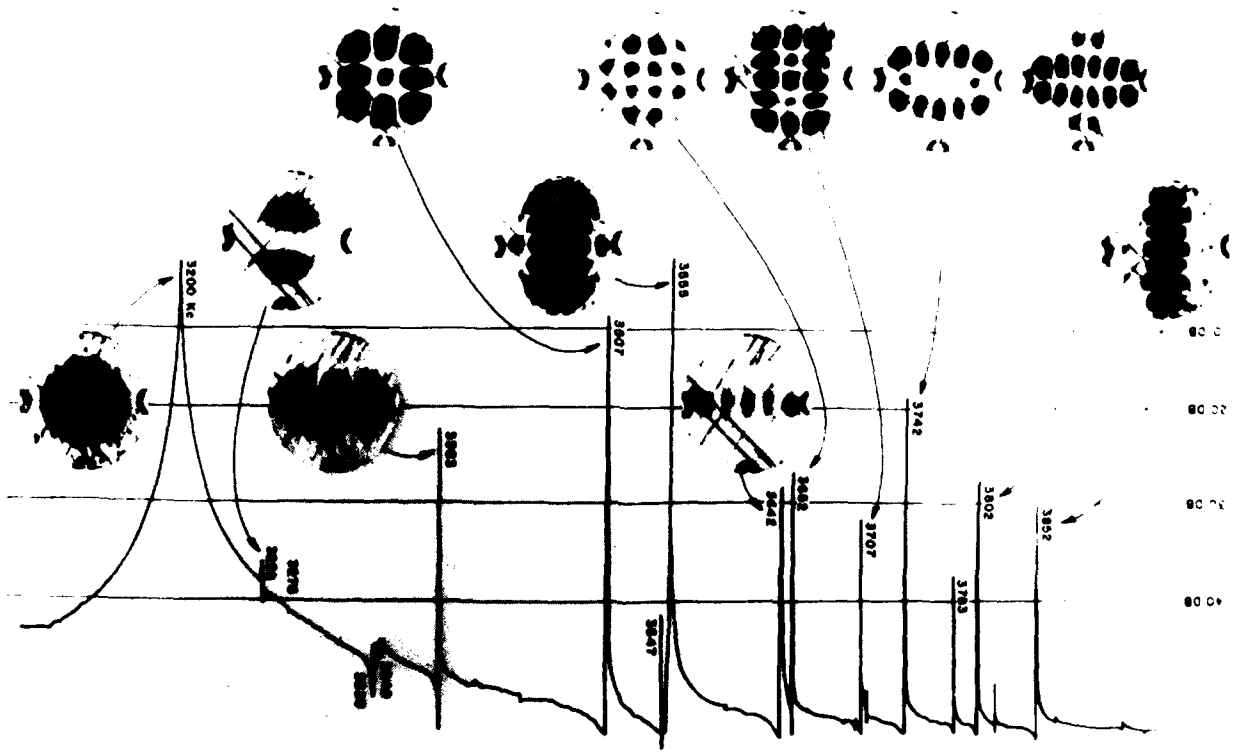


Figure 7. Mode Shapes Made Visible by X-Ray Topography (Bell Labs photo)

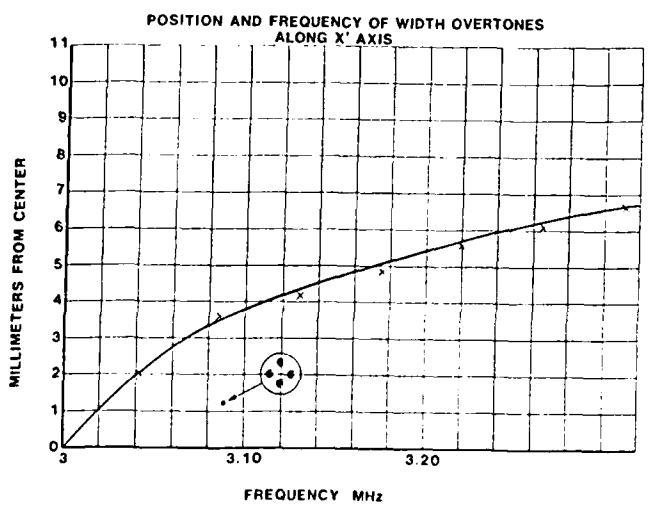



Figure 8. Position and Frequency of Width Overtones Along X' Axis

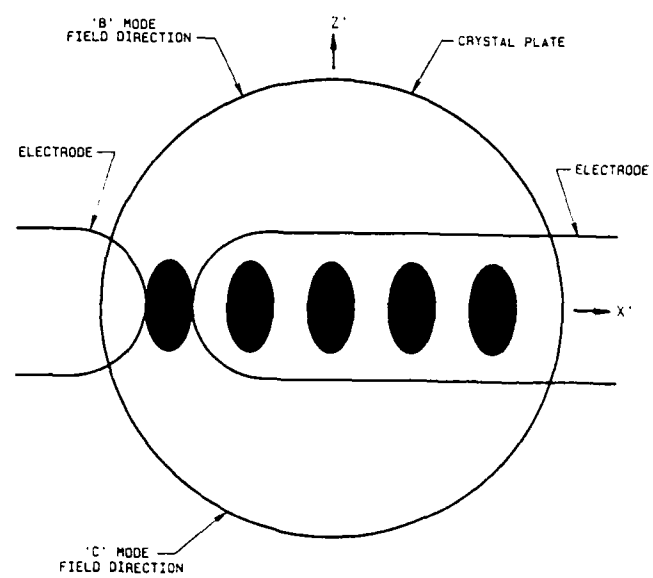



Figure 9. Method of Exciting a Width Overtone (Darkened areas are active. Field direction is along X')

## CONCLUSION

The use of the flat, shaped fixture on the flat side only of an SC crystal blank greatly increases the precision of blank measurement. The use of the same fixture for identification of other modes is an added attraction.

This work was supported by the U.S. Army Laboratory Command.

[1]R.J. Brandmayr and J.R. Vig, "Chemical Polishing in Etching Solutions that Contain Surfactants", Proc. 39th AFCS, 1985, pp.276-281.

[2]B. Goldfrank and A. Warner, "Lateral Field Resonators", Proc. 39th AFCS, 1985, pp. 473-474.

[3]R.J. Brandmayr and J.R. Vig, "Chemical Polishing in Etching Solutions that Contain Surfactants", Proc. 39th AFCS, 1985, pp.276-281.

42nd Annual Frequency Control Symposium - 1988

# USE OF ANNULAR SAW FOR THE CUTTING OF QUARTZ RESONATOR BLANKS AND SUBSTRATES

J Dowsett
R B Spencer
P Morley

STC Components, Quartz Crystal Division

Harlow, Essex, England, CM20 2DE

## Summary

For many years the Quartz Crystal Industry has used as a standard method the multiblade slurry saw for the cutting of Quartz resonator plates. Within the last few years STC Crystal Division has commissioned a large annular saw with a thin stressed blade for the cutting of tight tolerance blanks. A paper given in May 1987 at this symposium described the initial commissioning results. This paper is a report of the results now being achieved in full production of AT cut blanks. It also describes the extension of its use into cutting of doubly rotated cuts such as the SC cut.

## The Annular Saw

Recent advances in the Semi-conductor industry have lead to the development of large annular saws using a thin stressed blade with an inner diamond load edge, capable of cutting large diameter silicon bars. We have modified the machine to handle the cutting of large dimension quartz bars typically 43 x 24mm cross section. Figures 1-4 show views of the machine construction and illustrate position of the quartz block in the machine during the sawing process. The block is supported on a slide mechanism which is capable of correcting the ZZ and XX directions in the block. The target angle can be achieved by obtaining corrections from the first slice which is sufficiently flat and parallel and of good finish to obtain correction to 0.25 minutes of arc by X-ray orientation.

The machine is capable of indexing each feed step to maintain thickness control to better than 15 uM. Coupled with the cut feed mechanism is a blade distortion sensor which detects blade deflection on a continuous analogue plot and gives warning of blade wear and the need to redress the diamond cutting edge.

The quartz block is lumbered in the normal way, a reference face is prepared on the -X, this face is prepared with an angle tolerance of better than 5′ to minimise machine resetting between block loads. The block is cemented to a glass plate to ensure a complete cut through the quartz block. See figure 5.

The diamond loaded cutting blade is shown in figure 6; this blade has a 9" internal diameter and a thickness of 175 uM, the diamond cutting edge produces a Kerf of 300 uM and blank thickness of less than 200 uM have been achieved. Recent further work on the blade has resulted in a significant increase in the blade life by a further 50% with capability to cut quartz slices from which 22,000 blanks of 8mm diameter can be produced.

## Use for Cutting Doubly Rotated Cuts

The extension of the application of the annular saw to cutting of doubly rotated cuts has produced a further significant application of this type of saw. With the capability to adjust the cutting angle in two planes it was evident that it should be possible to use it in this mode. The results achieved have been very encouraging.

The method used was to produce a quartz block cut to include the first orientation. This was then mounted in the usual way and the second orientation set by the controls on the machine, however it was also possible to adjust for the first orientation. A small block was prepared and completely cut and all blanks produced were finished to complete resonators. The results from these were analysed, see figures 7-10, and it was found that the Frequency/Temperature characters were within the target specification for all resonators produced. The estimated angle distribution from this Frequency/Temperature data was found to be within 1.25′ of arc. The electrical parameter showed excellent control. This was achieved without any angle corrections on the cut blanks.

## Results

A comparison of results achieved from the annular saw initially on commissioning and in production with those from multiblade slurry saw are given in figures 10 to 13. Figures 10 and 11 show the angle distribution now being achieved in production. Figures 12 and 13 give a comparative summary of the economy and performance of the saws. Significant improvements are still being achieved by improved blade life and blade cutting accuracy with the consequent improvements in the efficiency and yield of the method. It is particularly notable that the losses of blanks for poor XX angle control have been practically eliminated and the yield of blanks within an angle tolerance of ±0.5′ on ZZ angle is now running at 3.5 times that previously achieved on the slurry saw.

## Conclusion

The advantages of the annular saw are:-

1) Ease of setting to target angle

2) Improved yield to ±0.5 angle tolerance

3) Improved thickness and parallelism control

4) Significant improvement in an operating economy

5) Capability to cut doubly Rotated cut crystals into angle tolerance

CH2588-2/88/0000-205 $1.00 C 1988 IEEE

Fig 1. General view of Annular Saw

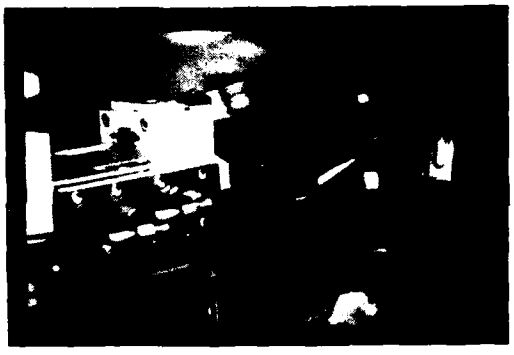


Fig 2. Close up of Quartz Block during cut

Fig 3. View of controls

Fig 4. Slice pick-up Mechanism

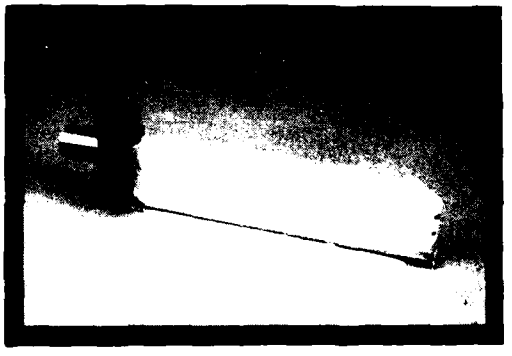


Fig 5. Mounting of Lumbered Quartz Block

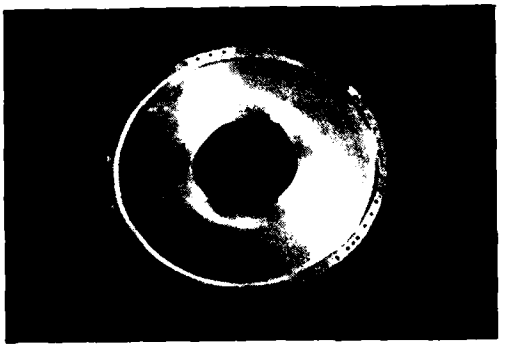


Fig 6. Saw blade

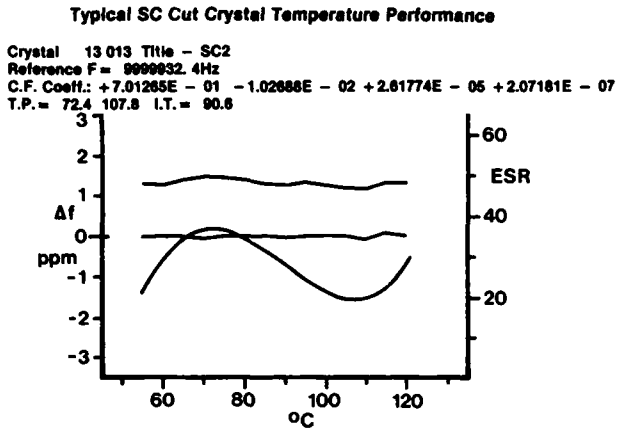



Fig 7.

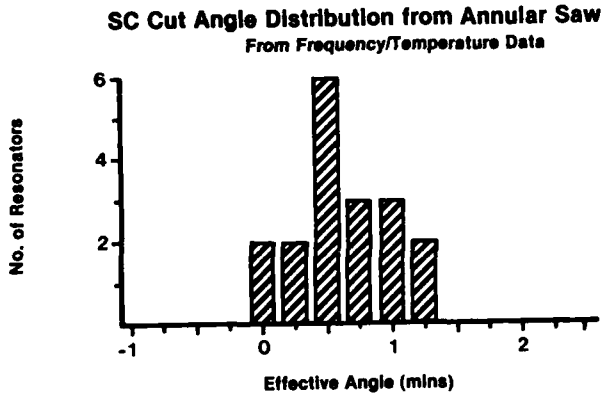



Fig 8.

## SC Cut Crystals

| Crystal code | fs (kHz) | drive (uW) | R1 (ohms) | C1 (fF) | Q | CO (pF) |
|---|---|---|---|---|---|---|
| 1 | 9999.799 | 23.84 | 45.82 | .27 | 1297978 | 4.41 |
| 2 | 9999.839 | 23.25 | 49.65 | .26 | 1213929 | 4.53 |
| 3 | 9999.781 | 23.81 | 45.97 | .26 | 1351175 | 4.40 |
| 4 | 9999.759 | 23.72 | 46.58 | .26 | 1308417 | 4.50 |
| 5 | 9999.782 | 23.34 | 49.06 | .25 | 1302103 | 4.51 |
| 6 | 9999.802 | 23.95 | 45.07 | .24 | 1456610 | 4.49 |
| 7 | 9999.805 | 23.80 | 46.06 | .25 | 1401493 | 4.57 |
| 8 | 9999.781 | 23.69 | 46.76 | .26 | 1293214 | 4.50 |
| 9 | 9999.779 | 23.75 | 46.41 | .25 | 1372580 | 4.50 |
| 10 | 9999.801 | 23.95 | 45.07 | .25 | 1403294 | 4.54 |
| 11 | 9999.782 | 23.85 | 45.71 | .26 | 1335856 | 4.52 |
| 12 | 9999.796 | 23.87 | 45.60 | .25 | 1395860 | 4.54 |
| 13 | 9999.795 | 23.67 | 46.89 | .26 | 1314931 | 4.40 |
| 14 | 9999.823 | 23.05 | 50.93 | .24 | 1291696 | 4.46 |
| 15 | 9999.804 | 23.54 | 47.74 | .24 | 1378587 | 4.56 |
| 16 | 9999.790 | 22.11 | 57.07 | .26 | 1052626 | 4.49 |
| 17 | 9999.793 | 18.63 | 82.51 | .25 | 762271 | 4.54 |
| 18 | 9999.816 | 23.60 | 47.39 | .27 | 1253322 | 4.45 |

Complete batch Electrical parameters

Fig 9.

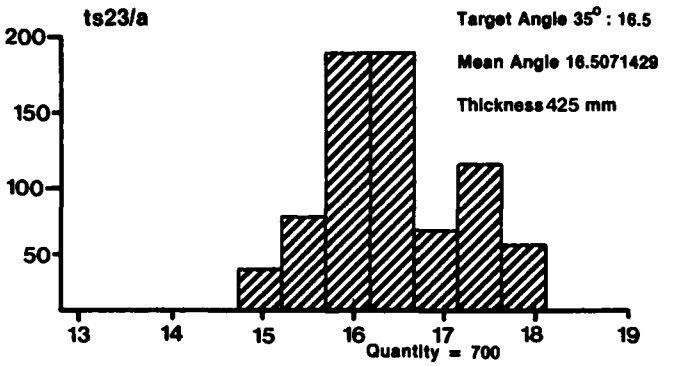



Fig 10.

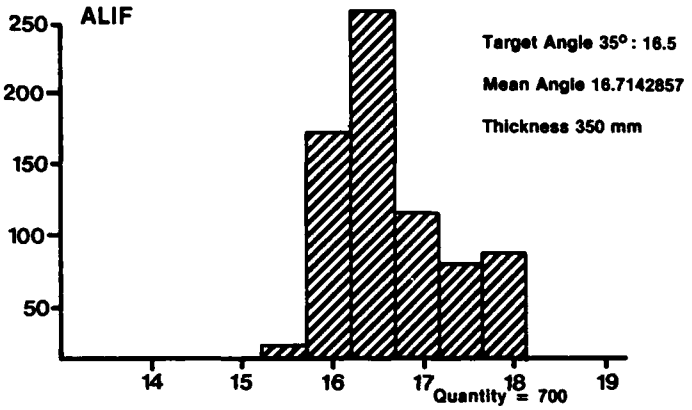



Fig 11.

## Summary of Results (Sawing Economy)

| | Multiblade Saw | Annular Saw Initial Results | Annular Saw Production Results |
|---|---|---|---|
| — Sawing rate | 3.2 cm²/min | 3.5 cm²/min | 3.8cm²/min |
| — Blade life | 8,515 cm² | 18,920 cm² | 28,000 cm² |
| — Quartz utilisation/block | 150 slices | 175 slices | 175 slices |
| — Blade change time | 0.75 hrs | 1.5 hrs | 1.5 hours |

Fig 12.

## Summary of Results (Blanks)

| | Multiblade Saw | Annular Saw | |
|---|---|---|---|
| | | Initial | Production |
| Angle Control z z angle | | | |
| Target Angle ± 0.25' | 75% | 95% | 95% + |
| Angle Spread | 3-4' | 2' | 1.5' |
| Blank yield ± 0.5' | 10% | 30% | 35% |
| Angle Control XX Angle ± 10' | | | |
| Yield | 10-15% | 95% | 98% |
| Blank Parallelism | 6.0μm | 5-10μm | 5-10μm |
| Blank Thickness Control | 100μm | 15μm | 15μm |

Fig 13.

42nd Annual Frequency Control Symposium - 1988

# X-RAY TOPOGRAPHIC STUDY OF VACUUM SWEPT QUARTZ CRYSTALS

A. ZARKA, M.T. SEBASTIAN and B. CAPELLE

Laboratoire de Minéralogie et Cristallographie
associé au CNRS
Universités Pierre et Marie Curie (Paris 6) et Paris 7
4 place Jussieu, 75252 PARIS CEDEX 05, FRANCE.

## Summary

Aluminium is invariably present in both synthetic and natural quartz crystals which exists as a substitution for silicon. The trivalent aluminium can make only three of the bonds normally made by a silicon and hence a negative charge exists on the site due to the fourth non-bonding oxygen. This charge is compensated by impurity ions which are bound electrostatically in an adjacent $\vec{c}$-axis channel. During sweeping the compensating impurity ions which are weakly bound near the aluminium atoms move toward the cathode. This result in the collection of negative space charges near the anode side of the crystal deforming the lattice. X-ray transmission topographs recorded from vacuum swept quartz crystals show a strong dark contrast near the anode side of the crystals due to lattice strain. This strong X-ray topographic contrast disappears when the crystals are subjected to prolonged X-ray irradiation. The negative charge formed near the anode side of the crystals during sweeping is compensated by forming aluminium-hole and aluminium-hydroxyl centres during the irradiation.

## Introduction

Quartz crystals as a consequence of their piezoelectric properties are widely used as frequency standards and hence the importance of the use of high quality sample crystals. In both synthetic and natural quartz crystals aluminium atoms invariably get substituted for silicon. From an ionic point of view aluminium is a +3 entity in a +4 site. The trivalent aluminium can make only three of the bonds normally made by a silicon and hence a negative charge on the site due to the fourth non-bonding oxygen. This charge is compensated by an alkali ion ($Li^+$, $Na^+$) or a proton ($H^+$) electrostatically bound in an adjacent $\vec{c}$-axis channel nearby or holes trapped at oxygen ions. The aluminium-hole centres are formed by the absence of an electron from a non-bonding orbital of an oxygen ion adjacent to the aluminium. Sodium is an important impurity in synthetic crystals since NaOH and $Na_2CO_3$ are used as mineralisers during the hydrothermal growth of the crystals. Because of the coulombic attractive force of the interstitial ions and the holes with the aluminium ions, and because of the high mobility of both interstitial ions and the holes, these charge compensators are usually located adjacent to the substitutional aluminium ions and this gives rise (1) to either Al-$OH^-$, Al-$Li^+$, Al-$Na^+$, or Al-hole centres.

Sweeping or electrodiffusion is a process which selectively exchanges monovalent impurity ions in α-quartz crystals and has been studied for many years (2-7). The sweeping is reported (6,8-11) to lower the production of etch channels significantly which is very important in the production of devices by photolithographic techniques and of very high frequency bulk wave oscillator crystals. Sweeping prior to device fabrication has been reported (6,12,13) to improve the performance of resonators and increases the radiation hardness of quartz resonators used in precision oscillators. Irradiation of the resonators fabricated from vacuum swept quartz is highly beneficial in reducing subsequent transcient and steady state frequency offsets in a radiation environment. The variations in frequency of unswept crystals are traceable to impurities and defects. Recently Hansom (14) made systematic X-ray topographic study of swept and unswept quartz using synchrotron radiation and reported that the basic dislocation network remains the same after sweeping. In the present paper we report the presence of an unusual X-ray topographic contrast in vacuum swept α-quartz crystals which disappears on prolonged X-ray irradiation.

## Experimental methods and results

Y-cut synthetic crystals (15x15x1 mm) were used for the sweeping experiments. The sweeping was done in vacuum at 500°C with platinum electrodes and a field of 2500 V/cm was applied parallel to the $\vec{c}$-axis of the crystals. The electrodiffusion process was carried out for 11 to 16 days. After sweeping, the crystals were studied by X-ray Lang projection topography with MoKα radiation. It is observed that the topographs of all the swept crystals show a strong dark contrast near the anode side of the crystals. Figure 1 shows the X-ray transmission topographs (10.0 reflection) of a typical crystal which was swept for 16 days. In order to study in detail the strong contrast which appeared on the topograph near the anode side, we took (15) a series of topographs from different X-ray reflections. Surprisingly it was found that this contrast disappeared gradually after recording few topographs (after 4 or 5 days exposure to X-rays). Notice the disappearance of the dark contrast in figure 2 which was initially present in figure 1. It is observed that the contrast appeared in all types of reflections (types 10.0 and 00.3).

## Discussion

Several workers observed (16-21) an anomalous increase in the X-ray and γ-ray diffracted intensity of α-quartz under the influence of electrostatic fields up to $10^4$V/cm. Charpa et al. (20) observed a change in the X-ray Bragg diffracted integrated intensity with clear hysteresis effects on slowly varying the electric field. By γ-ray diffraction experiments Dousse and Kern (19) showed that the reflectivity of (110) planes of α-quartz can be increased by 50% by applying a high electric field. Several investigators observed (20-23) a space charge collection in swept crystals.

CH2588-2/88/0000-208 $1.00 C 1988 IEEE

Alpha quartz crystals invariably contain aluminium ($Al^{3+}$) which exists as a substitution for silicon ($Si^{4+}$). The effective charge so created is compensated by positive interstitial ions such as alkalis and protons. When an electric field is applied along the $\vec{c}$-axis of the crystal in vacuum at elevated temperatures, the weakly bound alkali ions move toward the cathode. If additional mobile positive ions or charges are not available at the anode, sweeping would result in the accumulation of negative space charges near the anode. The strong dark X-ray contrast which appeared on the topographs recorded from vacuum swept crystals are due to lattice strain produced by the accumulation of negative charges. The contrast with striations parallel to the $\vec{c}$-axis, indicating inhomogenous lattice strain, is not due to dislocations since dislocation contrast vanishes on certain reflections depending on the Burgers vector and do not disappear on irradiation. The disappearance of the contrast on prolonged irradiation is due to the neutralisation of the negative charges. It is well known (1,3,13,24-27) that quartz crystals on irradiation produces Al-hole and Al-$OH^-$ centres. The ESR data of vacuum swept crystals showed (27) that the Al-hole concentration increases by a factor of four during irradiation.

## Conclusion

Electrodiffusion of impurity ions in alpha quartz crystals in vacuum leads to the accumulation of negative charges near the anode side of the crystals. This results in a strained crystal lattice which gives a strong X-ray topographic contrast near the anode side of the crystal. During irradiation of such crystals the negative charges created is neutralised by forming Al-hole and Al-$OH^-$ centres.


## Acknowledgements

The authors are grateful to Mr. Buisson and Mr. Bignon of SICN for their help in sweeping the crystals studied here.


## References


(1) L.E. Halliburton, N. Koumvakalis, M.E. Markes and J.J. Martin. J. Appl. Phys. 52 (1981) 3565.

(2) J.C. King. Bell System Techn. J. 38 (1959) 573.

(3) A. Kats. Philips Res. Repts 17 (1962) 133.

(4) H. Jain and A.S. Nowick. J. Appl. Phys. 53 (1981) 1477.

(5) E.H. Snow and P. Gibbs. J. Appl. Phys. 35 (1964) 2368.

(6) J.J. Martin. Proc. 41 Ann. Freq. Contl. Symp. (1987) 167.

(7) R.N. Brown, J.J. O'Connor and A.F. Armington. Mater. Res. Bull. 15 (1980) 1063.

(8) G.R. Johnson and R.A. Irvine. Proc. 41 Ann. Freq. Contl. Symp. (1987) 175.

(9) J.F. Balascio and A.F. Armington. 40 Ann. Freq. Contl. Symp. (1986) 70.

(10) J.R. Hunt and R.C. Smythe. Proc. 39 Ann. Freq. Contl. Symp. (1985) 292.

(11) J.G. Gualtieri. Proc. 39 Ann. Freq. Contl. Symp. (1985) 247.

(12) D.B. Fraser. Physical Accoustics, vol. 5 (1968) 59.

(13) J.C. King and D.R. Koehler. Precision Frequency Controll, Vol. 1 (1985) 147. Academic Press, New-York.

(14) W. Hansom. Proc. 41 Ann. Freq. Contl. Symp. (1987) 228.

(15) M.T. Sebastian, A. Zarka and B. Capelle. J. Appl. Cryst. 21 (1988) in press.

(16) W.J. Spencer. Physical Accoustics, vol. 5 (1968) Academic Press New-York.

(17) K. Yasuda and N. Kato. J. Appl. Cryst. 11 (1978) 705.

(18) K. Gouhara, Y. Bessho, K. Yasuda and N. Kato Jap. J. Appl. Phys. 21 (1981) L503.

(19) J.Cl. Dousse and J. Kern. Acta Cryst. A36 (1986) 966.

(20) E. Charpa, H. Ihringer, H. Jagodzinski and A. Kneifel. Z. Naturfor. A27 (1972) 469.

(21) M. Calamiotou, J. Psicharis, S.E. Filippakis and E. Anastassakis. J. Phys. C20 (1987) 5641.

(22) H.E. Wenden. Amer. Miner. 42 (1957) 859.

(23) E.L. Milne and P. Gibbs. J. Appl. Phys. 35 (1964) 2364.

(24) J.C. King and H.H. Saander. IEEE Trans. Nucl. Sci. NS-19 (1972) 23.

(25) J.A. Weil. Radiation Effects 26 (1975) 261.

(26) W.A. Sibly, J.J. Martin, M.C. Wintergill and J.D. Brown. J. Appl. Phys. 50 (1979) 5449.

(27) A. Kahan, F. Euler, H.G. Lipson, C.Y. Chen and L.E. Halliburton. Proc. 41 Ann. Freq. Contl. Symp. (1987) 216.

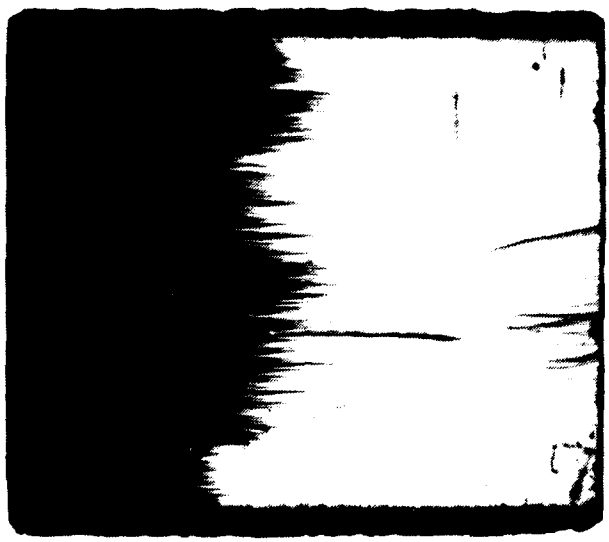


Fig. 1 - X-ray transmission topograph using 10.0 reflection of a crystal swept for 16 days.

Fig. 2 - X-ray transmission topograph using 10.0 reflection of the same crystal recorded on the seventh day of irradiation by X-rays. The contrast on Fig. 1 has disappeared here.

42nd Annual Frequency Control Symposium - 1988

# A NEW SHUTTER SYSTEM FOR FINE-TUNING COUPLED-DUAL CRYSTALS

Gerald E. Roberts
General Electric Company
Mountain View Road
Lynchburg, Virginia 24502

Abstract - One of the most critical operations in the fabrication of a coupled-dual resonator crystal is the fine-tuning operation used to adjust the crystal's two resonator frequencies and synchronous peak separation frequency (SPSF) to their final values. In this paper, a new process is discussed for precisely adjusting these three frequencies to their final values. This new process uses a masking and shutter arrangement, called the binary shutter system, which allows the two resonator frequencies to be lowered during the resonator adjustment by selective deposition onto the resonator electrodes and which also allows SPSF to be increased by deposition on the ground side of the crystal in the area between resonators. The process is controlled in such a way that both the resonator frequencies and SPSF may be adjusted very precisely.

The system consists of a frequency plate mask with an aperture opening onto each of the resonator electrodes on the active side of the crystal and an aperture opening onto the region on the ground side of the crystal shadowing the area between the resonator electrodes. A system of shutters then systematically controls the deposition through these apertures, thereby adjusting the frequencies of the two resonators and SPSF. Two different types of realization of this new system of shutters will be discussed.

## Introduction

In manufacturing coupled-dual crystals, the major parameters under consideration are the two resonator frequencies and coupling. Coupling, in the past, has been defined as the frequency difference between the zeros of the short-circuit driving-point impedance of a coupled-dual crystal. Unfortunately, this value of coupling is a function of the intrinsic coupling of the device and the difference between the resonator frequencies. Likewise, each resonator frequency ideally can be monitored by physically open-circuiting the other mesh in the well-known two-mesh equivalent circuit of a coupled-dual crystal and measuring the mesh frequency of the resonator in question. In practice, the condition is approximated by placing an external inductance across the static capacitance of the mesh to be open-circuited and tuning the parallel combination to the nominal resonant frequency of the resonator to be measured. This method of measuring resonator frequencies is undersirable due to inaccuracies, lack of repeatability, and difficulty in maintaining calibration.

With the discovery and development of the four-frequency measurement process [1,2], the two resonator frequencies can now be determined accurately and the value of the intrinsic coupling between two resonators can be obtained by determining the synchronous peak separation frequency (SPSF). One of several extremely useful features of the four-frequency process is that the values of the two resonator frequencies and SPSF can be monitored at any step in the fabrication process after the crystals are formed during the baseplate operation [3]. Because of this feature, the four-frequency measurement process is well-suited to monitor the crystal during the final fine-tuning process as well as because of the accuracy and repeatibility of its measurements, its need to monitor only one resonator, and its fixture not requiring calibration.

Typically, the final fine-tuning adjustment of a coupled-dual crystal is accomplished by depositing material on the resonator electrodes to lower their frequencies. Control of this deposition is critical. One way of controlling the deposition is to vary the deposition rate from an initial high rate to a lower rate as the resonator frequency approaches the target frequency as discussed by Rorick and Lewis [4]. A second approach is to interject a shutter between the vaporization source and the object to be plated to abruptly terminate the plating process as given in [5]. Another approach is to abruptly release the vacuum and inject an inert gas into the vaporization chamber as given by Buynak and Sauerland [6]. Of these systems, the shutter system seems to be the most easily controllable with previous systems suffering from mechanical complexity, lack of an accurate monitoring process, and an inefficient plating method.

Although it is known that resonator frequencies can be lowered by depositing material onto the resonator electrodes, it is important to know what happens to SPSF during the deposition process as well as how to adjust SPSF. Fair and Thompson [7] discuss adjusting the separation between the frequencies of the two zeros of the short-circuit driving-point impedance of a coupled-dual crystal by depositing material on or removing it from an electroded stripe between the resonators. This separation is the non-intrinsic coupling defined previously which is a function of both the intrinsic coupling and the difference between the resonator frequencies. Likewise, no real explicit technique for controlling the addition of the material or removal of it is discussed. Hokanson [8] is concerned only with the removal of material from the electroded stripe as well as the resonator electrodes using laser trimming as are Watanabe and Tsukamoto [9]. Thompson [10] suggests deposition of material on each resonator to establish a desired frequency separation between the two short-circuit zeros while lowering each resonator frequency. Then material is deposited in a solid pattern on the ground side of the crystal but shadowing both split electrodes on the active side to maintain the separation and adjust the resonators to their final values. No system for depositing material or controlling it is discussed. In [11], a similar scheme is presented in which both frequency and amplitude must be monitored. Also a complicated plating apparatus for controlling the deposition is presented which requires precise alignment since no frequency plating masks are used.

In this paper, the relationships between the resonator frequencies and SPSF during controlled deposition are discussed; and a new system [12] to control this deposition, called the binary shutter system, is presented. This system consists of a mask and shutter arrangement for controlling deposition during the adjustment of both SPSF and the two resonator frequencies. The most significant features

CH2588-2/88/0000-211 $1.00 C 1988 IEEE

of this system are its inherent simplicity of realization, ease of alignment of the deposition onto the crystal through the use of a mask, and the inherent precision with which the final values of SPSF and the two resonator frequencies can be obtained.

## Resonator Frequency Adjustment

One method of fine-tuning coupled-dual crystals which has been used often in the past is to plate only the resonators to their final frequencies. The material for plating typically comes from high purity silver or gold wires suspended upon filament wire[illegible] running generally parallel to each fac[illegible]r. A precision, non-conducting mask is hung [illegible] the wafer exposing all or part of one electrode of each electrode pair forming the input and output resonators, usually denoted as resonator A and resonator B, respectively. In this method, one electrode of each resonator is exposed with the two exposed electrodes being on opposite sides of the wafer. Figure 1 shows the specific electrode pattern of interest while Fig. 2 shows the exposed part of each electrode. After the vacuum chamber containing the unit is evacuated, power is applied to the filament via a power supply causing the deposition material to evaporate onto the unmasked portions of each face of the wafer where the frequency plate spots are deposited as illustrated in Fig. 2, thereby increasing electrode mass. The speed at which the silver evaporates is controlled partially by the filament power supply. A shutter system exposes the opening in the mask over one or both electrodes as the silver on the filament is evaporating.

It has been found experimentally that not only does the amount of deposited material on the electrodes effect the tuning of the crystal, but the position or location of the added material has an effect as well. In particular, a given mask which positions the frequency plate spots at certain points on the wafer produces its own intrinsic change in SPSF as a function of plateback. Plateback, in this context, is defined as the change in the mean value of the resonator frequencies (i.e., the difference between their mean frequency measured at any time in the plating process and the desired final mean frequency) as the resonators are adjusted to their final frequencies. This is illustrated in Fig. 3 where plateback for a particular mask and particular frequency plate spot locations is plotted against SPSF change ( $\Delta$SPSF) as material is plated on each electrode of the wafer as shown in Fig. 2. It can also be seen in Fig. 3 that the relationship between plateback and $\Delta$SPSF is generally a linear one. Experimentation has also shown the frequency plate spots which are positioned close to the unelectroded region between resonators produce the slope of line A in Fig. 3 steeper than when the frequency plate spots are positioned further away from the unelectroded region.

Consider Fig. 4 where a wafer is given having plated resonator electrodes separated by an unelectroded region. In this example, a pair of frequency plate spots shown by the solid lines denoted by FP Spot No. 2 produce a plateback vs SPSF change curve represented by line B in Fig. 5. As the location of the frequency plate spots is moved closer to the unelectroded region, as indicated by FP Spot No. 1, a curve having a steeper slope than line B as represented by line C is generated. Conversely, if the frequency plate spot were located at FP Spot No. 3, the slope would decrease as represented by line D. Thus, a family of curves is possible for a family of masks which have their frequency plate spots at different locations, relative to the unelectroded region between resonators. Also by placing frequency plate spots sufficiently close to the unelectroded region, the slope of the line becomes negative as represented by curves E and F in Fig. 5.

If a set of masks whose SPSF change profiles are such as given in Fig. 5 for a particular coupled-dual crystal design, if the coordinates of plateback and SPSF are known before final frequency adjust, and if the mask is chosen with a SPSF change profile closest to the initial coordinates of plateback and SPSF, then the crystal resulting from final frequency adjustment would have both resonators at the correct frequencies; and the value of SPSF will be close to its required value. However, errors will occur on crystals whose coordinates at the beginning of the fine-tuning operation does not fall on the particular SPSF change curve in Fig. 5 corresponding to the mask chosen.

## Frequency and SPSF Adjustment

By making the frequency plate mask such that both frequency adjust spots are on the front side or active side of the wafer, one can position a mask opening (usually rectangular) on the opposite side of the crystal which will allow the adjustment of SPSF. As shown in Fig. 6, this mask opening allows a SPSF adjust spot, denoted by SAS1, to be deposited on the ground side of the crystal. This spot should shadow the unelectroded region between electrodes on the active side. In addition, the SPSF adjust spot should be vertically larger than the vertical dimension of the resonators. It should also shadow a portion of the resonator electrodes adjacent to the unelectroded region between them on the active side and overlap each frequency adjust spot slightly as in Fig. 6.

As suggested in [12], a SPSF adjust spot used alone on the ground side will produce a SPSF change curve with a negative slope when plotted on the same axes as the plateback vs SPSF change plot obtained when only the resonators are plated as in Fig. 5. Such a curve is shown in Fig. 7. Therefore, a deposition algorithm can be devised to adjust SPSF until the dotted line in Fig. 7 intersects the solid line which can be produced by adjusting only the resonators. From this point, the resonators can be adjusted to their final frequencies producing the required value of SPSF as well.

In the discussion to follow, two types of realizations of the binary shutter system are described. They both basically expose the correct opening in the mask as outlined above having two frequency adjust spot openings on the active side of the crystal and one SPSF adjust opening on the ground side to adjust both resonator frequencies and SPSF to their final values.

## A Two-shutter System

The simplest realization of the binary shutter system consists of (1) a four-position shutter between a single filament and the frequency plate mask having two apertures opening onto the resonator electrodes on the active side of the crystal, and (2) a two-position shutter between a single filament and the frequency plate mask having one aperture opening primarily onto the area between resonator electrodes on the ground side as shown in Fig. 6. The four-position shutter allows exposing either opening or both openings in the frequency plating mask on the active side as well as shutting off the deposition of metal onto the resonator electrodes when the correct resonator frequencies (and SPSF) are obtained via the

solid curve in Fig. 7. The four-position shutter composed of a thin sheet of metal with two openings is shown in Fig. 8a and in a top view in Fig. 9 with the plating of the electrodes on the active side of the crystal depicted in Fig. 8b. A two-position shutter on the ground side simply exposes the mask opening or shuts off the plating when the correct coordinates of plateback and SPSF are obtained. The dotted curve in Fig. 7 shows the behavior of the crystal when a SPSF adjust spot is plated on the ground side. The two apertures in the shutter on the active side are separated laterally and vertically such that they can expose their respective resonator electrodes A and B only when in one of four positions. When the four-position shutter is in position 1, it exposes both electrodes. In position 1, the shutter has been retracted sufficiently through the plating fixture surface so that neither resonator is blocked. The partition shown in Figs. 8a, 8b, & 9 extends from the frequency plating mask to the shutter and prevents material from being inadvertently deposited on the other electrode when the shutter is in either position 2 or 4.

A key feature of this arrangement is that only one binary digit and only one opening is changed when the shutter sequences from one position to another. This provides relatively easy electronic monitoring of the two resonator frequencies and SPSF. An alternate four-position shutter system which involves changing more than one binary digit when switching between positions 2 and 3 is shown in Fig. 10a. The associated table depicting this operation appears in Fig. 10b.

The four-position shutter which controls deposition on the resonator electrodes on the active side of the crystal and the two-position shutter which controls deposition between the resonator electrodes on the ground side working in conjunction with each other allows the functional adjustment of the two resonator frequencies and SPSF in a dynamic mode. The fact that both resonators can be plated simultaneously, one resonator can be plated at a time, or both resonators can be completely masked out very quickly allows for extremely accurate frequency adjustment while working in conjunction with the plating of the SPSF adjust spot which changes SPSF (and frequency).

## A Three-shutter System

Another realization of the binary shutter system replaces the four-position shutter with two two-position shutters as shown in Fig. 11a. Basically, a modified shutter, denoted as the first shutter, with only positions 2 and 4 from Fig. 8a is used to plate one resonator or the other. Between it and the filament is a second shutter which consists of only ON or OFF positions, namely, positions 1 and 3 from Fig. 8a. This dual-shutter arrangement permits the plating of each resonator with even quicker cut-off of deposition at the expense of some small added system complexity. This arrangement has the particular advantage of assuring appropriate time for the search routine used in monitoring the crystal. A table showing the operation of this dual-shutter system is given in Fig. 11b. However, it is seen that only three shutter positions are available relative to our original four-position shutter. In this two-shutter realization, the condition in which both resonators are plated at the same time is not available, namely, position 1 in Fig. 8a.

Another two-position shutter arrangement which can replace the four-position shutter and realize all four operational positions is shown in Fig. 12a. Each shutter has only two positions. The first shutter has a single aperture which aligns with mask aperture for resonator electrode A and the second shutter has a single aperture which aligns with the mask aperture for resonator electrode B. When the two shutters are in their ON positions, they do not obstruct the deposition of material through the masks for either electrode. In their OFF positions, they mask both resonator electrodes, A and B, from the filament. When both shutters are in their OFF positions each blocking off one of the frequency plate mask openings, neither resonator can be plated and the equivalent of position 3 in Fig. 8a is obtained. If the first shutter is ON and the second shutter is OFF, the equivalent of position 4 in Fig. 8a occurs and only one resonator (B) is plated. If the first shutter is OFF and the second shutter is ON, the equivalent of position 2 in Fig. 8a is achieved and the other resonator (A) is plated. This four-position, two-shutter system on the active side of the crystal tracks the binary operation of the single shutter in Fig. 8a except that it has the added benefit of being able to turn off the plating at any time.

Both of these dual shutter arrangements can be used for exposing the resonator electrodes on the active side of the crystal in conjunction with an ON-OFF shutter on the ground side, thereby controlling plating of the SPSF adjust spot in Fig. 6.

## Conclusions

The actual binary shutter system provides a simple way of exposing the correct opening or openings in the frequency adjust mask on both the active side of the crystal and the ground side. In the two shutter version, the four-position shutter on the active side provides an extremely good conceptual picture of how the system works with the two-position shutter on the ground side. In general, the four-position shutter is more difficult to implement than the two two-position shutters. Moreover, the two-position shutters are better from a standardization point of view and they are more reliable in a production environment. In general, the binary shutter system provides an extremely simple, convenient, and precise means of exposing particular openings in the frequency plate mask for the deposition of metal onto the crystal or blocking off the openings from the vaporization source. It provides the basis for an extremely accurate system for the final fine-tuning adjustment of the two resonator frequencies and SPSF.

## References


1. H. J. Peppiatt and G. E. Roberts, "Method of measuring parameters of a crystal filter, "U. S. Patent 4,093,914, June 6, 1978.

2. G. E. Roberts, "A four-frequency process for accurately measuring coupled-dual resonator crystals, "Proc. 33nd Ann. Symp. on Frequency Control, May 30-June 1, 1978, pp. 159-165.

3. G. E. Roberts, "A practical metalization system for high frequency coupled-resonator crystals, "Proc. 9th Annual Quartz Devices Conference and Exhibition, Aug. 25-27, 1987, pp. 3-8.

4. W. G. Rorick and H. O. Lewis, "Quartz crystal resonator tuning control apparatus," U. S. Patent 3,670,693, June 20, 1972.

5. A. F. Turner and O. A. Ullrich, "Apparatus for making optical wedges," U.S. Patent 2,432,950, Dec. 16, 1947.

6. G. R. Buynak and F. L. Sauerland, "Vacuum deposition method for frequency adjustment of piezoelectric resonators," U. S. Patent 4,112,134, Sept. 5, 1978.

7. I. E. Fair and E. C. Thompson, "Crystal Filter," U.S. Patent 3,573,672, April 6, 1971.

8. J. L. Hokanson, "Laser machining thin film electroded arrays on quartz crystal substrates," Proc. 23rd Ann. Symp. on Frequency Control, 1969, pp. 163-170.

9. N. Watanabe and K. Tsukamoto, "High-performence monolithic crystal filters with stripe electrodes," Electron. Commun. Japan, Vol. 57, part A, pp. 53-60, 1975.

10. E. C. Thompson, "Method of manufacturing a monolithic crystal filter," U. S. Patent 4,112,147, Sept. 5, 1978.

11. E. C. Thompson, "Method of fine-tuning a monolithic crystal filter," U. S. Patent 4,343,827, Aug. 10, 1982.

12. G. E. Roberts and S. Toliver, "Shutter apparatus for fine-tuning a coupled-dual resonator crystal," U. S. Patent 4,627,379, Dec. 9, 1986.

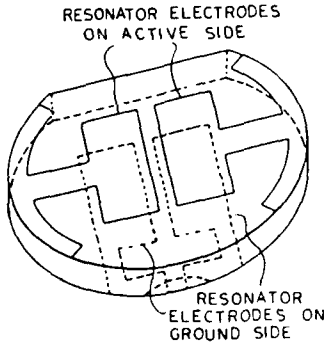



FIGURE 1. A simplified prospective view of a coupled-dual resonator crystal.

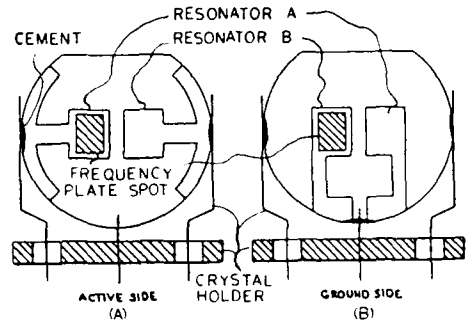



FIGURE 2. Plane views of a) The active side, and b) the ground side of a coupled-dual crystal.

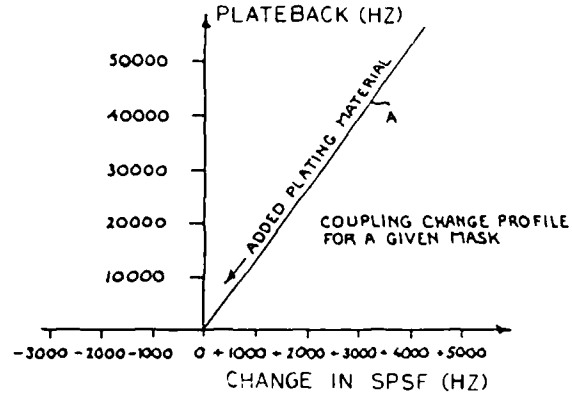



FIGURE 3. A graph of the linear relationship between plateback and SPSF change for a particular frequency plate mask when the resonator electrodes are plated.

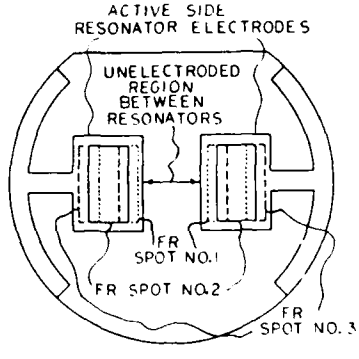



FIGURE 4. A front view of a coupled-dual crystal showing several locations for the frequency plate spots.

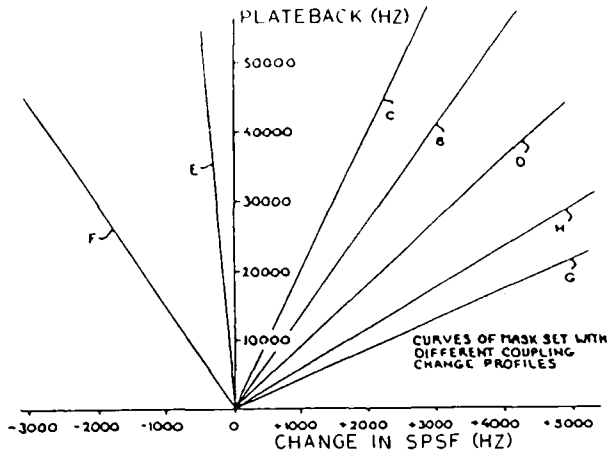



FIGURE 5. A family of curves of plateback vs SPSF change for a family of frequency plate masks.

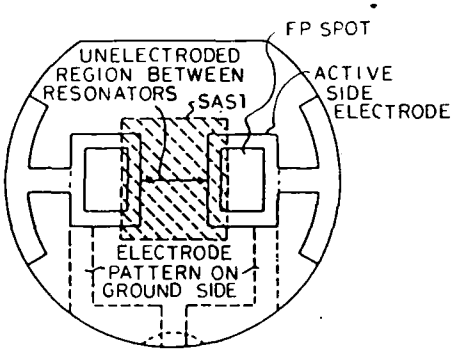



FIGURE 6. A view of the SPSF adjust spot on the ground side of the coupled-dual crystal.

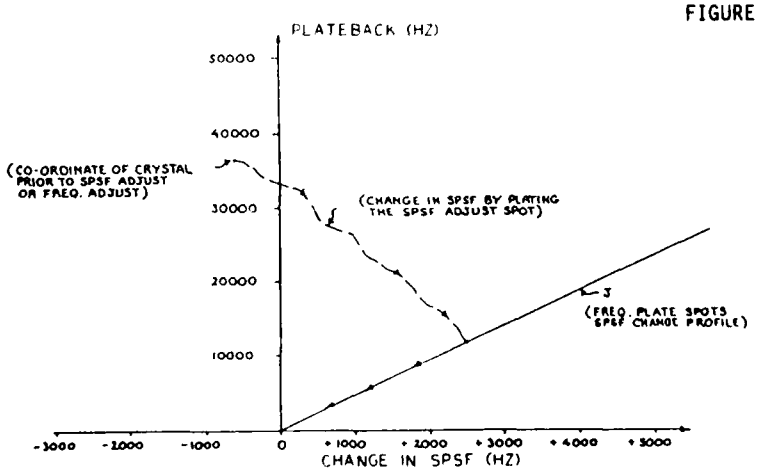
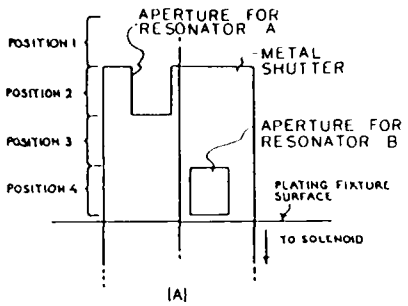
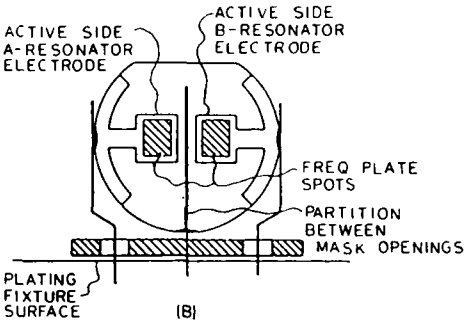



FIGURE 7. A graphical representation of how the SPSF adjustment on the ground side of the crystal works with the frequency adjustment on the active side of the crystal.




| SHUTTER POSITION | SHUTTER APERTURES LEFT | SHUTTER APERTURES RIGHT | DECIMAL VALUE | PLATING ACTION |
|---|---|---|---|---|
| 1 | 1 | 1 | 3 | PLATE BOTH RESONATORS |
| 2 | 1 | 0 | 2 | PLATE A-RESONATOR |
| 3 | 0 | 0 | 0 | STOP PLATING |
| 4 | 0 | 1 | 1 | PLATE B-RESONATOR |

(C) APERATURE STATUS
0 - CLOSED
1 - OPEN

FIGURE 8. a) The four-position shutter on the active side of the crystal between the filament and mask positioned on the crystal, b) view of the active side of the crystal with frequency adjust spots on the resonator electrodes relative to the plating fixture surface, c) table showing the operation of the four-position shutter.

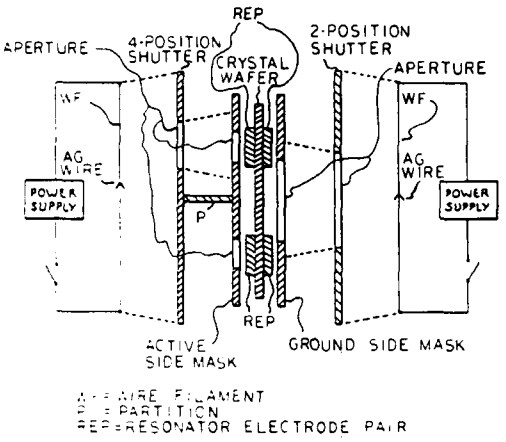
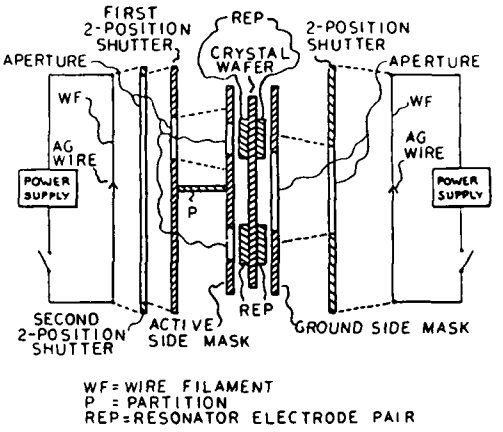



FIGURE 9. A top view schematic representation of a two-shutter binary shutter system relative to the coupled-dual crystal and mask.




(A)

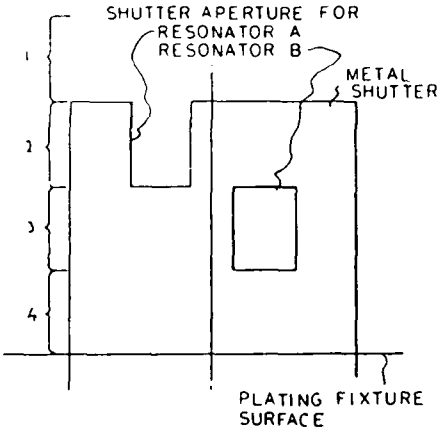
APERTURE STATUS

| SHUTTER POSITION | 1ST SHUTTER LEFT | 1ST SHUTTER RIGHT | 2ND SHUTTER | COMPOSITE A RESONATOR | COMPOSITE B RESONATOR | DECIMAL VALUE | PLATING ACTION |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 1 | 1 | 0 | 2 | PLATE A-RESONATOR |
| 2 | 0 | 1 | 1 | 0 | 1 | 1 | PLATE B-RESONATOR |
| 3 | X | X | 0 | 0 | 0 | 0 | STOP PLATING |

APERTURE STATUS
0 = CLOSED
1 = OPEN
X = DOESN'T MATTER

(B)

FIGURE 11. a) A top view schematic representation of a three-shutter binary shutter system relative to the coupled-dual crystal and mask, b) table showing the operation of the two shutters on active side of the crystal.

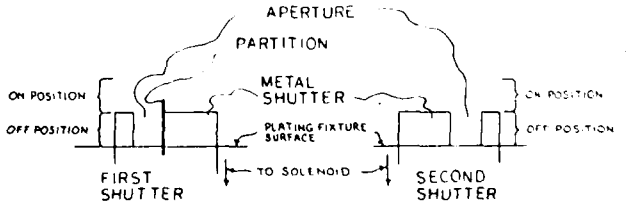



(A)

| SHUTTER POSITION | SHUTTER APERTURES LEFT | SHUTTER APERTURES RIGHT | DECIMAL VALUE | PLATING ACTION |
|---|---|---|---|---|
| 1 | 1 | 1 | 3 | PLATE BOTH RESONATORS |
| 2 | 1 | 0 | 2 | PLATE A - RESONATOR |
| 3 | 0 | 1 | 1 | PLATE B - RESONATOR |
| 4 | 0 | 0 | 0 | STOP PLATING |

(B)

FIGURE 10. a) An alternate four-position shutter arrangement, b) the corresponding table showing shutter operation.




(A)

| SHUTTER POSITION | 1ST SHUTTER LEFT | 1ST SHUTTER RIGHT | 2ND SHUTTER LEFT | 2ND SHUTTER RIGHT | RESONATORS A | RESONATORS B | DECIMAL VALUE | PLATING ACTION |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 1 | 3 | PLATE BOTH RESONATORS |
| 2 | 1 | 0 | 1 | 1 | 1 | 0 | 2 | PLATE A - RESONATOR |
| 3 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | STOP PLATING |
| 4 | 1 | 1 | 0 | 1 | 0 | 1 | 1 | PLATE B - RESONATOR |

1ST SHUTTER APERTURE STATUS
ON = 11
OFF = 10

2ND SHUTTER APERTURE STATUS
ON = 11
OFF = 01

(B)

FIGURE 12. a) An alternate arrangement for obtaining the four positions using two two-position shutters on the active side of the crystal, b) table showing the operations of this alternate two shutter arrangement.

42nd Annual Frequency Control Symposium - 1988

# ANALYSIS OF ANHARMONICS IN SURFACE SKIMMING BULK WAVE DEVICES

Janpu Hou

Electronic Material and Devices Laboratory
Allied-Signal Inc. Morristown, NJ 07960

Abstract - In this paper, the anharmonics of surface skimming bulk wave devices (SSBW) on rotated Y-cut quartz are analyzed by the two-dimensional theory developed by Lee, Syngellakis and Hou. The problem is solved in closed form taking into account the electrical and mechanical surface conditions. From this calculation, the dispersion curves of the horizontal shear mode are obtained. The applications of these dispersion curves to the bulk wave resonators and surface skimming bulk wave delay lines are compared and discussed. Delay lines have been fabricated and tested on ST-cut quartz with different wavelength-to-thickness ratios. Good agreement between experiment results and theoretical prediction has been obtained.

## Introduction

Depending on the orientation of the interdigital transducers(IDT) with respect to the crystal substrate, various wave modes can be generated and propagated in the rotated Y-cut quartz plates. When the waves are propagating in the x-direction, in addition to the Rayleigh surface wave, the interdigital transducer will generate longitudinal plate modes, shear vertical plate modes and shear horizontal plate modes [1,2]. When the waves are propagating in the z-direction, only the surface skimming bulk wave and shear horizontal plate modes are generated by the interdigital transducer. We may consider the surface skimming bulk wave mode as the fundamental shear horizontal plate mode.

In a regular surface acoustic wave delay line, different wave modes are launched by the IDT's. Some of these modes propagate into the bulk and reflect from the bottom of the crystal. Upon reflection, some of the energy can be coupled to a different mode. Mode conversion in the crystal makes it difficult to utilize those high frequency modes. Lewis pointed out that a clean plate mode response at a frequency up to at least six times that of the corresponding SAW device can be seen in a SSBW device because only the shear-horizontal mode is piezoelectricaly coupled to the IDTs[3].

One major attraction of these modes is that, for a given transducer period, it enables the operating frequency to be much higher than that of the fundamental SAW or SSBW modes[4-7]. How to design devices with the desirable center frequency is still a task of interest. These plate modes can also be utilized to form a natural comb filter. The practical concern is how to assure the frequency difference among passbands.

The device made by Lewis has a delay of 156 IDT periods and is on a plate of 85.3 periods thick. We notice that, in his device, as the frequency of the plate mode increases, the fractional bandwidth increases. We made a device with a delay of 52 IDT periods and fabricated it on a plate of only 1.6 periods thick. Responses of much higher frequency have been observed. But the higher the operating frequency the smaller the fractional bandwidth. Lewis used the ray approach to consider each plate mode separately. The plate mode may be generated by the first bounce or the second bounce. The first bounce can be considered as a plate of twice the thickness with an interdigital transducer on each side. The second bounce can be considered as a plate of four times the thickness with an interdigital transducer on each side. Both the group velocity and the phase velocity and their associated angles for each mode can be computed from the wave equations for plane waves traveling in infinite media. Once these are known, the resonant frequency of the mode and the time delay can be computed. When the thickness-to-period ratio becomes small, these modes shall be considered as modes traveling in a plate acoustic wave guide.

In the present paper, two-dimensional equations of motion of successively higher orders of approximation are applied to study these modes in SSBW delay lines[8,9]. Because either the face potential or the face charge might be specified as surface conditions, the frequency of these modes can be easily found without iteration for a given substrate thickness.

To understand the relationship between the operating frequency and the substrate thickness-to-IDT period ratio, we compute the dispersion curves of straight-crested wave propagation in the $x_3$ direction in a ST-cut quartz plate.

The frequency response of bulk wave resonators, surface skimming bulk wave delay lines and their anharmonics can all be analyzed from these dispersion curves. Several bulk wave and surface skimming bulk wave devices have been fabricated and tested. The close agreement between the prediction and testing results demostrate the usage of these approximate equations for further design of high frequency plate mode devices.

CH2588-2/88/0000-217 $1.00 C 1988 IEEE

## Analysis Method

The geometry under consideration is illustrated in Fig. 1. A piezoelectric plate has its length along the $x_3$-axis and thickness 2b along the $x_2$-axis, and the modal vibrations are taken to be independent of $x_1$.

In the two-dimensional approximate theory, the components of mechanical displacement and electric potential are expanded in an infinite series with the thickness dependence expressed by trigonometric functions

$$u_1 = \sum_{n=o}^{\infty} u_1^{(n)}(x_3,t) \cos \frac{n\pi}{2}\left(1 - \frac{x_2}{b}\right)$$

$$\phi = \sum_{}^{\infty} \phi^{(n)}(x_3,t) \cos \frac{n\pi}{2}\left(1 - \frac{x_2}{b}\right)$$

For clarity and ease of use, equations of zero-order and first-order approximation, which contain correction factor $\alpha_1$ and general equations of nth-order approximation, $(1 < n < N)$ are recapitulated and displayed in detail for the rotated Y-cut crystal with class 32 symmetry.

Zero-order equations

$$C_{55} u_{1,33}^{(o)} + e_{35} \phi^{(o)}_{,33} + \left(\frac{\alpha_1}{b}\right)\left(C_{56} u_{1,3}^{(1)} + e_{25} \phi^{(1)}_{,3}\right)$$

$$+ \frac{\pi}{4b} \sum_{m=3}^{N} mB_{mo} \left(C_{56} u_{1,3}^{(m)} + e_{25} \phi^{(m)}_{,3}\right)$$

$$+ \left(\frac{1}{2b}\right) F_1^{(o)} = \rho u_{1,tt}^{(o)}$$

$$e_{35} u_{1,33}^{(o)} - \varepsilon_{33} \phi^{(o)}_{,33} + \left(\frac{\alpha_1}{b}\right)\left(e_{36} U_{1,3}^{(1)} - \varepsilon_{32} \phi^{(1)}_{,3}\right)$$

$$+ \frac{\pi}{2b} \sum_{m=3}^{N} mB_{mo} \left(e_{36} u_{1,3}^{(m)} - \varepsilon_{32} \phi^{(m)}_{,3}\right)$$

$$+ \left(\frac{1}{2b}\right) D^{(o)} = o$$

First-order equations

$$C_{55} u_{1,33}^{(1)} + e_{35} \phi^{(1)}_{,33} - \left(\frac{2\alpha_1}{b}\right)\left(C_{56} u_{1,3}^{(o)}\right.$$

$$\left. + e_{36} \phi^{(o)}_{,3}\right) - \left(\frac{\pi}{2b}\right)^2 \left(C_{66} u_1^{(1)} + e_{26} \phi^{(1)}\right)$$

$$+ \frac{\pi}{2b} \sum_{m=1}^{N} [mB_{m1} \left(C_{56} u_{1,3}^{(m)} + e_{25} \phi^{(m)}_{,3}\right)$$

$$- B_{1m} \left(C_{56} u_{1,3}^{(m)} + e_{36} \phi^{(m)}_{,3}\right)]$$

$$+ \left(\frac{1}{b}\right) F_1^{(1)} = \rho u_{1,tt}^{(1)}$$

$$e_{35} u_{1,33}^{(1)} - \varepsilon_{33} \phi^{(1)}_{,33} - \frac{2\alpha_1}{b}\left(e_{25} u_{1,3}^{(o)} - \right.$$

$$\left. \varepsilon_{23} \phi^{(o)}_{,3}\right) - \left(\frac{\pi}{2b}\right)^2 \left(e_{26} u_1^{(1)} - \varepsilon_{22} \phi^{(1)}\right)$$

$$+ \left(\frac{\pi}{2b}\right) \sum_{m=3}^{N} [mB_{m1} \left(e_{36} u_{1,3}^{(m)} - \right.$$

$$\left. \varepsilon_{32} \phi^{(m)}_{,3}\right) - B_{1m} \left(e_{25} u_{1,3}^{(m)} - \varepsilon_{23} \phi^{(m)}_{,3}\right)]$$

$$+ \left(\frac{1}{b}\right) D^{(1)} = o$$

nth Order equations $(1 < n < N)$

$$C_{55} u_{1,33}^{(n)} + e_{35} \phi^{(n)}_{,33} - \left(\frac{n\pi}{2b}\right)^2 \left(C_{66} u_1^{(n)}\right.$$

$$\left. + e_{26} \phi^{(n)}\right) + \frac{\pi}{2b} \sum_{m=o}^{N} [m B_{mn} \left(C_{56} u_{1,3}^{(m)}\right.$$

$$\left. + e_{25} \phi^{(m)}_{,3}\right) - n B_{nm} \left(C_{56} u_{1,3}^{(m)} + e_{36} \phi^{(m)}_{,3}\right)]$$

$$+ \left(\frac{1}{b}\right) F_1^{(n)} = \rho u_{1,tt}^{(n)}$$

$$e_{35} u_{1,33}^{(n)} - \varepsilon_{33} \phi^{(n)}_{,33} - \left(\frac{n\pi}{2b}\right)^2 \left(e_{26} u_1^{(n)}\right.$$

$$\left. - \varepsilon_{22} \phi^{(n)}\right) + \frac{\pi}{2b} \sum_{m=o}^{N} [mB_{mn} \left(e_{36} u_{1,3}^{(m)}\right.$$

$$\left. - \varepsilon_{32} \phi^{(m)}_{,3}\right) - n B_{nm} \left(e_{25} u_{1,3}^{(m)} - \right.$$

$$\left. \varepsilon_{23} \phi^{(m)}_{,3}\right)] + \left(\frac{1}{b}\right) D^{(n)} = 0$$

where $\alpha_1 = \frac{2\sqrt{2}}{\pi}$

$$B_{mn} = \begin{cases} \frac{4m}{(m^2-n^2)\pi} & m + n = \text{odd} \\ 0 & m + n = \text{even} \end{cases}$$

Face Traction $F_1^{(n)} = T_{21}(b) - (-1)^n T_{21}(-b)$

Face Charge $D^{(n)} = D_2(b) - (-1)^n D_2(-b)$

In two-dimensional theory, the boundary conditions on the top and bottom of the plate are specified as part of the plate equations. The mechanical boundary conditions are related by the face tractions. The electric boundary conditions are related by the face charge.

The mechanical mass loading effect can be taken into account as

$$F_1^{(n)} = \bar{F}_1^{(n)} - 2b'\rho' u_{1,tt}$$

where $2b'$ is the thin film thickness and $\rho'$ is the film mass density.

The electrical effect on the surface can be specified for open conditions as

$$D^{(n)} = 0$$

or for short conditions as

$$D^{(n)} = \frac{2}{b} e_{26} u_1^{(n)}$$

$n = 1, 3, 5...$

## Analytical Solutions

The relationship between normalized frequency and normalized wavenumber can be found by substituting the displacement and potential in the equations of motion. For a given value of plate thickness, or a given value of normalized wavenumber, we can find the corresponding normalized frequency, or phase velocity of plate modes. The wavenumber can also be considered as the ratio of plate thickness and period of the interdigital transducer. Under the phase matching conditions, the wavelength of the bulk wave is the period times the cosine of the bulk wave propagating angle.

The matrices of the dispersion relations are symmetric and their elements may be classified into two groups: $A_{nn}$ for elements on the main diagonal; $B_{mn}$ for elements off the main diagonal. In terms of these groups the dispersion relations of the N-th order theory can be written as:

$$\begin{array}{c|cccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & \cdots & N \\ & & & & & m & & & \\ 0 & A_{00} & B_{01} & 0 & B_{03} & 0 & B_{05} & \cdots & \\ 1 & & A_{11} & B_{12} & 0 & B_{14} & 0 & \cdots & \\ 2 & & & A_{22} & B_{23} & 0 & B_{25} & \cdots & \\ 3 & & & & A_{33} & B_{34} & 0 & \cdots & \\ n\ 4 & & & & & A_{44} & B_{45} & \cdots & \\ \vdots & & & & & & & & \\ N & & & & & & & & A_{NN} \end{array} = 0$$

The terms in the groups are defined as follows ($0 < m, n < N$, for any $N > 0$):

$$A_{nn} = \begin{vmatrix} C_{55}Z^2 + n^2C_{66} - \Omega^2 & e_{35}Z^2 + n^2e_{26} \\ e_{35}Z^2 + n^2e_{26} & -\varepsilon_{33}Z^2 - n^2\varepsilon_{22} \end{vmatrix}$$

$$B_{nm} = \frac{16Z^2}{(m^2 - n^2)\pi^2} \begin{vmatrix} (m^2+n^2)C_{56} & m^2e_{25}+n^2e_{36} \\ m^2e_{25}+n^2e_{36} & -(m^2+n^2)\varepsilon_{23} \end{vmatrix}$$

$\Omega$ : Normalized Frequency

$Z$ : 2(Plate Thickness)/(Period of the IDT)

It shall be noted that the correction coefficient $\alpha_1$ is applied in terms of $A_{00}$, $A_{11}$, $B_{01}$, $B_{10}$ only.

## Theoretical and Experimental Results

The dispersion curves are computed and plotted in Fig. 2 for ST-cut Z-propagation SSBW delay lines. The circles correspond to the bulk wave resonator modes. The dotted line corresponds to the SSBW mode. The solid lines correspond to the SSBW anharmonics.

If we replace the interdigital transducers by simple electrodes, the device becomes a lateral field excited bulk wave resonator as shown in Fig. 3. The phase velocity in the $x_3$-direction approaches infinity; the group velocity in the $x_3$-direction approaches zero. The bulk waves generated by the electrodes are bouncing back and forth between the top and botton surfaces of the plate. No energy is propagating in the $x_3$-direction. Fig. 4 gives a typical frequency response of a lateral field excited bulk wave resonator. In addition to the main response, which corresponds to the circles in Fig. 2, there are several unwanted modes in the spectrum. These unwanted modes can be identified by imposing the edge boundary conditions and solving the plate equations. Here, we measured the first, third and fifth harmonics of the bulk wave resonator and compared them to the computed values.

| Measured (MHz) | Computed (MHz) | Difference |
|---|---|---|
| 2.968744 | 2.962975 | ( -0.19 % ) |
| 8.906231 | 8.852439 | ( -0.60 % ) |
| 14.843720 | 14.755393 | ( -0.59 % ) |

In this case, the frequency is determined by the thickness of the plate.

For the surface skimming bulk wave or the fundamental plate mode, the phase velocity is almost a constant as shown by the dotted line in Fig. 2. The mechanical displacement profile across the thickness can be considered as a constant as shown in Fig. 5. Fig.6 shows the frequency response of a surface skimming bulk wave delay line on a thin plate. We use different IDT periods on the plates with the same thickness. The measured and computed frequencies are listed below:

| Measured (MHz) | Computed (MHz) | Difference |
|---|---|---|
| 15.97 | 16.14 | ( 0.06 % ) |
| 47.91 | 47.03 | ( -1.84 % ) |
| 79.85 | 79.76 | ( -0.11 % ) |

In this case, the operating frequency depends on the period of IDT only.

For surface skimming bulk wave anharmonics or the high frequency plate modes, we may normalize the dispersion curves to the fundamental surface skimming bulk wave mode. We obtain a set of curves as shown in Fig. 7, and then relate the thickness-to-IDT period ratio to the operating frequency. Each of the anharmonics can be considered as a bulk wave travelling in the $x_3$-direction with different variation profile across the thickness as shown in Fig.8. Their phase velocity is a function of frequency. Point A in Fig. 6 is the first anharmonic and point B is the second anharmonic. When the thickness-to-IDT period ratio is small, the separation between each mode is large. When the thickness-to-IDT period ratio becomes big, the frequencies of anharmonics become closer. Fig.9, 10, 11 are the experimental frequency response of surface skimming bulk wave devices with a different thickness-to-IDT period ratio. Fig. 12, 13, 14 are the mode spectra predicted by the two-dimensional plate equations. Fig. 9 and 12 have a thicknessto-IDT period ratio of 16; Fig. 10 and 13 have a thickness-to-IDT period ratio of 8 and Fig. 11 and 14 have a thickness-to-IDT period ratio of 2. As we can see, the computed results and experimental results agree well.

Fig. 15 gives a close-up of one of the devices. At the frequency near to the surface skimming bulk wave mode, the anharmonics are very close to each other. The predicted mode spectrum is given in Fig. 16. Again, they agree with each other. The frequencies of each mode are listed below:

| Measured (MHz) | Computed (MHz) | Difference |
|---|---|---|
| 79.848 | 79.760 | ( .110% ) |
| - | 80.072 | ( - ) |
| - | 80.286 | ( - ) |
| 80.792 | 80.643 | ( .185% ) |
| 81.148 | 81.140 | ( .010% ) |
| 81.916 | 81.774 | ( .174% ) |
| 82.472 | 82.542 | ( -.085% ) |
| 83.056 | 83.441 | ( -.462% ) |
| 84.340 | 84.467 | ( -.150% ) |
| 85.532 | 85.614 | ( -.096% ) |
| 86.816 | 86.878 | ( -.072% ) |
| 88.328 | 88.254 | ( .083% ) |
| 89.844 | 89.738 | ( .119% ) |
| 91.532 | 91.322 | ( .229% ) |
| 93.156 | 93.004 | ( .164% ) |
| 95.024 | 94.776 | ( .261% ) |
| 96.608 | 96.635 | ( -.028% ) |
| 98.324 | 98.576 | ( -.255% ) |
| 100.324 | 100.593 | ( -.267% ) |
| 102.476 | 102.682 | ( -.201% ) |
| 104.692 | 104.839 | ( -.140% ) |
| 106.880 | 107.060 | ( -.168% ) |
| 108.948 | 109.341 | ( -.359% ) |

In this device, the first two anharmonics are smeared into the fundamental surface skimming bulk wave mode. In general, the agreement between the experimental and the theoretical results is less than 0.5 percent.

## Conclusions

1. The two-dimensional plate equations are used to compute the frequencies of anharmonics in the surface skimming bulk wave delay lines.

2. In the thin plates, instead of the reflected bulk wave modes, the anharmonics become the predominant modes generated by the interdigital transducers.

3. In the bulk wave resonator, the frequency is determined by the thickness-to-wavelength ratio; in the surface wave or surface skimming bulk wave devices, the frequencies are determined by the periods of the interdigital transducers; and in the plate mode devices, the frequencies are determined by both the period of the IDT and thickness-to-IDT period ratio of the plate.

4.Comb filters can be realized by using plates with small thickness-to-IDT period ratio.

## Acknowledgements

The author wishes to acknowledge the fabrication skills of David D. Badding of Allied-Signal Inc. He also values the fruitful discussions with Dr. C. S. Lam of AT&T Bell Laboratories.

## References

1. Hou, J.: "Approximation method for plate modes in surface acoustic wave devices" Proc. 41st Annual Symposium on Frequency Control, May 1987, pp. 365-370

2. Hou, J., van de Vaart, H.: "Mass sensitivity of plate modes in surface acoustic wave devices and their potential as chemical sensors" Proc. of the 1987 IEEE Ultrasonics Symposium pp. 573-578

3. Lewis, M. F.:"High frequency acoustic plate mode device employing interdigital transducers" Electronics Letters Oct. 1981 Vol. 17 No. 21 pp. 819-821

4. Goodberlet, M. A., Lee, D. L.:"Reflected bulk wave delay lines" Proc. of the 1983 IEEE Ultrasonics Symposium pp. 548-553

5. Goodberlet, M. A., Lee, D. L.:"The excitation and detection of surface generated bulk waves" IEEE Trans. on Sonics and Ultrasonics Vol. SU-31, No.2. March 1984 pp. 67-75

6. Josse, F., Vetelino, J. F., Soluch, W., Andle J.C."On the reflection of bulk acoustic wave in rotated Y-cut quartz" Proc. of the 1985 Ultrasonic Symposium. pp. 319-324

7. Elias, E., Jhunjhunwala, A.:"Analysis of planar bulk acoustic wave devices"Electronics Letters Feb. 1988 Vol. 24 No.4 pp. 240-242

8. Lee, P. C. Y., Syngellakis, S. Hou, J.:"A two-dimensional theory for high frequency vibration of piezoelectric crystal plates with or without electrodes" J. Appl. Phys. 61(4) 15. Feb. 1987 pp.1249-1262

9. Mindlin, R. D.:"Thickness-twist vibrations of an infinite monoclinic crystal plate" Int. J. Solid Structures, 1965 Vol.1 pp. 141-145

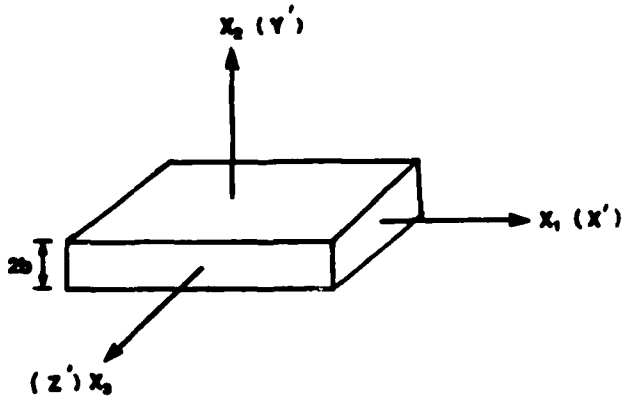



Figure 1. Plate coordinates

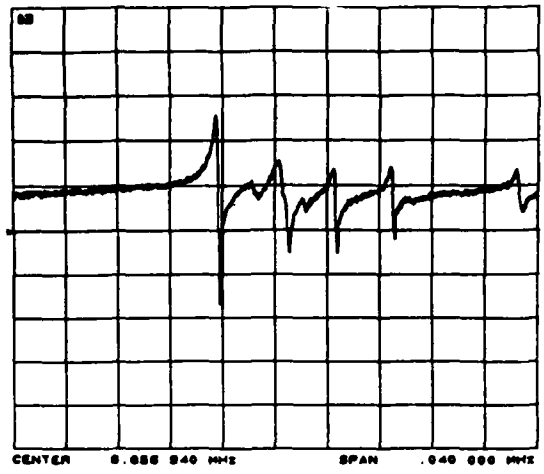



Figure 4. Frequency response of a lateral field excited bulk wave resonator

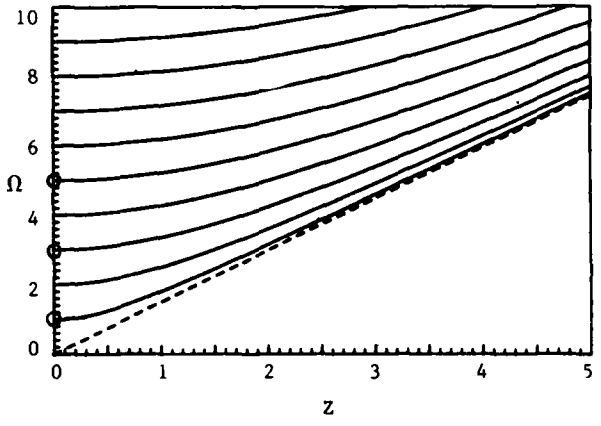



Figure 2. Dispersion curves for waves in x3-direction in a ST-cut quartz plate

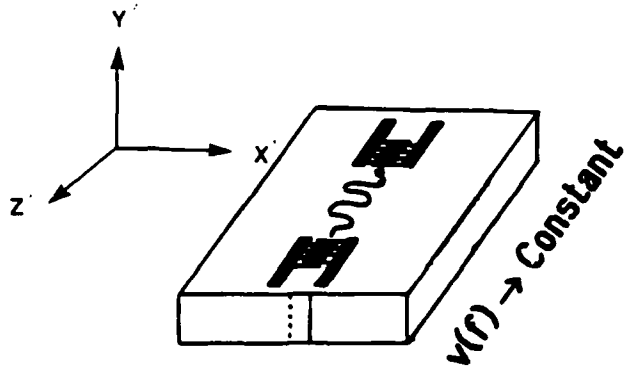



Figure 5. Surface skimming bulk wave devices

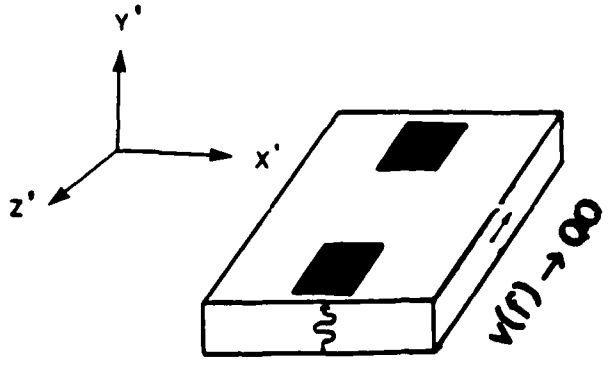



Figure 3. Lateral field excited bulk wave resonator

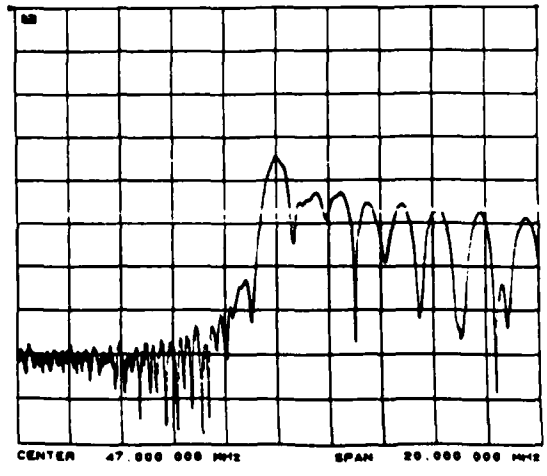



Figure 6. Frequency response of a surface skimming bulk wave delay line

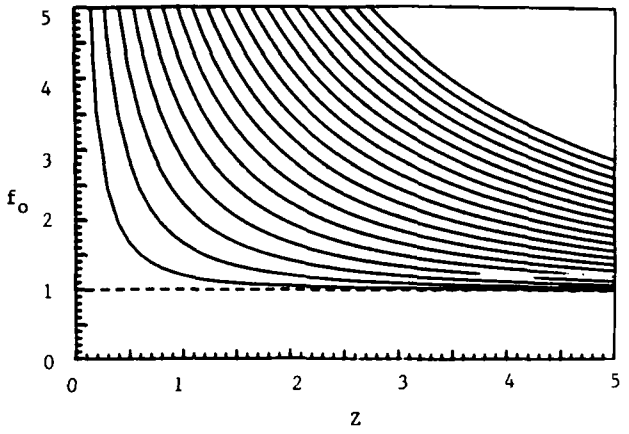



Figure 7. Frequency spectrum of surface skimming bulk wave anharmonics

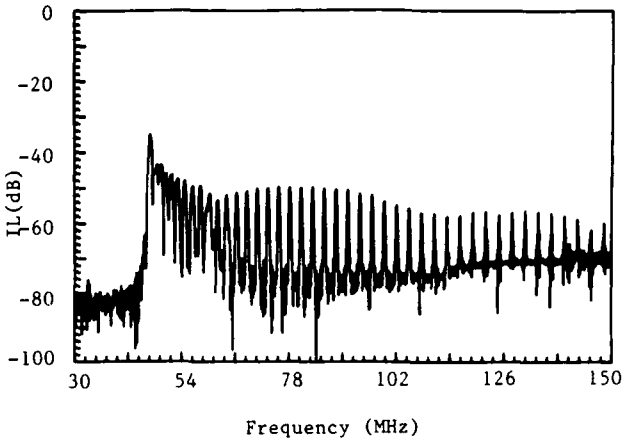



Figure 10. Frequency response of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 8

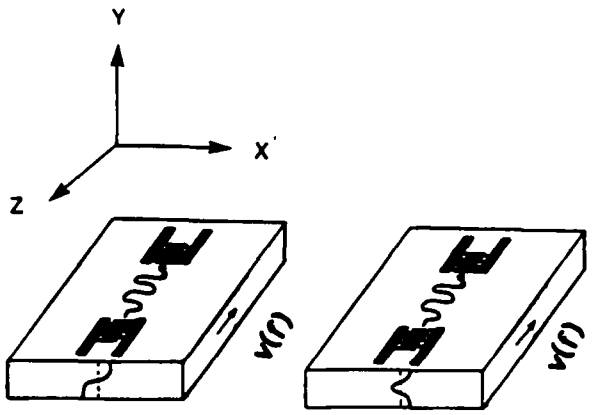



Figure 8. Anharmonics in surface skimming bulk wave devices

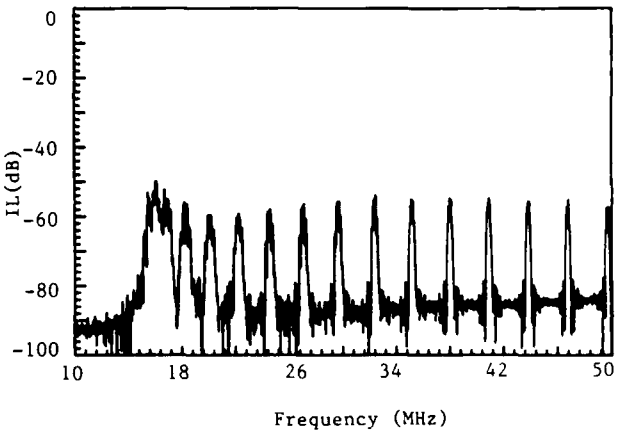



Figure 11. Frequency response of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 2

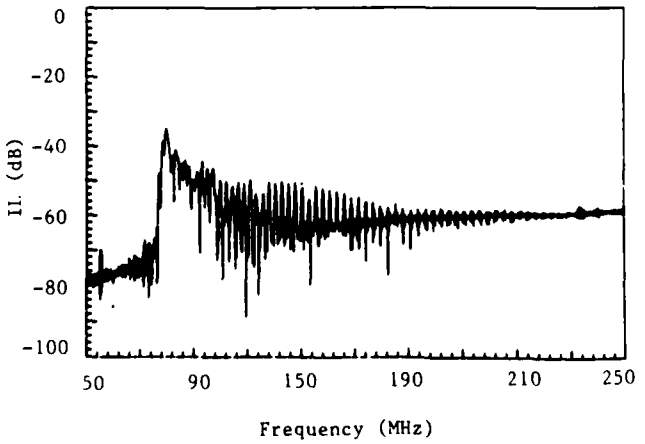



Figure 9. Frequency response of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 16

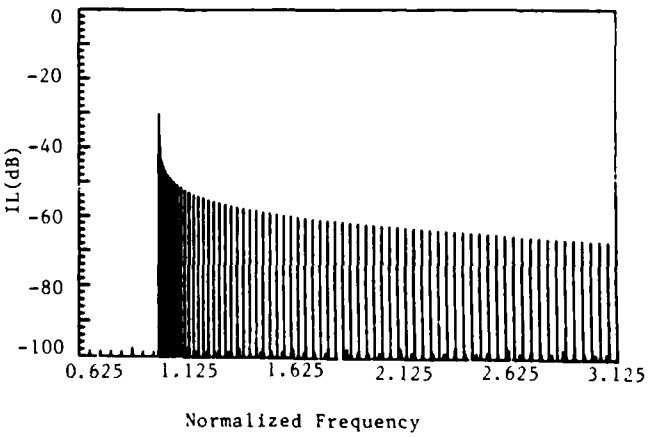



Figure 12. Mode spectrum of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 16

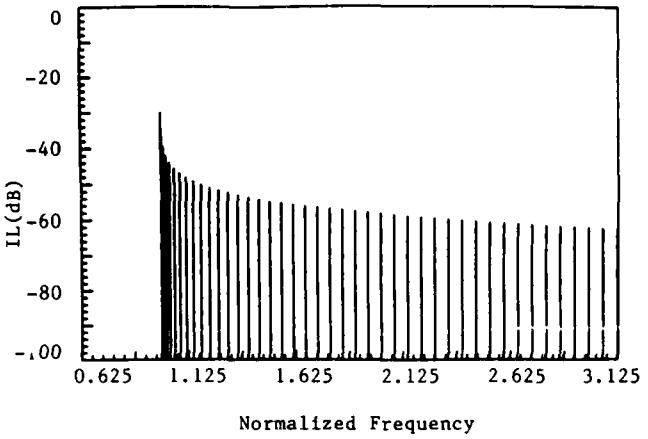



**Figure 13.Mode spectrum of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 8**

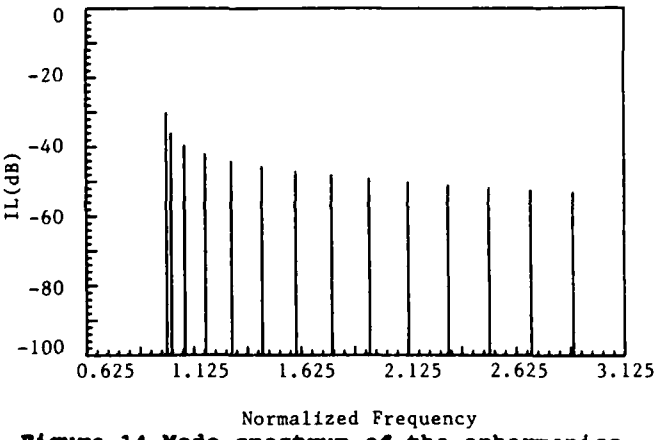



**Figure 14.Mode spectrum of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 2**

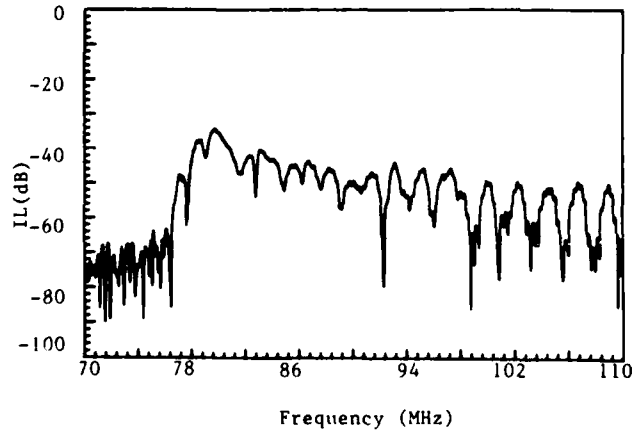



**Figure 15.Frequency response of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 16**

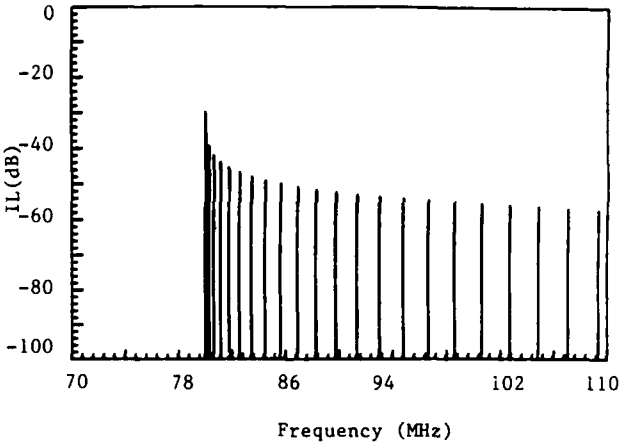



**Figure 16.Mode spectrum of the anharmonics in a surface skimming bulk wave delay line with thickness-to-IDT period ratio of 16**

42nd Annual Frequency Control Symposium - 1988

# EXPERIMENTAL TEMPERATURE AND STRESS SENSITIVITIES OF SURFACE ACOUSTIC WAVES QUARTZ CUTS

E. Bigler, R. Coquerel, 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 sensitivity of surface acoustic wave devices to quasi-static in-plane stresses has been determined in the case of thin quartz crystal plates, as a function of crystal anisotropy. Stress sensitivity is compared to static temperature sensitivity calculated by the perturbation method proposed by Tiersten. Several orientations of quartz crystal exists, exhibiting both zero temperature sensitivity and zero stress sensitivity for planar isotropic stresses. Two interesting singly rotated quartz crystal cuts ($\Phi = 0$) have been theoretically determined, and one experimentally tested. First-order temperature compensation was checked and stress sensitivity was measured by a diametral compression of circular quartz delay lines. A comparison between experimental results and theoretical predictions is discussed.

## Introduction

Temperature[1,2] and forces[3,4] effects are the main sources of instability of surface acoustic wave (SAW) devices. Static temperature effects may be avoided by the choice of a temperature-compensated cut like the classical ST-cut. The effects of forces depending closely on the geometry of the device (shape and fixation conditions), one may think to study this problem for each particular device. Another way is to consider the effects of stresses as the tensorial equivalent of temperature effects. In this approach,[5,6] stress sensitivity may be defined whatever the origin (thermoelastic effects, external forces) might be, introducing six independent stress coefficients ${}^{s}\alpha_{ij}$ according to the relation

$$\frac{\Delta V}{V_o} = \sum_{i,j=1}^{3} {}^{s}\alpha_{ij}\,\overline{T_{ij}} \qquad (1)$$

In eq. (1), $\Delta V/V_o$ represents the relative velocity shifts of the SAW induced by a quasi-static mechanical bias $T_{ij}$.

Using crystal anisotropy, it is possible to look for particular crystallographic orientations and propagation directions minimizing the velocity shifts for a given set of stress components $T_{ij}$. We shall consider here the case of a thin plate of quartz crystal submitted to **planar isotropic stresses** (this configuration of stresses appears in the case of thermoelastic effects).

Under these assumptions, it will be shown that several quartz cuts are both temperature- and stress-compensated, defining for SAWs the equivalents of the SC-cut[7] of bulk waves.

This work is sponsored by RADC/EEAC, Hanscom Field AFB under AFOSR grant # 87-0241

CH2588-2/88/0000-224 $1.00 C 1988 IEEE

## Theoretical temperature and stress-compensated quartz cuts

Stress and temperature coefficients are defined using the formalism developed by Tiersten for wave propagation in a prestrained medium.[8]

Considering a SAW propagating in a very small area (with respect to the dimensions of the plate) and thus submitted to an homogeneous distribution of mechanical or thermal bias, the relative velocity shifts of the SAW may be expressed using the perturbation method[4,9,10] :

a - Case of a thermal bias in a free-expanding plate

$$\frac{\Delta V}{V_o} = {}^{\theta}\alpha\,(\theta - \theta_o) \qquad (2)$$

where $(\theta-\theta_o)$ is the temperature elevation with respect to the reference state and ${}^{\theta}\alpha$ is the first-order temperature coefficient defined as

$${}^{\theta}\alpha = \{C_{iskurv}\,\ell_{uv} + C_{qskr}\,\ell_{iq} + C_{isqr}\,\ell_{kq} + \frac{d}{d\theta}(C_{iskr})\}\,U_{iskr} \qquad (3)$$

where $\ell_{uv}$ are the coefficients of linear expression
$C_{qskr}$ are the 2nd order stiffness constants
$C_{iskruv}$ are the 3rd order elastic constants
$d/d\theta\,(C_{iskr})$ are the first derivatives of the fundamental elastic constants
$U_{iskr}$ depends only on the propagation terms of the SAW on the unperturbed crystal.

b - Case of a mechanical bias at constant temperature

$$\frac{\Delta V}{V_o} = {}^{s}\alpha_{pq}\,\overline{T_{pq}} \qquad (4)$$

where $\overline{T_{pq}}$ are the components of the quasi-static stress tensor and
${}^{s}\alpha_{pq}$ the stress coefficients defined as

$${}^{s}\alpha_{pq} = \{\delta_{ik}\delta_{sp}\delta_{rq} + C_{iskruv}\,s_{uvpq} + C_{\ell skr}\,s_{i\ell pq} + C_{is\ell r}\,s_{k\ell pq}\}\,U_{iskr} \qquad (5)$$

where $\delta_{ik}$ is the Kronecker delta
$s_{uvpq}$ are the 2nd order compliance constants.

In the case of a thin plate of quartz crystal submitted to in-plane, isotropic stresses, eq.(4) takes the simpler form

$$\frac{\Delta V}{V_o} = ({}^{s}\alpha_{11} + {}^{s}\alpha_{33})T \qquad (6)$$

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 ${}^s\alpha_{11} + {}^s\alpha_{33}$ for planar isotropic stresses has been computed for quartz, in the non-piezoelectric approximation, as a function of the cut angles ($\Phi,\theta$) and propagation direction $\Psi$ corresponding to a (YXwℓt)$\Phi\theta\Psi$ plate, [11] SAW propagation in the ℓ-direction.

Taking the angle $\Psi$ as a parameter, the loci of stress-compensated cuts are defined in the ($\theta,\Phi$) plane by the condition

$$({}^s\alpha_{11} + {}^s\alpha_{33})(\Phi, \theta, \Psi) = 0 \qquad (7)$$

These loci are represented in solid lines on Fig. 1 for different propagation directions : $\Psi = 27°$ (Fig. 1a) $\Psi = 30°$ (Fig. 1b) $\Psi = 36°$ (Fig. 1c) $\Psi = 45°$ (Fig. 1d). On the same graphs are represented in dotted lines the loci of zero first-order temperature coefficients.

The intersections between the two loci define quartz cuts satisfying both conditions of stress and temperature compensation.

As a first criterium to choose between the different intersection points shown on Fig. 1, it is interesting to look for singly rotated quartz cuts ($\Phi = 0$), since machining tolerances are easier to obtain than for doubly rotated cuts.

The evolution of the curves shows that there are two possibilities of intersection at $\Phi = 0$

- for $\Psi$ close to 30° (point P30 on Fig. 1b)
- for $\Psi$ close to 45° (point P45 on Fig. 1d).

This last cut (point P45 : $\Phi = 0_{12}\theta = 40°$, $\Psi \simeq 45°$) is close to B.K. Sinha's STC cut [12] ($\Phi = 0$, $\theta = 41{,}8°$, $\Psi = 46{,}9°$).

This cut is compensated for planar isotropic stresses, and the angular discrepancy may be explained by the fact that we use a non-piezoelectric approximation for stress and temperature effects.

The first cut corresponding to the point P30 has been tested experimentally, as exposed now.

## Experimental method

Experiments are performed on quartz delay lines built on circular plates (diameter : 22 mm, thickness : 1 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 $\Psi$ was adjusted to keep the inversion point of the frequency/temperature curve between 20-25 °C (Fig. 2).

### In-plane stress effects

A delay line satisfying the condition of temperature compensation is tested with a mechanical system represented on Fig. 3.

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 $\beta$ between the propagation direction of the SAW and the direction of the force (Fig. 4).

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 5N 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.

## Experimental results

### Temperature effects

Four singly-rotated cuts were tested for cut angles $\theta = -2.5°$, $-5°$, $-7°$, $-10°$ and the propagation angles $\Psi$ corresponding to an inversion point at room temperature are 32.5°, 31.5°, 31°, 30° respectively (Fig. 3). As a reference the theoretical curve ($34 \times 10^{-9}$ ppm.K$^{-2}$) 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.

### Planar stress effects

The force-frequency effect in a circular disk may be derived from the stress sensitivity coefficients (eqs 5 and 6) and from the classical theory of elasticity. [13] In the case of a single diametral compression, the relative frequency shifts are given by

$$\frac{\Delta F}{F_o} = \frac{-2P}{\pi dt}\left(({}^s\alpha_{11} + {}^s\alpha_{33}) + 2\,({}^s\alpha_{11} - {}^s\alpha_{33})\cos 2\beta - 2\,{}^s\alpha_{13}\sin 2\beta\right) \qquad (8)$$

where

P is the applied force

d is the diameter and t the thickness of the disk

${}^s\alpha_{11}$, ${}^s\alpha_{33}$, ${}^s\alpha_{13}$ the in-plane stress coefficients

$\beta$ the azimuthal angle between the SAW and the force.

Eq. (8) shows a sinusoïdal dependence with respect to the angle $\beta$, with an offset value $-2P/\pi dt$ $({}^s\alpha_{11}+{}^s\alpha_{33})$ corresponding to the case of a symmetrical compression with two orthogonal forces of magnitude P/2.

Fig. 4 shows a comparison between experimental data and the theoretical curves computed from the following values (Table I).

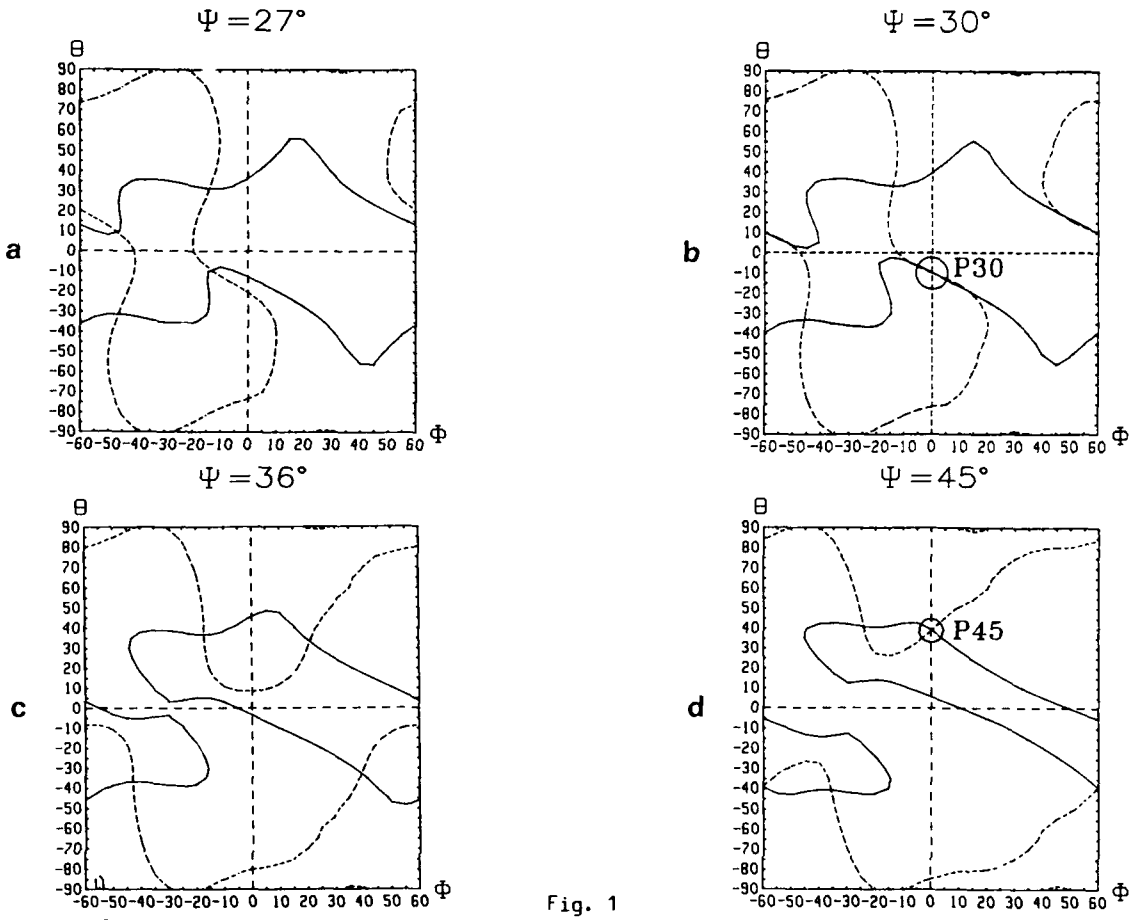



Fig. 1

- 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 (YXwℓt)Φθψ plate, SAW propagating in the ℓ-direction

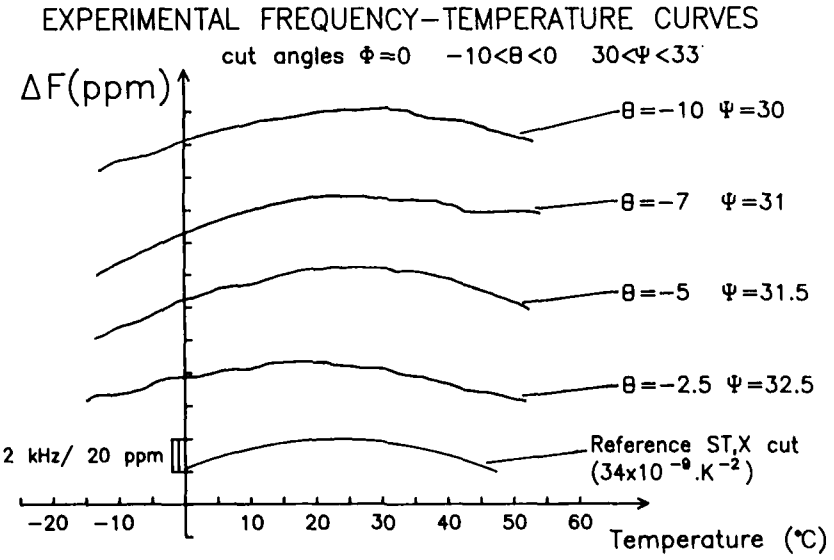



Fig. 2

Experimental static temperature characteristics for the cuts Φ = 0, $-10^{\circ} < \theta < -2.5^{\circ}$, $30^{\circ} < \Psi < 33^{\circ}$

Comparison with the classical ST,X cut

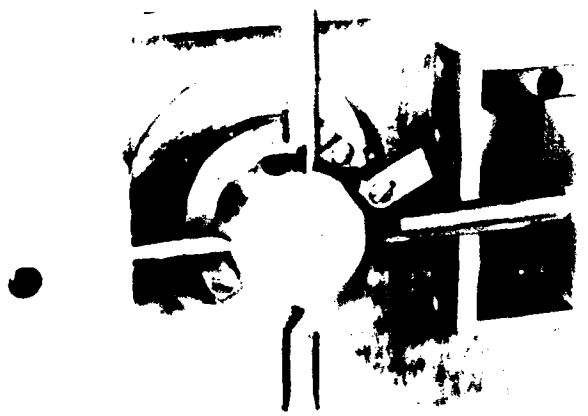
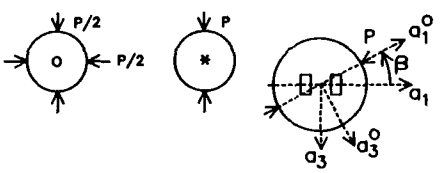


Fig. 3
Experimental set-up for single and symmetrical diametral compression
of a circular quartz delay line



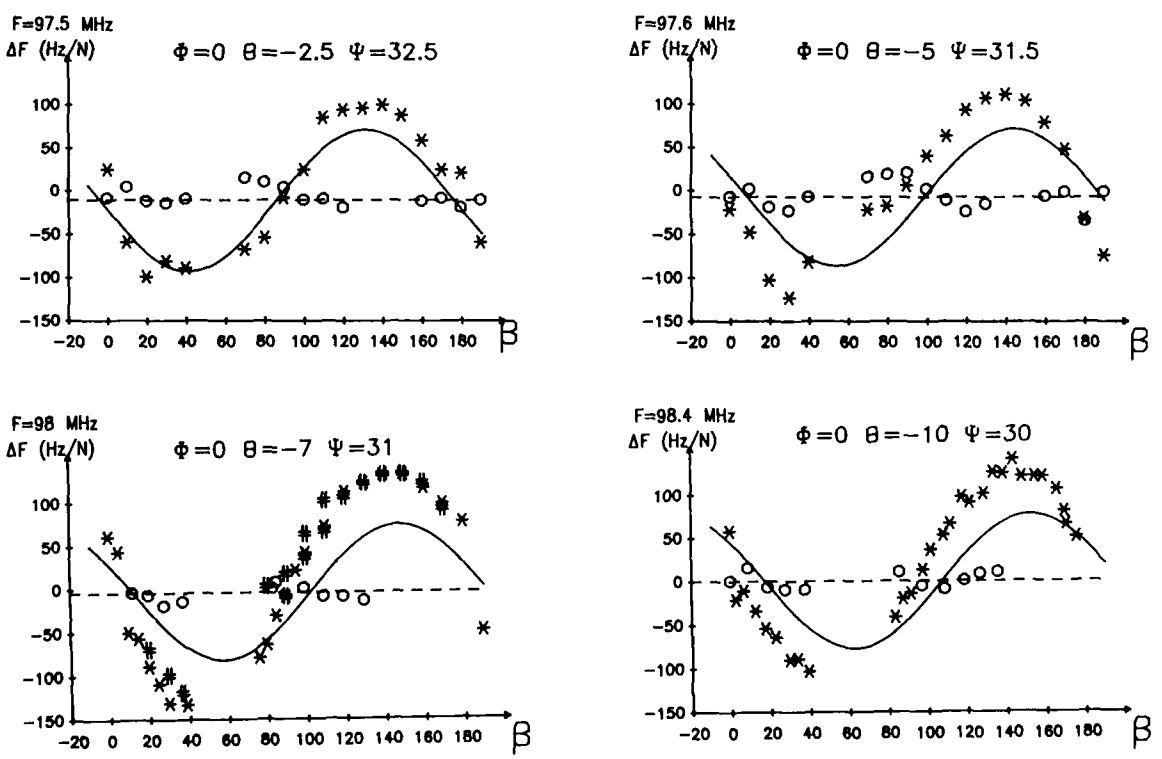




Fig. 4
Experimental force-frequency effect on temperature-compensated circular quartz delay lines
with a low sensitivity to planar isotropic stresses :
$\Phi = 0 \quad -10^\circ < \theta < 2.5^\circ \quad 30^\circ < \Psi < 33^\circ$
single diametral compression : * represents experimental values
solid sinusoïdal curve : theoretical values
double diametral compression : o represents experimental values
dashed horizontal line : theoretical values

| cut angles ($\Phi$ = 0) | | Theoretical values of stress coefficients units : $10^{-12}$ $m^2.N^{-1}$ | | | Theoretical values of force-frequency effects $\Delta F/F$ ($\beta$) (diameter 22 mm - thickness 1 mm) units : $10^{-8}$ $Hz.N^{-1}$ |
|---|---|---|---|---|---|
| $\theta$ | $\Psi$ | $s_{\alpha_{11}}$ | $s_{\alpha_{33}}$ | $s_{\alpha_{13}}$ | |
| - 2.5 | 32.5 | + 0.8 | + 2.9 | - 14.3 | - 10.9 - 11.8 cos2$\beta$ - 83.6 sin2$\beta$ |
| - 5 | 31.5 | - 0.7 | + 3.4 | - 13.3 | - 7.7 + 24.0 cos2$\beta$ - 77.2 sin2$\beta$ |
| - 7 | 31 | - 2.0 | + 3.6 | - 12.5 | - 4.4 + 32.6 cos2$\beta$ - 72.2 sin2$\beta$ |
| - 10 | 30 | - 3.8 | + 3.8 | - 11.1 | 0.1 + 45 cos2$\beta$ - 64.3 sin2$\beta$ |

Table I
Theoretical values of force-frequency effects in the angular range
($-10 < \theta < -2.5$) ($30 < \Psi < 33$)

## Discussion

- Temperature-compensated cuts were experimentally found close to the theoretical values ($|\Psi_{theory} - \Psi_{exp}| < 0.5°$) given by the non-piezoelectric model (Fig. 5)

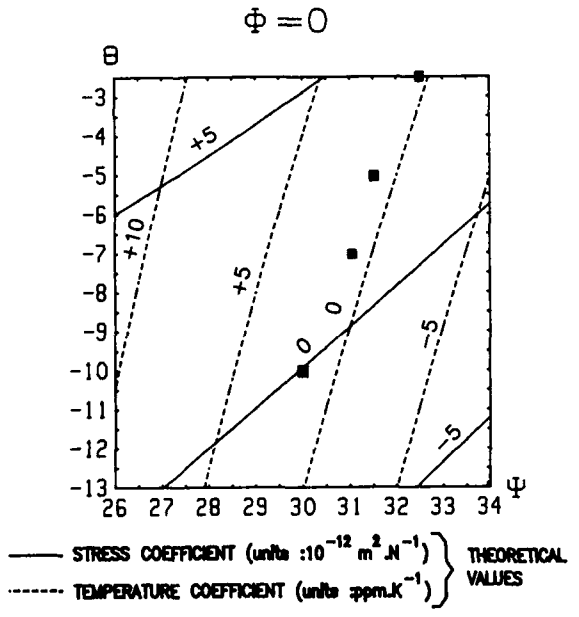



Fig. 5
Contour line map of temperature and stress coefficients (non-piezoelectric approximation)
cut angles : $\Phi = 0$, $-13° < \theta < -2.5°$
propagation direction $26° < \Psi < 34°$
comparison with experimental temperature-compensated cuts

- Stress compensation effects are clearly shown when the circular delay lines are submitted to a pair of orthogonal forces of the same magnitude. But in the case of a single diametral compression, the experimental points do not exactly match the theoretical sinusoïdal curve. The reasons of this discrepancy might be either a systematic error due to parasitic stresses induced by the experimental set-up, or due to the fact that the dimensions of the delay line ($9.5 \times 4.5$ $mm^2$) are not small with respect to the diameter of the plate (22 mm). Thus, it would be necessary to use a model taking into account non-homogeneous distributions of stresses and strains.

Anyway the cut ($\Phi = 0$ $\theta = -10$ $\Psi = 30$) is interesting since the sensitivity to planar isotropic stresses is significatively reduced with respect to the classical ST,X cut, as summed up below.

| | ST,X[12] $\Phi$=0 $\theta$=42.75 $\Psi$=0 | $\Phi$=0 $\theta$=-10° $\Psi$=30° |
|---|---|---|
| SAW velocity (m/s) | 3 158 | 3 396 |
| Coupling factor $\Delta V/V$ | $5.8 \times 10^{-4}$ | $6 \times 10^{-4}$ [14] |
| Turnover temp. | 23 °C | 25 °C |
| First order temp. coefficient | 0 | 0 |
| 2nd order temp. coefficient | $34 \times 10^{-9}$ $K^{-2}$ | $30 \times 10^{-9}$ $K^{-2}$ |
| sensitivity to planar isotropic stresses * | $s_{\alpha_{11}}+s_{\alpha_{33}} = -12\times10^{-12}$ $m^2.N^{-1}$ | $\lvert s_{\alpha_{11}}+s_{\alpha_{33}}\rvert < 3\times10^{-12}$ $m^2.N^{-1}$ |

Table II

* for the ST cut, the value of $s_{\alpha_{11}}+s_{\alpha_{33}}$ is calculated by the model of stress sensitivity coefficients, and is in good agreement with theoretical and experimental values given by Sinha et al.[4]
For the cut ($\Phi = 0$ $\theta = -10$ $\Psi = 30$) the given value is experimental.

## Conclusion

Stress and temperature coefficients have been used to determine singly rotated quartz cuts with low temperature sensitivity and low sensitivity to in-plane isotropic stresses. Two interesting cuts were determined ($\theta = -10$, $\phi = 30$) and ($\theta = 40°$, $\Psi = 45°$).

Experimental tests were performed in the angular range ($-10 < \theta < -2.5$, $30 < \Psi < 33$) and temperature compensated cuts were found in good agreement with theoretical previsions. Stress compensation was observed on these cuts in the case of symmetrical diametral compression. The upper limit of sensitivity to planar isotropic stresses for the cut ($\theta \simeq -10$, $\Psi = 30$) was estimated experimentally four times lower than the classical ST,X cut (the coupling factor and 2nd order temperature coefficients are very similar to the ST cut).

Comparison with the cut ($\theta = 40^\circ$, $\Psi = 45^\circ$) is presently under test in order to find the best trade-off between coupling factor, 2nd order temperature coefficient, angular machining tolerances.

Next step will be to investigate dynamical temperature effects to confirm by a different method the potentialities of stress compensated SAW quartz cuts.

## Acknowledgements

The authors are very grateful to C. Bonjour for SAW delay lines fabrication.

## References

1. H.F. Tiersten and B.K. Sinha, "Temperature derivatives of the fundamental elastic constants of quartz", Proc. 32nd Ann. Freq. Cont. Symp., Atlantic-City (N.J.), pp. 150-154, 1978.

2. D. Hauden, M. Michel and J.J. Gagnepain, "Higher order temperature coefficients of quartz SAW oscillators", Proc. 32nd Ann. Freq. Cont. Symp., Atlantic-City (N.J.), pp. 77-86, 1978.

3. 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, pp. 342-348, 1981.

4. B.K. Sinha, W.J. Tanski, T. Lukaszek, A. Ballato, "Influence of biasing stresses on the propagation of surface waves", J. Appl. Phys., vol. 57, pp. 767-776, 1985.

5. E. Bigler, G. Théobald and D. Hauden, "SAW quartz cuts with low stress and temperature sensitivity", Electron. Lett., vol. 23, pp. 514-516, 1987.

6. E. Bigler, G. Théobald and D. Hauden, "Stress-sensitivity mapping for surface acoustic waves on quartz", to be published in IEEE Trans. Ferro. Freq. Cont.

7. E.P. Eernisse, "Calculations on the stress-compensated (SC-cut) quartz resonator", Proc. 30th Ann. Freq. Cont. Symp., pp. 8-11, 1976.

8. H.F. Tiersten and J.C. Baumhauer, "Nonlinear electroelastic equations for small fields superposed on a bias", J. Acoust. Soc. Am., vol. 54, p. 1017, 1973.

9. H.F. Tiersten and B.K. Sinha, "A perturbation analysis of the attenuation and dispersion of surface waves", J. Appl. Phys., vol. 49, p. 87, 1978.

10. B.K. Sinha and H.F. Tiersten, "On the temperature dependence of the velocity of surface waves in quartz", J. Appl. Phys., vol. 51, pp. 4659-4665, 1980.

11. "1949 IRE Standard on Piezoelectric crystals", Proc. IRE, p. 1378 (1949).

12. B.K. Sinha, "A stress and temperature compensated orientation and propagation direction for surface acoustic wave devices", IEEE Trans. Ultrason., Ferro. and Freq. Cont., vol. UFFC-34, pp. 64-74, 1987.

13. S.P. Timoshenko and J.W. Goodier, in "Theory of Elasticity", p. 122, 3rd edition (International Student Edition), McGrawHill, ed. (1970).

14. M. Feldmann and J. Henaff, Revue de Physique Appliquée, vol. 12, pp. 1775-1778, 1977.

42nd Annual Frequency Control Symposium - 1988

# AN ANALYSIS OF THE IN-PLANE ACCELERATION SENSITIVITY OF ST-CUT QUARTZ SURFACE WAVE RESONATORS WITH THE SUBSTRATE EXTENDING BEYOND THE SUPPORTS

D.V. Shick and H.F. Tiersten
Department of Mechanical Engineering,
Aeronautical Engineering & Mechanics
Rensselaer Polytechnic Institute
Troy, New York 12180-3590

## Abstract

An analysis of the in-plane acceleration sensitivity of ST-cuts quartz surface wave resonators with the substrate extending beyond the supports is performed Since with this support configuration a small amount of flexure occurs in the active substrate for the case of in-plane acceleration, the dominant extensional biasing state is found first and used to conservatively define the flexural biasing state. Since the influence of the conservative estimate of the flexural biasing state is quite small, this is an appropriate procedure. The calculated results show that very low (less than a few parts in $10^{11}$) in-plane acceleration sensitivities can be obtained, and since the normal acceleration sensitivity can readily be made to vanish, very low resultant acceleration sensitivities can be achieved with this support configuration.

## 1. Introduction

It has recently been shown[1] that the normal acceleration sensitivity of ST-cut quartz surface wave resonators with the substrate extending beyond rectangular supports vanishes for certain planar geometric aspect ratios, as it does in the simply- and rigidly-supported cases[2,3]. It has further been shown[1] that in the vicinity of the zero-crossings the slope of the sensitivity versus aspect ratio curve for the configuration with the overhangs is much shallower than that for the case of rigid supports[3]. Consequently, with this support configuration, which is the one employed in the work at Raytheon[4,5], very low normal acceleration sensitivities can be obtained over a reasonably wide range of aspect ratio. This has recently been verified experimentally[6]. However, for the case of in-plane acceleration with this support configuration, a small amount of flexure is induced in the active substrate, which serves to increase the in-plane acceleration sensitivity somewhat.

In this work an analysis of the in-plane acceleration sensitivity of this support configuration is performed. The biasing state is determined by decomposing the problem into an extensional and flexural part. First the extensional biasing state is determined by means of the new variational approximation procedure[2,3], which is very accurate and convenient to use. The resulting solution for the extensional biasing state defines the associated flexural problem. The flexural biasing state is obtained by solving the equation for anisotropic static flexure, which is possible because for the flexural part the relatively unimportant corner regions are left out of account. The resulting extensional and flexural biasing states are employed in the existing perturbation integral[7] along with the proper continuous representation of the acoustic surface wave mode shape[8] to calculate the in-plane acceleration sensitivity for this support configuration. Results are presented for a range of planar aspect ratio and the calculations show that the resultant in-plane, and, hence, absolute resultant, acceleration sensitivity can be kept below a few parts in $10^{11}$ per g.

## 2. Perturbation Equations

For purely elastic nonlinearities the equation for the first perturbation of the eigenvalue obtained from the perturbation analysis[7] mentioned in the Introduction may be written in the form

$$\Delta_\mu = H_\mu/2\omega_\mu, \quad \omega = \omega_\mu - \Delta_\mu, \tag{2.1}$$

where $\omega_\mu$ and $\omega$ are the unperturbed and perturbed eigenfrequencies, respectively, and

$$H_\mu = -\int_V \hat{K}^n_{L\gamma} g^\mu_{\gamma,L}\, dV, \tag{2.2}$$

where V is the undeformed volume of the piezoelectric plate. In (2.2) $g^\mu_\gamma$ denotes the normalized mechanical displacement vector, and $\hat{K}^n_{L\gamma}$ denotes the portion of the Piola-Kirchhoff stress tensor resulting from the biasing state in the presence of the $g^\mu_\gamma$, and is given by

$$\hat{K}^n_{L\gamma} = \hat{c}_{L\gamma M\alpha} g^\mu_{\alpha,M}, \tag{2.3}$$

where

$$\hat{c}_{L\gamma M\alpha} = T^1_{LM}\delta_{\gamma\alpha} + c_{3L\gamma M\alpha KN}E^1_{KN} + c_{2L\gamma KM}w_{\alpha,K} + c_{2LKM\alpha}w_{\gamma,K} \tag{2.4}$$

and

$$T^1_{LM} = c_{2LMKN}E^1_{KN}, \quad E^1_{KN} = \frac{1}{2}(w_{K,N} + w_{N,K}). \tag{2.5}$$

The quantities $T^1_{LM}$, $E^1_{KN}$ and $w_K$ denote the static biasing stress, strain and displacement field, respectively. Thus, in this description the present position $\underset{\sim}{y}$ is related to the reference position $\underset{\sim}{X}$ by

$$\underset{\sim}{y}(X_L,t) = \underset{\sim}{X} + \underset{\sim}{w}(X_L) + \underset{\sim}{u}(X_L,t). \tag{2.6}$$

The coefficients $c_{2LMKN}$ and $c_{3L\gamma M\alpha KN}$ denote the second and third order elastic constants, respectively.

The normalized eigensolution $g^\mu_\gamma$ and $\hat{f}^\mu$ is defined by

$$g^\mu_\gamma = \frac{u^\mu_\gamma}{N_\mu}, \quad \hat{f}^\mu = \frac{\tilde{\varphi}^\mu}{N_\mu}, \quad N^2_\mu = \int \rho u^\mu_\gamma u^\mu_\gamma\, dV, \tag{2.7}$$

where $u^\mu_\gamma$ and $\tilde{\varphi}^\mu$ are the mechanical displacement and electric potential, respectively, which satisfy the equations of linear piezoelectricity

CH2588-2/88/0000-230 $1.00 C 1988 IEEE

$$\hat{K}^{\ell}_{L\gamma} = c_{L\gamma M\alpha} u_{\alpha,M} + e_{ML\gamma} \tilde{\varphi}_{,M},$$

$$\hat{\mathcal{D}}^{\ell}_{L} = e_{LM\gamma} u_{\gamma,M} - \epsilon_{LM} \tilde{\varphi}_{,M}, \qquad (2.8)$$

$$\hat{K}^{\ell}_{L\gamma,L} = \rho \ddot{u}_{\gamma}, \quad \hat{\mathcal{D}}^{\ell}_{L,L} = 0, \qquad (2.9)$$

subject to the appropriate boundary conditions, and $\rho$ is the mass density. Equations (2.8) are the linear piezoelectric constitutive relations and (2.9) are the stress equations of motion and charge equation of electrostatics, respectively. The upper cycle notation for many dynamic variables and the capital Latin and lower case Greek index notation is being employed for consistency with Ref.7, as is the remainder of the notation in this section.

The substitution of (2.3) in (2.2) yields

$$H_{\mu} = - \int_{V} \hat{c}_{L\gamma M\alpha} g^{\mu}_{\alpha,M} g^{\mu}_{\gamma,L} \, dV. \qquad (2.10)$$

Since $g^{\mu}_{\alpha}$ denotes the normalized surface wave mode shape and from (2.4) $\hat{c}_{L\gamma M\alpha}$ depends on the biasing state, $H_{\mu}$ can be evaluated when the surface wave mode shape and biasing state are known.

## 3. Decomposition into Extensional and Flexural Biases

A plan view and cross-section of the substrate extending beyond the supports is shown in Fig.1 along with the coordinate system. Since the active surface wave region is in the interior of the plate away from the corner regions, certain simplifying approximations may be made in the treatment while still retaining a high degree of accuracy in the interior region of the plate. First the corner regions are ignored and we obtain the first plan view in Fig.2, in which we have introduced the notation C for the central region, S for the two side regions and T for the top and bottom regions. Since the plate is thin, the equations of anisotropic plane stress apply in the extensional portion of the problem, for which we have taken the resisting forces due to the frits to lie in the center-plane of the plate rather than below the plate as shown in the cross-section in Fig.1. As a consequence of this the plan views shown in Fig.2 refer to the extensional portion of the problem only and we must add a flexural portion to the problem resulting from the actual location of the resisting forces in the frit below the plate rather than at the center-plane of the plate. For anisotropic plane stress all $K_{2L}$ vanish and only $K_{AB}$ exist, where we have introduced the convention that A, B, C, D take the values 1 and 3 and skip 2.

Now we note that we are interested in determining the biasing deformation only in region C, and then only in the central portion where the surface wave is located. Consequently, we are interested only in the influence of the S and T regions on the deformation in the C region, which affect the flexural portion of the deformation in the C region by virtue of the offset between the center-plane of the plate and the location of the resisting force in the frit. In the plan views of Fig.2 we have adopted the convention that solid lines refer to traction free edges along which we have the conditions

$$N_A K_{AB} = 0, \qquad ((3.1)$$

and dotted lines refer to the frits along which we have support. As a consequence of the geometry of the supporting frit, we now take the displacement $w_s$ in the tangential direction of the supporting frit to vanish. In view of this unquestionably accurate assumption, which for the tangential condition serves to separate the C, S and T regions, we make the somewhat inconsistent but, we believe, quite accurate simplifying assumption that the displacement $w_n$ in the direction normal to the frit need not be continuous across the frit. This assumption separates the extensional problem shown in the first plan view in Fig.2 into the three separate extensional problems shown in the second plan view in Fig.2.

On account of the stubbiness of the cross-section of a frit, we model its resistance to the normal displacement $w_n$ of the center-plane of the plate as a simple variable shear strip, which yields

$$K_{(n)(n)} = \mp k w_n, \qquad (3.2)$$

where[9]

$$k = 2\mu t_f / 3h^2, \qquad (3.3)$$

in which $\mu$ is the appropriate shear modulus, i.e., either $c_{44}$ or $c_{66}$ for the ST-cut depending on the direction in which the frit runs, and we have ignored the deformation in the frit to this approximation because $h_f \ll h$. Furthermore, in accordance with the discussion in the previous paragraph concerning the tangential displacement along the frit, we have

$$w_s = 0 \qquad (3.4)$$

at the frit.

As already noted, in view of the location of the resisting forces in the frit below the center-plane of the plate, for the C-region we have the first cross-section shown in Fig.2, which results in the extensional problem shown at the right that has already been discussed and the associated flexural problem is shown below. In the flexural problem we make the very conservative assumption of simple-supports, which maximizes the very small flexural deformation in the C-region but is inconsistent with the quite accurate shear-strip model for the evaluation of k because this model implies that the resulting applied moment is transmitted through the frit directly to the base, which is almost true. However, since the influence of the overestimated flexural deformation is very small, this inconsistent assumption of simple-supports is appropriate in this case. The couple associated with the flexural problem is given by

$$M_c = F_c h, \qquad (3.5)$$

where we have neglected $h_f$ since its inclusion is clearly unwarranted in this overestimate of $M_c$. For the influence of the S-regions on the overestimated flexure of the C-region we refer to the bottom cross-sections in Fig.2. From the resulting flexure of the C-region shown at the bottom-right of Fig.2, we obtain

$$M_s = F_s h, \qquad (3.6)$$

which with (3.5) yields the total applied moment to the C-region $M^{(1)}$, where

$$M^{(1)} = M_C^{(1)} + M_S^{(1)} = \mp 2h^2 (K_{11}^{C} + K_{11}^{S}) \text{ at } X_1 = \pm a. \qquad (3.7)$$

Clearly, by similar reasoning in a similar way including the influence of the T-region on the C-region, we obtain $M^{(3)}$, where

$$M^{(3)} = M_C^{(3)} + M_S^{(3)} = \mp 2h^2(K_{33}^C + K_{33}^T) \text{ at } X_3 = \pm b . \quad (3.8)$$

## 4. Extensional Biasing Deformation

As already noted, the equations of anisotropic plane stress apply, which may be written in the form

$$K_{AB,A} - \rho a_B = 0 , \quad (4.1)$$

where we have introduced the convention that A, B, C, D take the values 1 and 3 and skip 2 and $a_B$ denotes the acceleration in the plane of the plate and the anisotropic constitutive equations for plane stress may be written in the form

$$K_{AB} = \gamma_{ABCD}E_{CD}, \quad E_{CD} = \frac{1}{2}(w_{C,D} + w_{D,C}) , \quad (4.2)$$

where $\gamma_{ABCD}$ are Voigt's anisotropic plate elastic constants which are given by

$$\gamma_{RS} = c_{RS} - c_{RW}c_{WV}^{-1}c_{VS} ; \; R,S = 1,3,5 ; \; W,V = 2,4,6 . \quad (4.3)$$

From the relaxation of the stresses $K_{2L}$, we have

$$E_W = - c_{WV}^{-1}c_{VS}E_S , \quad (4.4)$$

which enables us to obtain all the strains $E_{KL}$ when a solution is known as well as the rotation $\Omega_{13}$ from the relation

$$\Omega_{13} = \frac{1}{2}(w_{3,1} - w_{1,3}) . \quad (4.5)$$

However, the rotations $\Omega_{2A}$ cannot be obtained from a solution because of the relaxation of the $K_{2L}$. Nevertheless, the flexural type plate rotations $\Omega_{2A}$ are taken to vanish in the case of uncoupled anisotropic plane stress, which is the kinematic uncoupling assumption for anisotropic plane stress. Accordingly, we have

$$\Omega_{2A} = 0 , \quad (4.6)$$

and now with (4.2) - (4.6) and

$$w_{K,L} = E_{KL} + \Omega_{LK} , \quad (4.7)$$

we have the required three-dimensional displacement gradients for the extensional portion of the deformation.

Plan views of the three regions of the plate denoted C, S and T are shown on the right-hand top of Fig. 2 and we note that the respective coordinate systems are placed in the center of each region in the treatment. In accordance with Section 3 the edge conditions for the C-region take the form

$$w_3 = 0, \; K_{11} = \mp kw_1 \text{ at } X_1 = \pm a ,$$
$$w_1 = 0, \; K_{33} = \mp kw_3 \text{ at } X_3 = \pm b , \quad (4.8)$$

while for the S region they take the form

$$w_3 = 0, \; K_{11} = -kw_1 \text{ at } \bar{X}_1 = +a_s ,$$
$$K_{11} = 0, \; K_{13} = 0 \text{ at } \bar{X}_1 = -a_s ,$$
$$w_1 = 0, \; K_{33} = \mp kw_3 \text{ at } X_3 = \pm b , \quad (4.9)$$

and for the T region they take the form

$$w_3 = 0, \; K_{11} = \mp kw_1 \text{ at } X_1 = \pm a ,$$
$$K_{31} = 0, \; K_{33} = 0 \text{ at } \bar{X}_3 = +b_T ,$$
$$w_1 = 0, \; K_{33} = kw_3 \text{ at } \bar{X}_3 = -b_T , \quad (4.10)$$

where the bars are used because of the placement of the coordinate system in the center of each region.

In all three problems we transform the inhomogeneities from the differential equations, which consist of (4.1) with (4.2), into the boundary conditions, which are given in either (4.8), (4.9) or (4.10) by writing

$$w_A = f_A + \hat{w}_A , \quad (4.11)$$

where for the C and S regions we have

$$f_1^{C,S} = A_1(X_3^2 - b^2), \; f_3^{C,S} = A_3(X_3 - b)^2 , \quad (4.12)$$

and for the T region we have

$$f_1^T = A_1(\bar{X}_3^2 - 2b_T\bar{X}_3 - 3b_T^2), \; f_3^T = A_3(\bar{X}_3 + b_T)^2 . \quad (4.13)$$

Since for ST-cut quartz $\gamma_{15} = \gamma_{35} = 0$, if for each region we select

$$A_1 = \rho a_1/2\gamma_{55}, \; A_3 = \rho a_3/2\gamma_{33} , \quad (4.14)$$

we obtain the homogeneous equations

$$\gamma_{11}\hat{w}_{1,11} + (\gamma_{13} + \gamma_{55})\hat{w}_{3,13} + \gamma_{55}\hat{w}_{1,33} = 0 ,$$
$$\gamma_{55}\hat{w}_{3,11} + (\gamma_{13} + \gamma_{55})\hat{w}_{1,13} + \gamma_{33}\hat{w}_{3,33} = 0 , \quad (4.15)$$

in each of the three regions. The further substitution of (4.11) with (4.12) into (4.8) and (4.9) yields the edge conditions

$$\hat{w}_1 = 0 \text{ at } X_3 = \pm b , \; \gamma_{13}\hat{w}_{1,1} + \gamma_{33}\hat{w}_3 + k\hat{w}_3 = 0 \text{ at } X_3 = +b ,$$
$$\gamma_{13}\hat{w}_{1,1} + \gamma_{33}\hat{w}_{3,3} - k\hat{w}_3 = 4A_3b(\gamma_{33} + kb) \text{ at } X_3 = -b ,$$
$$\hat{w}_3 = -A_3(X_3 - b)^2, \; \gamma_{11}\hat{w}_{11} + \gamma_{13}\hat{w}_{3,3} \pm kw_1 = \mp A_1(X_3^2 - b^2)$$
$$- 2A_3\gamma_{13}(X_3 - b) \text{ at } X_1 = \pm a , \quad (4.16)$$

for the C region and

$$\hat{w}_1 = 0 \text{ at } X_3 = \pm b , \; \gamma_{13}\hat{w}_{1,1} + \gamma_{33}\hat{w}_{3,3} + k\hat{w}_3 = 0 \text{ at } X_3 + b ,$$
$$\gamma_{13}\hat{w}_{1,1} + \gamma_{33}\hat{w}_{3,3} - k\hat{w}_3 = 4A_3b(\gamma_{33} + kb) \text{ at } X_3 = -b ,$$
$$\hat{w}_3 = -A_3(X_3 - b)^2 \text{ at } \bar{X}_1 = +a_s ,$$
$$\hat{w}_{1,3} + \hat{w}_{3,1} = -2A_1X_3 \text{ at } \bar{X}_1 = -a_s ,$$
$$\gamma_{11}\hat{w}_{1,1} + \gamma_{13}\hat{w}_{3,3} + k_1\hat{w}_1 = -kA_1(X_3^2 - b^2) - 2A_3\gamma_{13}(X_3 - b)$$
$$\text{at } \bar{X}_1 = +a_s ,$$
$$\gamma_{11}\hat{w}_{1,1} + \gamma_{13}\hat{w}_{3,3} = -2A_3\gamma_{13}(X_3 - b) \text{ at } \bar{X}_1 = -a_s , \quad (4.17)$$

for the S region. The substitution of (4.11) with (4.13) into (4.10) yields the edge conditions for the T region, which are of a similar nature to those for the S region and for the sake of brevity we do not bother to write here[9]. We now note that in each case there are exactly eight edge conditions, three of which are homogeneous in the transformed variables.

As a solution of (4.15) for each of the three regions consider

$$\hat{w}_A = \sum_m \beta_{AM} e^{i(\bar{\nu}_m \bar{X}_1 + \bar{\eta}_m \bar{X}_3)}, \qquad (4.18)$$

where

$$\bar{\nu}_m = \frac{m\pi}{2\sigma a}, \ \bar{X}_1 = X_1, \ \bar{X}_3 = X_3 \text{ in } C \text{ region},$$

$$\bar{\nu}_m = \frac{m\pi}{2\sigma a_s}, \ \bar{X}_1 = X_1 + a + a_s, \ \bar{X}_3 = X_3 \text{ in S region},$$

$$\bar{\nu}_m = \frac{m\pi}{2\sigma a}, \ \bar{X}_1 = X_1, \ \bar{X}_3 = X_3 - (b + b_T) \text{ in T region}, \qquad (4.19)$$

and m takes all + and - integers considered in the sum, and $\sigma$ is a parameter which must be an irrational number so that neither sin nor cos vanish at the appropriate edges $\bar{X}_1 = \pm a$ or $\bar{X}_1 = \pm a_s$ and can be chosen to improve convergence. It has been found that $\sigma = \pi$ works well. The substitution of (4.18) into (4.15) yields

$$(\gamma_{11}\bar{\nu}_m^2 + \gamma_{55}\bar{\eta}_m^2)\beta_{1m} + (\gamma_{13} + \gamma_{55})\bar{\nu}_m\bar{\eta}_m\beta_{3m} = 0,$$

$$(\gamma_{13} + \gamma_{55})\bar{\nu}_m\bar{\eta}_m\beta_{1m} + (\gamma_{55}\bar{\nu}_m^2 + \gamma_{33}\bar{\eta}_m^2)\beta_{3m} = 0, \qquad (4.20)$$

for each m. Equations (4.20) constitute a system of two linear homogeneous algebraic equations in $\beta_{1m}$ and $\beta_{3m}$, which yields nontrivial solutions when the determinant of the coefficients vanishes. If we define $\bar{\eta}_m = h\bar{\nu}_m$, then the vanishing of the determinant yields

$$\gamma_{33}\gamma_{55}h^4 + (\gamma_{11}\gamma_{33} - 2\gamma_{13}\gamma_{55})h^2 + \gamma_{11}\gamma_{55} = 0, \qquad (4.21)$$

which has four roots $h^{(p)}$ $(p = 1,2,3,4)$ and corresponding sets of amplitude ratios $\beta_1^{(p)} : \beta_3^{(p)}$, which are independent of m. This holds for all regions. Hence, as a solution of the boundary value problem in each region we take

$$\hat{w}_A = \sum_m \sum_{p=1}^{4} \bar{D}_p^m \beta_A^{(p)} e^{i\bar{\nu}_m \bar{X}_1} e^{i\bar{\eta}_m^{(p)} \bar{X}_3}, \qquad (4.22)$$

where the $\bar{D}_p^m$ are amplitude coefficients still to be determined and $\bar{\eta}_m^{(p)} = h^{(p)}\bar{\nu}_m$. By requiring the homogeneous equations in either (4.16), (4.17) or the equivalent equations for the T region, which are not shown, to be satisfied for each m, we can reduce the number of unknown coefficients from 4 to 1 for each m. Accordingly, we obtain

$$\bar{D}_p^m = \bar{g}_p^m \bar{D}_1^m, \qquad (4.23)$$

where $\bar{g}_1^m = 1$ and the other $\bar{g}_p^m$ are lengthy expressions, which are too cumbersome to present here. The substitution of (4.23) into (4.22) yields the reduced form

$$\hat{w}_A = \sum_m \bar{D}_1^m e^{i\bar{\nu}_m \bar{X}_1} \sum_{p=1}^{4} \bar{g}_p^m \beta_A^{(p)} e^{i\bar{\eta}_m^{(p)} \bar{X}_3}, \qquad (4.24)$$

for the solution function.

The solution functions in (4.24) satisfy the differential equations and the homogeneous boundary conditions in either (4.16), (4.17) or the equivalent equations for the T region, but they do not satisfy the inhomogeneous boundary conditions in the transformed variables in either (4.16), (4.17) or those for the T region. We satisfy these latter conditions approximately by means of a variational approximation procedure. Since the solution functions for each region satisfy the differential equations exactly, all that remains of the appropriate variational Principle in which all conditions are unconstrained[10] is

$$-\int_{S_N} N_A K_{AB} \delta w_B \, dS + \int_{S_C} w_{(s)} \delta K_{(n)(s)} \, dS$$

$$-\int_{S_N} (k w_{(n)} \pm K_{(n)(n)}) \delta w_{(n)} \, dS = 0, \qquad (4.25)$$

for each region C, S and T. In (4.25) $N_A$ denotes the outwardly directed unit normal to a free edge, the integral along $S_C$ is for no tangential displacement along the frit and the last integral is for the spring resistance to the normal displacement of the frit. Substituting from (4.11) and (4.12) into (4.25) and using (4.16), we obtain

$$\int_{-a}^{a} \left[\hat{w}_1 \delta\hat{K}_{13}\right]_{-b}^{b} dX_1 + \int_{-b}^{b} \left[(\hat{w}_3 + A_3(X_3 - b)^2)\delta\hat{K}_{13}\right]_{-a}^{a} dX_3$$

$$+ \int_{-a}^{a} \left[(-k\hat{w}_3 - \hat{K}_{33})\delta\hat{w}_3\right]_{b} dX_1 + \int_{-a}^{a} \left[(-k\hat{w}_3 - k4b^2A_3 + \hat{K}_{33} - 4bA_3\gamma_{33})\delta\hat{w}_3\right]_{-b} dX_1$$

$$+ \int_{-b}^{b} \left[(-k\hat{w}_1 - kA_1(X_3^2 - b^2) - \hat{K}_{11} - 2A_3\gamma_{13}(X_3 - b))\delta\hat{w}_1\right]_{a} dX_3$$

$$+ \int_{-b}^{b} \left[(-k\hat{w}_1 - kA_1(X_3^2 - b^2) + \hat{K}_{11} + 2A_3\gamma_{13}(X_3 - b))\delta\hat{w}_1\right]_{-a} dX_3 = 0, \qquad (4.26)$$

for the C region. In a similar way, but using (4.17) instead of (4.16), we obtain

$$\int_{-a_s}^{a_s} \left[\hat{w}_1 \delta\hat{K}_{13}\right]_{-b}^{b} dX_1 + \int_{-b}^{b} \left[(\hat{w}_3 + A_3(X_3 - b)^2)\delta\hat{K}_{13}\right]_{a_s}^{a_s} dX_3$$

$$- \int_{-a_s}^{a_s} \left[(k\hat{w}_3 + \hat{K}_{33})\delta\hat{w}_3\right]_{b} dX_1$$

$$+\int_{-a_s}^{a_s}\left[(-k\hat{w}_3 - 4kb^2A_3 + \hat{K}_{33} - 4bA_3\gamma_{33})\delta\hat{w}_3\right]_{-b}\,dX_1$$

$$+\int_{-b}^{b}\Big[(-k\hat{w}_1 + kA_1(X_3^2 - b^2) - \hat{K}_{11} - 2A_3\gamma_{13}(X_3 - b))\delta\hat{w}_1\Big]_{a_s}\,dX_3 + \int_{-b}^{b}\Big[[\hat{K}_{11} + 2A_3\gamma_{13}(X_3 - b)]\delta\hat{w}_1 + (\hat{K}_{13} + 2A_1\gamma_{55}X_3)\delta\hat{w}_3\Big]_{-a_s}\,dX_3 = 0\,, \qquad (4.27)$$

for the S region. The equivalent equation for the T region is somewhat similar and we do not bother to write it here to save space[9]. Substituting from (4.24) into (4.26) for the C region, (4.27) for the S region and the equivalent, which has not been written, for the T region, we obtain an equation of the form

$$\sum_{n=1}^{N}\sum_{m=1}^{N}\bar{A}_{mn}\bar{D}_1^m\delta\bar{D}_1^n + \sum_{n=1}^{N}\bar{b}_n\delta\bar{D}_1^n = 0\,, \qquad (4.28)$$

for each region, where the expressions for $A_{mn}$ and $b_n$ are too lengthy to present here. Since the variations $\delta D_1^n$ are arbitrary for each region, we have

$$\sum_{m=1}^{N}\bar{A}_{mn}\bar{D}_1^m = -\bar{b}_n\,, \qquad (4.29)$$

which constitute N inhomogeneous linear algebraic equations in the N unknowns $D_1^m$ for each region. Convergence is determined by increasing the number N and comparing the solution for N with that for N - 1.

## 5. Flexural Biasing Deformation

With the aid of Mindlin's plate equations[11-13] it has been shown that the equation of anisotropic static flexure with $X_2$ normal to the major surfaces can be written in the form[14]

$$K_{AB,AB}^{(1)} + F_{B,B}^{(1)} + F_2^{(0)} = 0\,, \qquad (5.1)$$

where

$$K_{AB}^{(1)} = \int_{-h}^{h} X_2K_{AB}\,dX_2\,,\quad F_B^{(1)} = h[K_{2B}(h) + K_{2B}(-h)]\,, \qquad (5.2)$$

and in this instance

$$F_B^{(1)} = 0\,,\quad F_2^{(0)} = K_{22}(h) - K_{22}(-h) = 0\,, \qquad (5.3)$$

since $K_{2K}$ vanishes on the major surfaces. From Eqs.(4.25) of Ref.14 the constitutive equations for the stress-resultants take the form

$$K_{AB}^{(1)} = \frac{2}{3}h^3\gamma_{ABCD}E_{CD}^{(1)}\,. \qquad (5.4)$$

The plate strains $E_{CD}^{(1)}$ in (5.4) are given by

$$E_{CD}^{(1)} = \frac{1}{2}\left(w_{C,D}^{(1)} + w_{D,C}^{(1)}\right)\,, \qquad (5.5)$$

and from the relaxation of the stress resultants $K_{2L}^{(1)}$, the vanishing of the plate shear strains $E_{2A}^{(0)}$ and (5.5) we have the respective relations

$$E_W^{(1)} = -c_{WV}^{-1}c_{VS}E_S^{(1)}\,,\quad w_A^{(1)} = -w_{2,A}^{(0)}\,,\quad E_{CD}^{(1)} = -w_{2,CD}^{(0)}\,, \qquad (5.6)$$

which, respectively, are given in Eqs.(4.24), (3.34) and (3.37) of Ref.14. Since from Eqs.(3.31) of Ref.15 we have $E_{KL}^{(0)} = 0$, when the plate deflection $w_2^{(0)}$ has been found, we know the three-dimensional strain field from the relation

$$E_{KL} = \frac{1}{2}(w_{L,K} + w_{K,L}) = X_2E_{KL}^{(1)}\,, \qquad (5.7)$$

as in Eqs.(3.32) of Ref.15.

Although we now have the plate strains from (5.7), we cannot yet determine the $\hat{c}_{L\gamma M\alpha}$ from (2.4) because, as noted in Ref.15, we need the displacement gradients $w_{L,K}$ or equivalently the rotations $\Omega_{KL}$, which are defined by

$$\Omega_{KL} = \frac{1}{2}(w_{L,K} - w_{K,L})\,. \qquad (5.8)$$

As in Ref.15, we require that the associated plate rotations that accompany the plate strains that arise from the relaxation of the plate stress resultants $K_{2L}^{(1)}$ satisfy the appropriate three-dimensional rotation gradient-strain gradient relations, i.e.,

$$\Omega_{KL,M} = E_{ML,K} - E_{MK,L}\,. \qquad (5.9)$$

From (3.22) and (3.23) of Ref.15 we have the expressions for the plate rotations

$$\Omega_{AB}^{(n)} = \frac{1}{2}\left(w_{B,A}^{(n)} - w_{A,B}^{(n)}\right)\,,$$

$$\Omega_{2A}^{(n)} = \frac{1}{2}\left[(n+1)w_A^{(n+1)} - w_{2,A}^{(n)}\right]\,. \qquad (5.10)$$

From (3.33) of Ref.15, we have

$$\Omega_{13}^{(0)} = \text{arbitrary constant} = 0,\ \Omega_{13}^{(1)} = 0\,, \qquad (5.11)$$

and from (3.35) and (3.34), respectively, of Ref.15, we have

$$\Omega_{2A}^{(0)} = -w_{2,A}^{(0)}\,,\quad \Omega_{2A}^{(1)} = E_{2A}^{(1)}\,, \qquad (5.12)$$

which with (3.24) of Ref.15 enables us to write

$$\Omega_{2A} = -w_{2,A}^{(0)} + X_2E_{2A}^{(1)}\,,\quad \Omega_{13} = 0\,, \qquad (5.13)$$

which is the same as (3.36) of Ref.15. Now, from (5.7), (5.13) and

$$w_{K,L} = E_{KL} + \Omega_{LK}\,, \qquad (5.14)$$

we have the desired three-dimensional displacement gradients for the flexural portion of the deformation when the plate deflection $w_2^{(0)}$ has been found.

For the case of flexure of the C region due to extension of the C, S and T regions, the differential equation (5.1) with (5.3) and constitutive equations (5.4) with $(5.6)_3$ take the form

$$K^{(1)}_{AB,AB} = 0,\ K^{(1)}_{AB} = -\frac{2}{3}h^3\gamma_{ABCD}w^{(0)}_{2,CD}, \quad (5.15)$$

and since the plate is simply-supported along the edges, from flexure due to extension of the C and T regions the boundary conditions take the form

$$w^{(0)}_2 = 0,\ K^{(1)}_{11} = 0 \ \text{ at } \ X_1 = \pm a,$$
$$w^{(0)}_2 = 0,\ K^{(1)}_{33} = \mp\bar{m}_3 = -2h^2(K^C_{33}+K^T_{33}),\ \text{at } X_3 = \pm b, \quad (5.16)$$

from $M^{(3)}$ given in (3.8). Since $\gamma_{15} = \gamma_{35} = 0$, we can obtain a series solution in the form

$$w^{(0)}_2 = \sum_{m=1,3,5}^{\infty} f_m(X_3)\cos\frac{m\pi X_1}{2a}, \quad (5.17)$$

the substitution of which in (5.15) yields

$$\gamma_{11}\left(\frac{m\pi}{2a}\right)^4 f_m - 2(\gamma_{13}+2\gamma_{55})\left(\frac{m\pi}{2a}\right)^2 f''_m + \gamma_{33}f''''_m = 0, \quad (5.18)$$

for each m. Since the problem is antisymmetric in $X_3$ we take

$$f_m = \sinh\zeta_m X_3, \quad (5.19)$$

the substitution of which in (5.18) yields

$$\zeta^{(1)}_{m2} = \frac{m\pi}{2a}\left[\frac{\gamma_{13}+2\gamma_{55}}{\gamma_{33}} \pm \sqrt{\frac{(\gamma_{13}+2\gamma_{55})^2 - \gamma_{11}\gamma_{33}}{\gamma^2_{33}}}\right]^{1/2}, \quad (5.20)$$

along with the negatives, which do not yield independent antisymmetric solutions. Thus, for each m, we take a sum of the two solutions in the form

$$f_m = A_m \sinh\zeta^{(1)}_m X_3 + B_m \sinh\zeta^{(2)}_m X_3. \quad (5.21)$$

The satisfaction of the edge conditions $(5.16)_3$ for each m yields

$$B_m = S_m A_m, \quad (5.22)$$

where

$$S_m = -\sinh\zeta^{(1)}_m b/\sinh\zeta^{(2)}_m b. \quad (5.23)$$

In order to satisfy the remaining edge conditions $(5.16)_4$ we take the form

$$w^{(0)}_2 = \sum_{m=1,3,5}^{\infty} A_m(\sinh\zeta^{(1)}_m X_3 + S_m \sinh\zeta^{(2)}_m X_3)\cos\frac{m\pi X_1}{2a}, \quad (5.24)$$

and from the orthogonality of the $\cos m\pi X_1/2a$ in the interval $-a < X_1 < a$, we obtain

$$A_m = M_{3m}\left[-\frac{2}{3}h^3\gamma_{33}(\zeta^{(1)2}_m - \zeta^{(2)2}_m)\sinh\zeta^{(1)}_m b\right]^{-1}, \quad (5.25)$$

where

$$M_{3m} = \frac{1}{a}\int_{-a}^{a}\bar{m}_3\cos\frac{m\pi X_1}{2a}dX_1. \quad (5.26)$$

For the case of the edge moments due to extension of the C and S region the solution is obtained in the same way, but with the treatment of the $X_1$- and $X_3$-coordinates interchanged. Consequently, we simply write the resulting solution by analogy with the previous one, thus

$$w^{(0)}_2 = \sum_{m=1,3,5}^{\infty}\bar{A}_m(\sin h\bar{\zeta}^{(1)}_m X_1 + \bar{S}_m \sin h\bar{\zeta}^{(2)}_m X_1)\cos\frac{m\pi X_3}{2b}, \quad (5.27)$$

where

$$\bar{\zeta}^{(1)}_{m2} = \frac{m\pi}{2b}\left[\frac{\gamma_{13}+2\gamma_{55}}{\gamma_{11}} \pm \sqrt{\frac{(\gamma_{13}+2\gamma_{55})^2 - \gamma_{11}\gamma_{33}}{\gamma^2_{11}}}\right], \quad (5.28)$$

$$\bar{A}_m = M_{1m}\left[-\frac{2}{3}h^3\gamma_{11}\left(\bar{\zeta}^{(1)2}_m - \bar{\zeta}^{(2)2}_m\right)^2\sin h\bar{\zeta}^{(1)}_m a\right]^{-1},$$

$$M_{1m} = \frac{1}{b}\int_{-b}^{b}\bar{m}_1\cos\frac{m\pi X_3}{2b}dX_3,$$

$$\bar{S}_m = -\sin h\bar{\zeta}^{(1)}_m a/\sin h\bar{\zeta}^{(2)}_m a,\ \bar{m}_1 = -2h^2(K^C_{11}+K^S_{11}) \ \text{ at } X_1 = \pm a. \quad (5.29)$$

Now, from the combined $w^{(0)}_2$, which consists of the sum of (5.24) and (5.27), we can obtain the biasing $w_{K,L}$ due to flexure from Eqs. $(5.6)_3$, $(5.6)_1$, (5.7), (5.13) and (5.14). The total biasing $w_{K,L}$ is obtained by adding the one due to flexure to the one due to extension, which may be obtained from (4.11), (4.12), (4.14) and (4.24), all for the C region.

## 6. Resonant Surface Wave Mode Shape

In this section we present the proper continuous representation of the acoustic surface wave mode shape in resonators with grooved reflectors, which was obtained in recent work[8]. The straight-crested surface wave displacement field may be written in the known form[16,17],

$$u_j = \alpha_j(X_2)e^{i\xi(X_1-Vt)},\ \alpha_j = \sum_{m=1}^{4}C^{(m)}A^{(m)}_j e^{i\beta_m\xi X_2}. \quad (6.1)$$

A plan view of the resonator showing the reflecting arrays of grooves, the coordinate system and the associated planar geometry is shown in Fig.1, and we note that the plane $X_2 = 0$ denotes the ungrooved surface of the substrate and the axis points down. It has been shown[8] that the variable-crested resonant surface wave mode shape with variable amplitude along the transmission path is very accurately approximated by

$$u_j = \cos\frac{\pi X_3}{2w}\,\mathrm{Re}\left[\alpha_j(X_2)\hat{C}^R(X_1)e^{i\xi(X_1-s)} + \alpha^*_j(X_2)\hat{C}^L(X_1)e^{-i\xi(X_1-s)}\right]e^{-i\omega t}, \quad (6.2)$$

where the variations along the transmission path are given by

$$\hat{C}^R(X_1) = \left[ r_1 e^{-\alpha_1 N} e^{-\beta_2 X_1} - r_2 e^{-\alpha_2 N} e^{-\beta_1 X_1} \right] \bar{C}/d ,$$

$$\hat{C}^L(X_1) = r_1 r_2 \left[ e^{-\alpha_1 N} e^{-\beta_2 X_1} - e^{-\alpha_2 N} e^{-\beta_1 X_1} \right] \bar{C}/d . \quad (6.3)$$

In (6.3) $\bar{C}$ denotes the amplitude of the input wave, N denotes the number of grooves,

$$d = r_1 e^{-\alpha_1 N} - r_2 e^{-\alpha_2 N}, \quad \beta_{1,2} = \alpha_{1,2} N/(\ell - s) , \quad (6.4)$$

and $\alpha_{1,2}$ and $r_{1,2}$ are given in Eqs.(32) and (34) of Ref.8. The amplitude of the standing surface wave mode, i.e., the part multiplying cos $\omega t$, along the transmission path is plotted in Fig.3.

## 7. Acceleration Sensitivity

From Sec.6 we now know $g_\gamma^\mu$ and from Sec.4 we know $\hat{c}_{L\gamma M\alpha}$ for extension and from Sec.5 we know $\hat{c}_{L\gamma M\alpha}$ for flexure, both for in-plane acceleration for this configuration. Hence, we can now evaluate $H_\mu$ in Eq.(2.2). Such calculations have been performed using the known values of the second order[18] and third order[19] elastic constants of quartz. For acceleration in the propagation direction the calculated sensitivity due to extension is plotted in Fig.4 as a function of aspect ratio a/b for b = 5.0 mm. The calculated curve shows that the acceleration sensitivity decreases first rapidly and gradually more slowly with increasing aspect ratio, reaching a value of $5 \times 10^{-11}$ per g at a/b = 3 and becoming asymptotic to $3 \times 10^{-11}$ per g at a/b $\approx$ 6. This result is independent of the overhangs $2a_s$ and $2b_T$ because of the inconsistent assumption made concerning the lack of continuity of $w_n$ across the frits. We intend to eliminate this simplifying but unnecessary assumption in later work. The overestimated acceleration sensitivity due to the associated flexure for this acceleration direction is plotted in Fig.5 for overhangs $2a_s$ = 5 mm and $2b_T$ = 5 mm. The figure shows that the overestimated sensitivity due to flexure is at most .002 of the sensitivity resulting from extension and, hence, is negligible. For acceleration normal to the propagation direction the calculated in-plane sensitivity due to extension is plotted in Fig.6 as a function of aspect ratio a/b for b = 5.0 mm. The calculated curve shows that the sensitivity decreases very rapidly with aspect ratio at first, reaches a minimum of about $1 \times 10^{-11}$ per g at a/b $\approx$ 3 and then increases to an asymptotic value of about $1.7 \times 10^{-11}$ per g at an aspect ratio of about 6. The overestimated acceleration sensitivity due to the associated flexure is shown in Fig.7 for overhangs $2a_s$ = 5 mm and $2b_T$ = 5 mm. This figure shows that this overestimated sensitivity due to flexure is also negligible. The resultant in-plane acceleration sensitivity is plotted in Fig.8, which shows that for b = 5.0 mm it reaches a minimum of about $1.2 \times 10^{-11}$ per g at a/b $\approx$ 3. Since for a fixed aspect ratio a/b the in-plane acceleration sensitivity in both directions varies with b, we have shown this variation in the propagation direction and normal to it in Figs.9 and 10, respectively.


## Acknowledgement

This work was supported in part by the Army Research Office under Contract No. DAAG-29-85-K-0075.


## References


1. H.F. Tiersten and D.V. Shick, "An Analysis of the Normal Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators with the Substrate Extending Beyond the Supports," to be published in the Proceedings of the 2nd European Frequency and Time Forum (1988).

2. D.V. Shick and H.F. Tiersten, "An Analysis of the Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators Supported Along the Edges," Proceedings of the 40th Annual Symposium on Frequency Control, U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, New Jersey and Institute of Electrical and Electronics Engineers, New York, IEEE Cat. No. 86CH2330-9, 262 (1986).

3. H.F. Tiersten and D.V. Shick, "An Analysis of the Normal Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators Rigidly Supported Along the Edges," Proceedings of the 41st Annual Symposium on Frequency Control, U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, New Jersey and Institute of Electrical and Electronics Engineers, New York, IEEE Cat. No. 87CH2427-3, 282 (1987).

4. G.K. Montress, T.E. Parker and J. Callerame, "A Miniature Hybrid Circuit SAW Oscillator Using an All Quartz Packaged Resonator," 1985 Ultrasonics Symposium Proceedings, IEEE Cat. No. 85CH2209-5, Institute of Electrical and Electronics Engineers, New York, 277 (1985).

5. J.A. Greer, T.E. Parker, M. Rothschild and D.J. Ehrlich, "A Novel Technique for Trimming the Frequency of a Sealed Surface Acoustic Wave Resonator," Proceedings of the 41st Annual Symposium on Frequency Control, U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, New Jersey and Institute of Electrical and Electronics Engineers, New York, IEEE Cat. No. 87CH2427-3, 351 (1987).

6. J.A. Greer and T.E. Parker, "Improved Vibration Sensitivity of the All Quartz Package Surface Acoustic Wave Resonator," Proceedings of the 42nd Annual Symposium on Frequency Control (these proceedings).

7. H.F. Tiersten, "Perturbation Theory for Linear Electroelastic Equations for Small Fields Superposed on a Bias," J. Acoust. Soc. Am., 64, 832 (1978).

8. H.F. Tiersten, J.T. Song and D.V. Shick, "A Continuous Representation of the Acoustic Surface Wave Mode Shape in Arrays of Reflecting Grooves," J. Appl. Phys., 62, 1154 (1987).

9. For more detail see H.F. Tiersten and D.V. Shick, "An Analysis of the Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators with the Substrate Extending Beyond the Supports," to be issued as a technical report, Rensselaer Polytechnic Institute, Troy, New York.

10. H.F. Tiersten, Linear Piezoelectric Plate Vibrations (Plenum, New York, 1969), Chap.6, Sec.4, Eq.(6.44).

11. R.D. Mindlin, "An Introduction to the Mathematical Theory of the Vibration of Elastic Plates," U.S. Army Signal Corps Eng. Lab., Fort Monmouth, New Jersey (1955). Signal Corps Contract DA-36-03956-56772.

12. R.D. Mindlin, "High Frequency Vibrations of Crystal Plates," Quart. Appl. Math., 19, 51 (1961).

13. Ref.10, Chap.13.

14. H.F. Tiersten and B.K. Sinha, "Temperature Dependence of the Resonant Frequency of Electroded Doubly-Rotated Quartz Thickness-Mode Resonators," J. Appl. Phys., 50, 8038 (1979).

15. D.S. Stevens, H.F. Tiersten and B.K. Sinha, "Temperature Dependence of the Resonant Frequency of Electroded Contoured AT-Cut Quartz Crystal Resonators," J. Appl. Phys., 54, 1709 (1983).

16. J.J. Campbell and W.R. Jones, "A Method for Estimating Optimal Crystal Cuts and Propagation Directions for Excitation of Piezoelectric Surface Waves," IEEE Trans. Sonics Ultrason., SU-15, 209 (1968).

17. B.K. Sinha and H.F. Tiersten, "Elastic and Piezoelectric Surface Waves Guided by Thin Films," J. Appl. Phys., 44, 4830 (1973). Sec.II.

18. R. Bechmann, "Elastic and Piezoelectric Constants of Alpha-Quartz," Phys. Rev., 110, 1060 (1958).

19. R.N. Thurston, H.J. McSkimin and P. Andreatch, Jr., "Third Order Elastic Constants of Quartz," J. Appl. Phys., 37, 267 (1966).

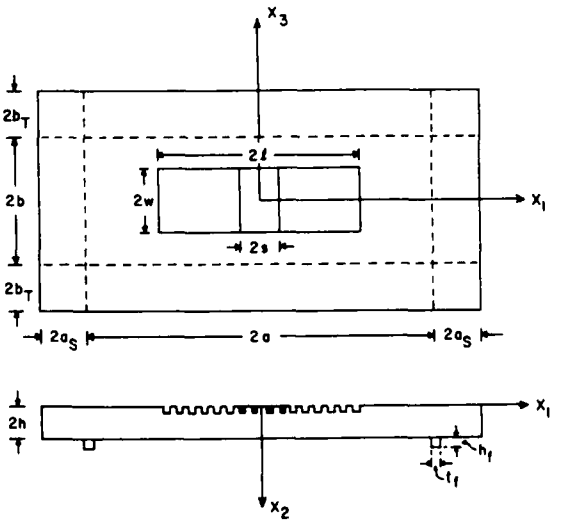



Figure 1 Plan View and Cross Section of Rectangular ST-Cut Quartz Plate with the Substrate Extending Beyond Rectangular Supports

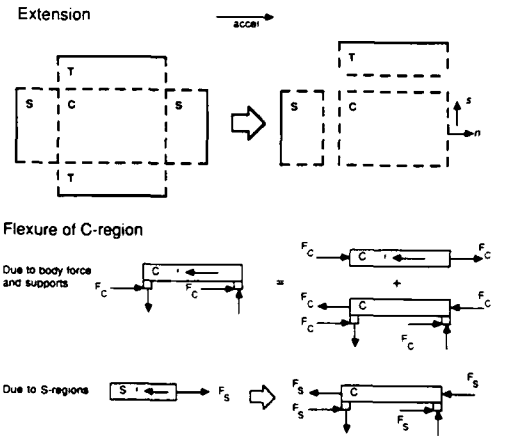



Figure 2 Plan Views and Cross-Sections of Plate for Simplifying Approximations Made in the Analysis for the Biasing State

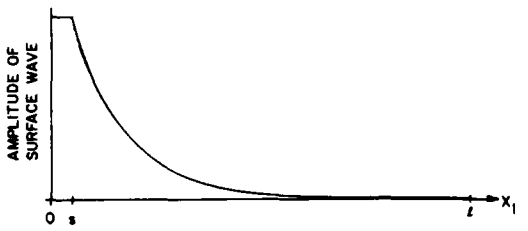



Figure 3 Amplitude of Resonant Surface Wave Mode Along Transmission Path

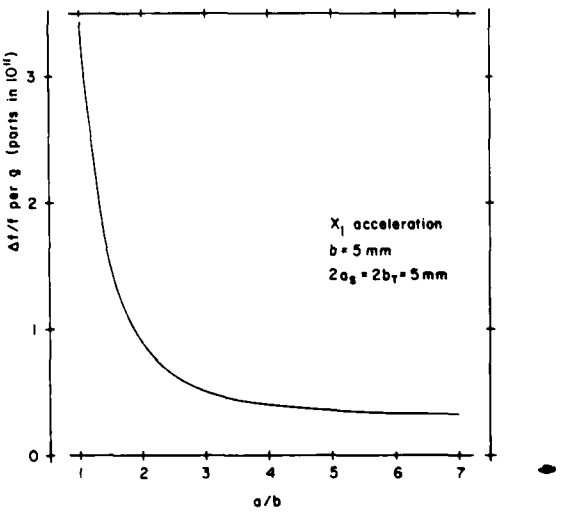



Figure 4 Calculated In-Plane Sensitivity Versus Planar Aspect Ratio a/b for Acceleration in the Propagation Direction

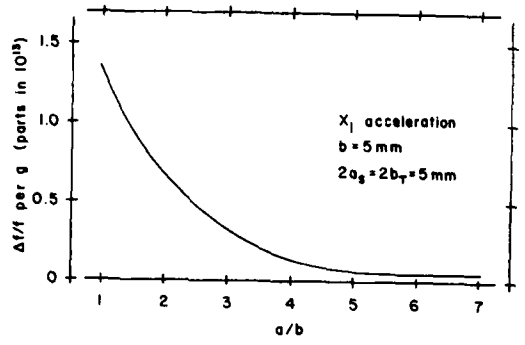



Figure 5 Calculated Sensitivity Due to Overestimated Flexural Deformation for Acceleration in the Propagation Direction for Addition to Curve in Figure 4

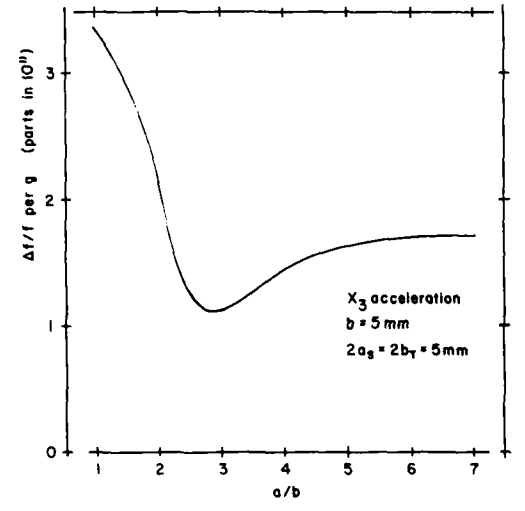



Figure 6 Calculated In-plane Sensitivity Versus Planar Aspect Ratio a/b for Acceleration Normal to the Propagation Direction

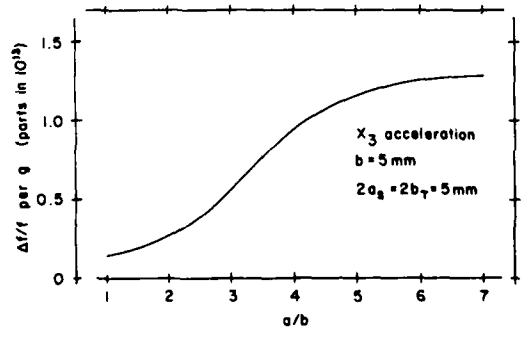



Figure 7 Calculated Sensitivity Due to Overestimated Flexural Deformation for Acceleration Normal to the Propagation Direction for Addition to Curve in Figure 6

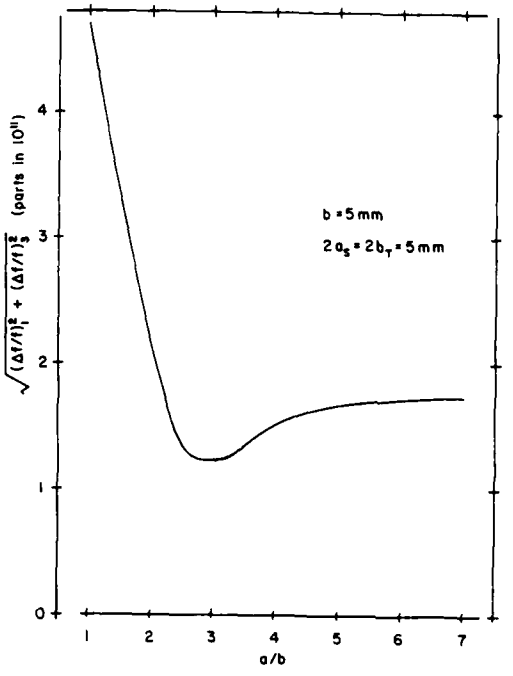



Figure 8 Resultant In-plane Acceleration Sensitivity Versus Planar Aspect Ratio a/b Obtained from Figures 4 and 6

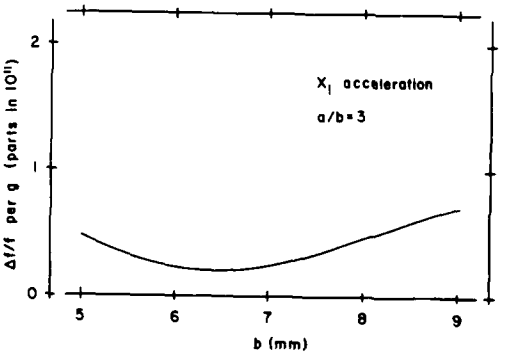



Figure 9 Calculated In-plane Sensitivity for Acceleration in the Propagation Direction Versus b for a Fixed Aspect Ratio a/b

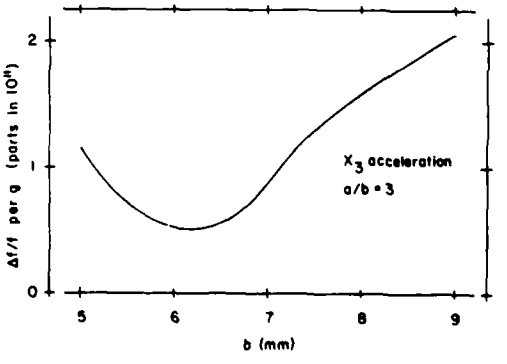



Figure 10 Calculated In-plane Sensitivity for Acceleration Normal to the Propagation Direction Versus b for a Fixed Aspect Ratio a/b

42nd Annual Frequency Control Symposium - 1988

# IMPROVED VIBRATION SENSITIVITY OF THE ALL QUARTZ PACKAGE SURFACE ACOUSTIC WAVE RESONATOR

J. A. Greer and T. E. Parker
Raytheon Research Division
131 Spring Street
Lexington, MA 02173

ABSTRACT: By exploiting the unique nature of the all quartz package (AQP) surface acoustic wave resonator (SAWR), we have significantly and reproducibly reduced the magnitude of the three dimensional vibration sensitivity, $|\Gamma_{3D}|$, of SAWR devices to below $3X10^{-10}/g$. First, by simply altering the geometry of the glass frit which forms the seal between the SAW substrate and matching quartz cover, the vibration sensitivity can be significantly reduced. Also, by fabricating two matched SAWR devices on opposite sides of the same substrate and connecting them in parallel the vibration sensitivity can also be reduced, even with non-ideal mounting support structures. These results are the lowest values reported for all 3 vibration axes published to date for SAWRs and represent a significant advance in the state of the art for this technology.

INTRODUCTION: As a result of the need to detect smaller targets at greater range, specifications for advanced radar systems are placing stringent requirements on the oscillator components in order to obtain a stable, spectrally pure signal. Such specifications include quiescent phase noise and vibration sensitivity. The specification for vibration sensitivity becomes more extreme for airborne applications such as missiles, re-entrant vehicles, helicopters, and high performance aircraft. This tighter specification is driven by the harsh vibration environment of these platforms since the quiescent phase-noise of the oscillator will be degraded by the vibration spectrum of the vehicle through the oscillator's vibration sensitivity. For this reason most airborne applications to date have utilized bulk acoustic wave (BAW) oscillators in the system since these devices have demonstrated vibration sensitivity levels down to a few parts in $10^{10}/g$ [1]. This vibration sensitivity is about one order of magnitude lower than the best published results for surface acoustic wave resonator (SAWR) devices [2], [3], [4]. However, SAWR oscillators have a significantly lower quiescent phase-noise spectrum than BAW oscillators for carrier offset frequencies greater than 1 kHz (when compared at the same frequency). Thus, except for slow moving targets, SAWR oscillators allow detection of smaller or more distant objects. A SAW oscillator multiplied to X-Band will weigh less and also require less power than a comparable BAW X-Band oscillator since it will not require as many stages of multiplication and amplification. Thus, a reduction in the vibration sensitivity of SAWR oscillators by about an order of magnitude (making the vibration sensitivity of SAW compatible t< BAW) would provide the system engineer with an attractive alternative to BAW oscillators.

SAWR oscillator vibration sensitivities have been dominated by the sensitivity of the SAW device itself. Typical experimental values [2],[3], [4] for the vibration sensitivity of SAWRs have been on the order of a few parts in $10^9/g$. These values have been obtained for both cold weld TO-8 and all quartz package (AQP) [5] type resonators and are more than two orders of magnitude higher than theoretical predictions [6],[7]. One proposed explanation for the discrepancy between theory and experiment has been the inability of the various packaging schemes to eliminate the bending stresses from the SAW device [3], [4]. The vibrationally-induced bending moments are the most significant for SAW devices since they produce stresses which are maximum at the substrate surface, exactly where the devices acoustic energy is concentrated. Given the difficulty in providing a SAWR substrate with a truly uniform support, one must look to other means of reducing the microphonic sensitivity of these devices. One recent theory has proposed taking advantage of the fact that frequency shifts from the two mutually perpendicular strain distributions in the plane of the device (parallel and perpendicular to the acoustic propogation axis) can be made to cancel for the AQP SAWR when the package is vibrated normal to the plane of the device [7], generally the most sensitive direction to vibration. Thus, if a SAWR could be supported in such a way that these strain distributions would be mutually offsetting, a lower vibration sensitivity would result. Another approach to cancel the affect of vibrationally induced bending stress is to fabricate two SAWR's on opposite sides of the same

CH2588-2/88/0000-239 $1.00 C 1988 IEEE

substrate. When the substrate is subjected to forces which produce bending moments, one substrate surface will be subjected to a tensile strain distribution and the second surface will be subjected to an equal but opposite compressive strain distribution. When properly connected together, either in series or parallel (forming a dual resonator), the overall effect of vibration normal to the plane of the device on the frequency of the dual resonator will be significantly reduced due to cancellation of the two equal but opposite strain distributions.

We have taken both of these approaches to reduce the microphonic sensitivity of the AQP SAWR. The AQP is an attractive packaging technique for SAWRs as it provides a number of advantages over the more conventional cold-weld TO-8 enclosure. These advantages included a low profile (making the device attractive for hybrid oscillators) and its transparent nature. This latter advantage allows accurate frequency adjustment of the SAWR after hermetic sealing using a laser-based frequency trimming technique [8]. The unique nature of the AQP SAWR device also allows us to readily implement the two vibration reduction approaches discussed above to cancel the effect of bending stresses on the resonant frequency of the device. First, the theory of Tiersten and Shick [7] indicates that there is an optimum geometry for the glass frit of the AQP which will yield a low vibration sensitivity normal to the plane of the device. Second, the nature of the AQP has allowed us to fabricate a novel dual resonator design with two SAWR devices hermeticaly sealed on opposite sides of the same substrate. The experimental results obtained from devices fabricated utilizing either of these approaches have yielded significant reductions in the vibration sensitivity of SAWRs. SAWRs have now demonstrated equivalent vibration sensitivities to those of the best BAW devices and may now be considered as lower noise and smaller volume alternatives for airborne applications.

METHODS: The AQP [5] SAWR consists of a single crystal quartz substrate supporting a surface wave device. This device is then hermeticaly sealed in high vacuum to a matching quartz cover with a glass frit. The reduction in the microphonic sensitivity of the AQP has been demonstrated using two independent techniques. First, the effect of optimizing the geometry of the glass frit of the AQP to provide the cancellation described above has been studied. Second, dual reasonator devices have been fabricated and tested. Both techniques have been successful.

Optimized Frit Geometry: Tiersten and Shick [6],[7] have shown that for a given resonator design (device frequency, number of interdigital fingers and reflective gratings, and width of the acoustic aperature) with the proper substrate size, there exists an optimum AQP frit geometry such that the vibration sensitivity normal to the plane of the device can be made to vanish, if the substrate is supported only by the glass frit. This cancellation will occur when both the aspect ratio R=a/b and the overhangs, c and d, are chosen such that the two mutually perpendicular strain distributions in the plane of the device produce equal but opposite effects on the resonant frequency of the SAWR device. A schematic defining the details of the quartz substrate, SAWR, and frit geometry is shown in Figure 1. The AQP SAWR substrate can readily be supported by the glass frit by simply mounting the package cover onto the carrier (such as a hybrid package or PC board). Another approach is to make a three piece AQP which consists of a base, substrate, and cover. The base (sealed to the bottom of the

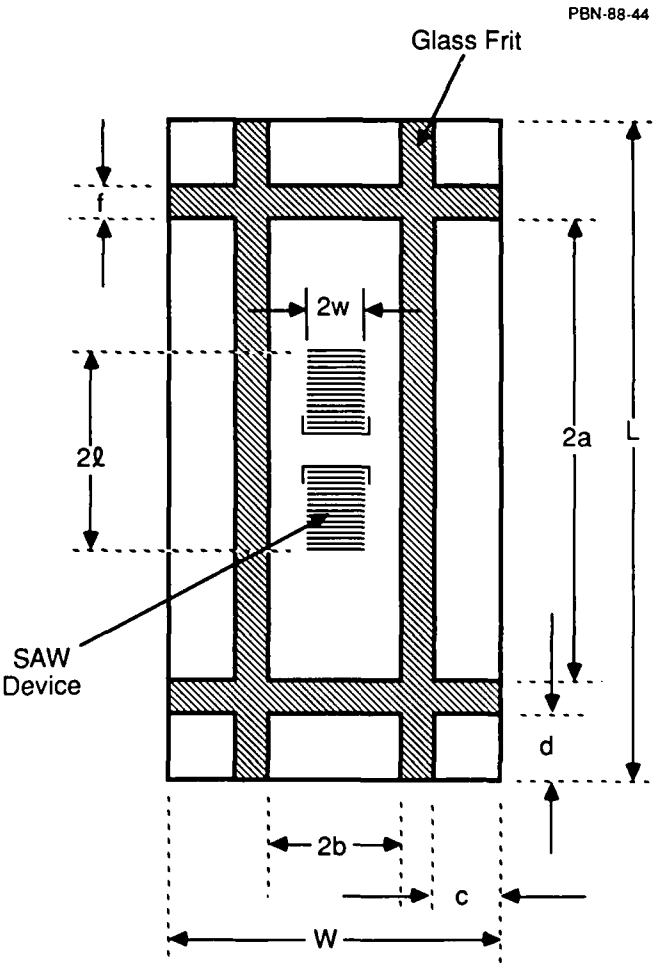



Figure 1. Schematic defining the relevant geometry of the AQP SAWR device, substrate, glass frit, and overhangs.

substrate with the glass frit) can then be mounted to the carrier, thus providing the substrate support only through the glass frit.

We fabricated a number of 2 port AQP SAWR devices with different mask designs on rotated Y-cut single crystal quartz substrates. Special care was taken to ensure that the device was printed in the center of the substrate (± 10μm) and that the glass frit was also accurately positioned (± 25 μm). These devices were fabricated either as a two piece (substrate and cover) or three piece unit (base, substrate, and cover). All substrates, bases, and covers were 0.89 mm thick. The frit geometries used for sealing the AQP were such that the ratio, R, varied from 1.0 to 4.5. A summary of the relevant parameters for these devices is shown in Table 1. Given a fixed substrate size it is not possible to change the ratio R without also changing the overhangs, c and/or d, simultaneously. Thus, quantitative correlation between changes in experimental parameters and theoretical predictions [7] are not possible with the data presented in this report.

Dual Resonator Devices: The second approach to reducing the microphonic sensitivity of the AQP SAWR was to fabricate a unique dual resonator device as shown in Figure 2. Each individual resonator was fabricated in the center of the quartz substrate (± 10μm) using a mask (402 MHz and an acoustic aperture of 130λ) with special alignment marks. The substrate and covers were 0.89 mm thick. The as-fabricated SAW devices were individually reactive ion etched [9] to the same frequency before sealing. However, due to variations in the sealing-induced frequency shift of SAWRs hermeticaly sealed in the AQP, the individual resonators will usually have slightly different resonant frequencies after sealing. Therefore, the matching quartz cover for each resonator contained an offset laser-trim pad as shown in Figure 2. This trim pad allowed utilization of a post-seal, laser-based frequency trimming technique to accurately adjust the

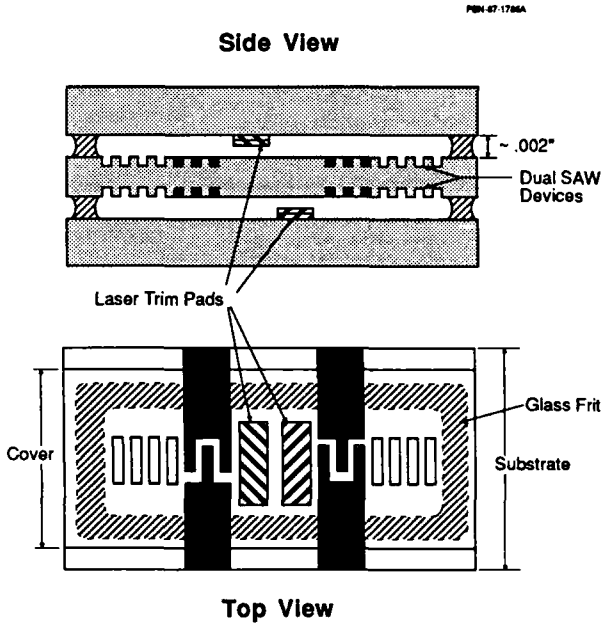



Figure 2. a) Cross section of an AQP Dual Resonator; b) Top view.

TABLE 1

| Device ID | R | a (mm) | b (mm) | c (mm) | d (mm) | L (mm) | W (mm) | l (mm) | w (mm) |
|---|---|---|---|---|---|---|---|---|---|
| E-482A† | 1.0 | 6.09 | 6.09 | 2.01 | 2.01 | 17.78 | 17.78 | 4.94 | 0.51 |
| E-168A† | 1.8 | 6.35 | 3.57 | 1.98 | 0.00 | 15.20 | 12.70 | 4.94 | 0.51 |
| E-169A† | 1.8 | 6.35 | 3.57 | 1.98 | 0.00 | 15.20 | 12.70 | 4.94 | 0.51 |
| E-132A† | 3.0 | 7.62 | 2.54 | 3.02 | 4.28 | 25.40 | 12.70 | 4.94 | 0.51 |
| E-133A† | 4.5 | 11.43 | 2.54 | 3.02 | 0.48 | 25.40 | 12.70 | 4.94 | 0.51 |
| Q-3081A↑ | 2.0 | 7.62 | 3.81 | 1.75 | 0.48 | 17.78 | 12.70 | 5.56 | 0.88 |
| Q-3080B↑ | 2.0 | 5.08 | 2.54 | 3.02 | 3.02 | 17.78 | 12.70 | 5.56 | 0.88 |

Frit width f=0.795 mm for all devices.
Thickness of all substrates, bases, and covers=0.89 mm
† = 402 MHz two piece AQP resonator with acoustic aperture of 130 λ.
↑ = 360 MHz three piece AQP resonator with acoustic aperture of 200 λ.

TABLE 1. The relevant data for the AQP device geometries tested. The definitions of the parameters are seen in Figure 1.

resonant frequency of either device [8]. After sealing, each individual device was tested with a network analyzer to document its resonant frequency, insertion loss (I.L.), loaded and unloaded Q values. The two SAWR devices were then connected in parallel to form a dual resonator by bonding 3 mil gold wires to the buss bars on opposite sides of the substrate. The dual resonators were then tested for the relevant parameters mentioned above. Detailed data for both as sealed and laser-trimmed dual resonators, and the individual devices from which they are composed, are shown in Table 2. An * in the Table indicates that the device has been laser-trimmed. Insertion loss plots of an as sealed (E151D) and laser-trimmed (E151D*) dual resonator AQP device are shown in Figures 3a and 3b respectively. Figure 3a clearly shows two resonant peaks indicating that each individual resonator underwent a different sealing-induced frequency shift. Figure 3b, shows the insertion loss plot of the same dual resonator after one device was trimmed down in frequency to accurately match its mate.

TABLE 2: DUAL RESONATOR PARAMETERS

| DEVICE I.D. | $F_o$ (MHz) | I.L (dB) | $Q_1$ | TYPE |
|---|---|---|---|---|
| E-151B | 401.551 | -4.42 | 8,234 | dual |
| E-151B-S1 | 401.550 | -6.26 | 13,517 | single |
| E-151B-S2 | 401.542 | -6.59 | 13,511 | single |
| E-151B* | 401.546 | -4.14 | 9,587 | dual |
| E-151B*S1 | 401.542 | -6.28 | 13,153 | single |
| E-151B-S1 | 401.542 | -6.59 | 13,511 | single |
| E-151D | 401.576 | -5.73 | 5,085 | dual |
| E-151D-S1 | 401.571 | -7.16 | 12,167 | single |
| E-151D-S2 | 401.558 | -6.27 | 13,622 | single |
| E-151D* | 401.563 | -4.73 | 9,082 | dual |
| E-151D*S1 | 401.559 | -7.20 | 12,213 | single |
| E-151D-S2 | 401.558 | -6.27 | 13,622 | single |

* = Laser Trimmed

Substrates were 36.7 cut quartz with identical dimensions and frit geometry as device E-168A in Table 1.

TABLE 2. The frequency, insertion loss, and loaded Q of both the individual and dual resonators before and after laser-trimming.

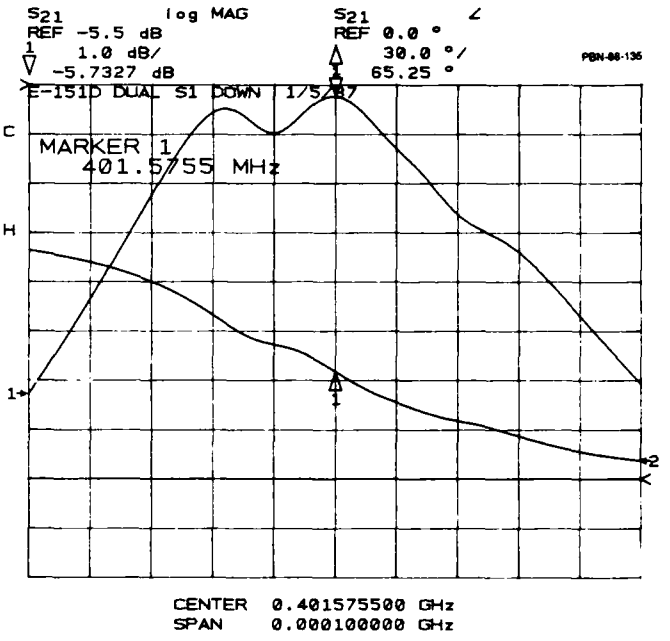



Figure 3. Insertion loss plot of both a) an as-sealed dual resonator showing two individual resonances and b) of the same dual resonator after laser-trimming.

Vibration Measurements: After fabrication, each SAWR device was placed onto the center of one face of a 63.5 mm aluminum cube using double sided tape. The cube, shown in Figure 4, was machined such that it could be rigidly fixed to a shaker table in each of three mutually perpendicular directions. Thus, three orthogonal vibration sensitivities for the mounted SAWR device could be measured as defined in Figures 5a and 5b. $\Gamma_1$ is measured when the AQP is vibrated perpendicular to the plane of the device. $\Gamma_2$ is measured when the package is vibrated in the plane of the device, parallel to the interdigital fingers, and $\Gamma_3$ is measured when the package is vibrated in the plane of the device parallel to the acoustic propagation axis. The unique nature of the AQP SAWR allows us to mount the device on the cube (or carrier) either in the right-side-up (RSU) or up-side-down (USD) position as shown in Figure 6a and 6b, respectively. The two piece single resonator AQPs were typically mounted in the USD position, while the three piece

single resonator devices were typically mounted RSU. These mounting schemes support the substrate only at the glass frit allowing the device to freely flex during acceleration. Thus, both techniques for vibration sensitivity cancellation could be tested.

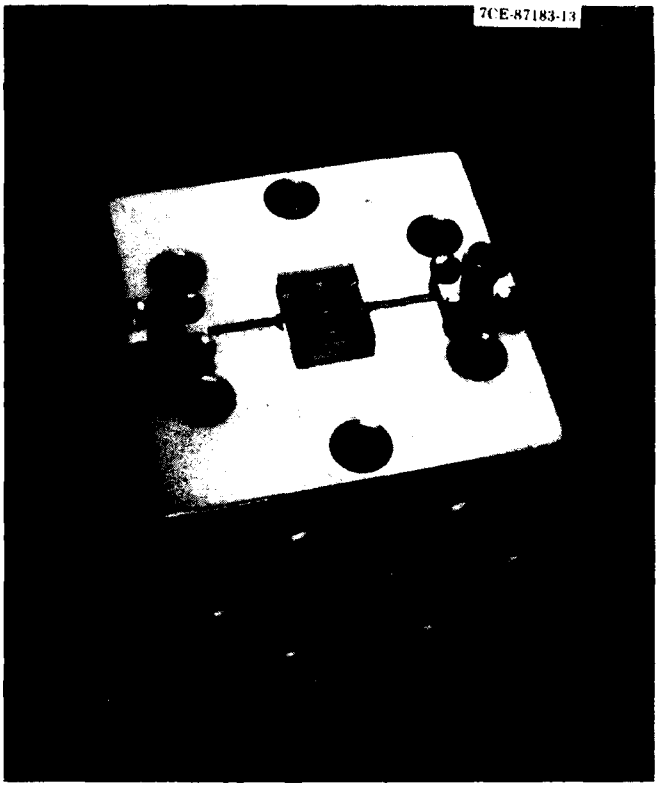



Figure 4. Photograph of an AQP SAWR mounted onto the aluminum cube.

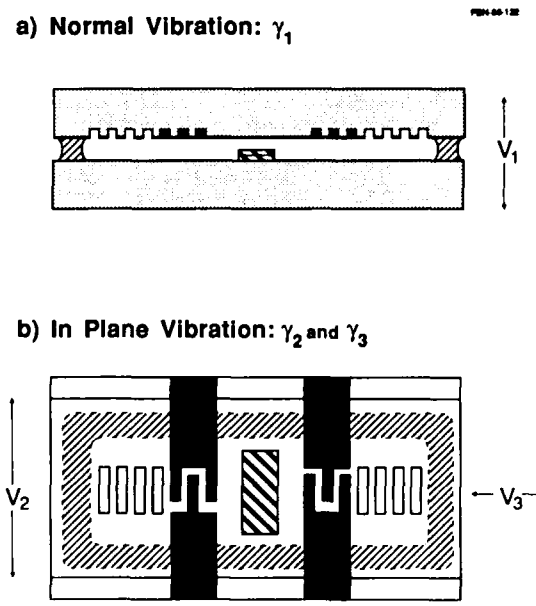



Figure 5. Schematic of the AQP defining: a) the normal vibration sensitivity $\gamma_1$ and b) the in-plane sensitivities, $\gamma_2$ and $\gamma_3$. $V_{1,2,3}$ represents the vibration directions for $\gamma_{1,2,3}$ respectively.

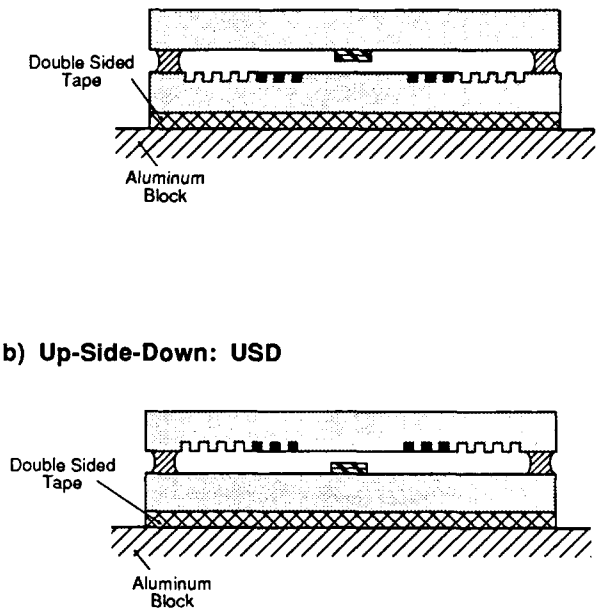



Figure 6. Schematic of a two piece AQP defining: a) RSU support and b) USD support.

Both open- and closed-loop techniques were used to measure the vibration sensitivity of the SAWR devices [3],[4]. The open-loop measurement provides for better sensitivity than the closed-loop technique and therefore the majority of the data was taken using this method. A schematic of the open-loop measurement is shown in Figure 7. The details of both open- and closed-loop methods have been recently reviewed in the literature [4] and therefore are not discussed here. The formula used for calculating the open-loop vibration sensitivity is given by:

$$\gamma = V_{rms}/(\sqrt{2}\ C\ Q_1\ G_p) \qquad (1)$$

Here $V_{rms}$ is the rms voltage measured by the spectrum analyzer, $G_p$ is the peak acceleration placed on the device, and $Q_1$ is the loaded Q of the device measured while it was mounted on the aluminum block. C is a calibration constant with units of volts/radian obtained by placing a known phase shift into the loop and measuring the corresponding voltage change. All devices were driven at +13 dBm incident on the resonator at the frequency of minimum insertion loss.

The peak acceleration of the block was measured with a Bruel and Kjaer Model 4375 accelerometer. All the data presented here was obtained using a sinusoidal acceleration of 10 g's peak for 13 fixed vibration frequencies ranging from 100 Hz up to 6 kHz.

As will be shown below, the sensitivity of these measurements was limited by phase fluctuations in the cables used to connect the device to the external electronics. Therefore, ruggedized cables with rigid connectors were used to make all the critical connections in the circuit.

The magnitude of the vibration sensitivity vector, $|\vec{\Gamma}_{3D}|$, was calculated from:

$$|\vec{\Gamma}_{3D}| = (\gamma_1^2 + \gamma_2^2 + \gamma_3^2)^{1/2} \qquad (2)$$

DUAL RESONATOR RESULTS: Figure 8 shows the normal vibration sensitivity, $\gamma_1$, obtained from device E151D wired as a dual resonator before and after the individual devices had their frequencies matched with the laser-trimming technique. Both sets of data were obtained with the dual resonators driven at the point of minimum insertion loss (see Figures 3a and 3b). As seen in Figure 8, the untrimmed dual resonator has a sensitivity of about $3X10^{-10}/g$. However, after laser-trimming the vibration sensitivity of the matched dual resonator, E-151D*, has been reduced to about $1X10^{-10}/g$.

The effect of cancellation due to the dual resonator approach can be elucidated by the following experiment. After the data for all three axes from a dual resonator device was taken, the small gold bonding wires which connect the buss bars of the two individual SAW devices in parallel were carefully cut. This leaves only one SAWR in the package electrically active. The package was then shaken again and the vibration sensitivity for the identical mount could be obtained for an individual resonator. The results of this experiment are also displayed in Figure 8. The single resonator, device E-151D-S1, was tested with S1 mounted in the USD position and the device was driven at its point of minimum insertion loss. The single resonator has a sensitivity about three times higher than that of the laser-trimmed dual resonator. As will be discussed below, the actual data for the dual resonator may be lower than that indicated.

As seen in Figure 8, the data for the untrimmed dual resonator and the single resonator from the same package are almost identical. This can readily be explained by looking at the insertion loss plot of the untrimmed dual resonator in Figure 3a. Here the higher frequency resonance is due to E151D-S1, and E151D-S2 has an insertion loss at this frequency that is down by about 3 dB. Also, the phase slope of E151B-S2 at the driving frequency will be less than E151-S1. Thus, the

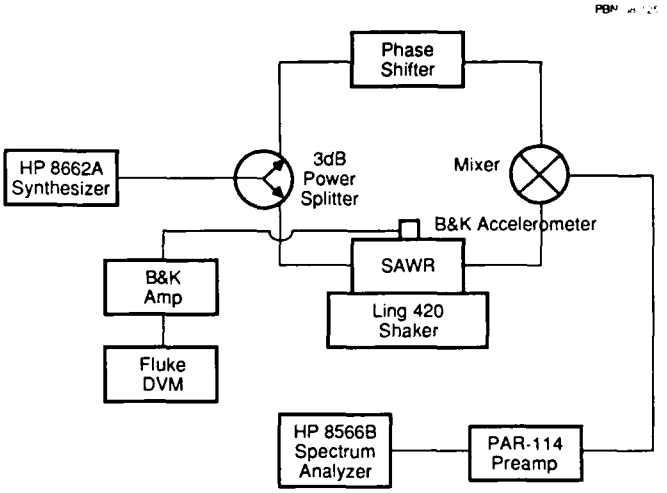



Figure 7. Schematic of the open-loop measurement system.

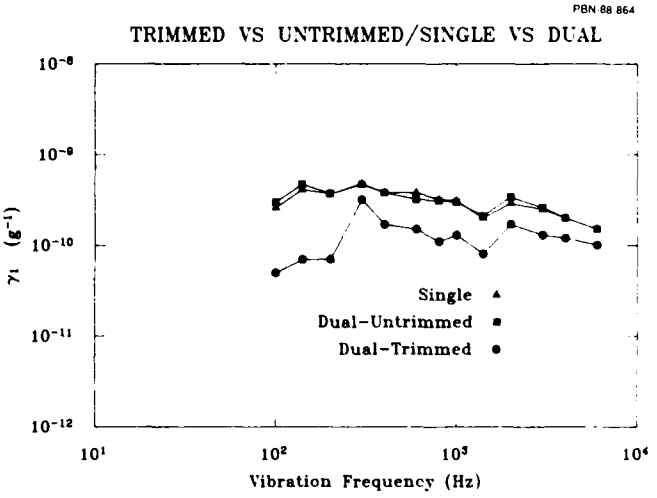



Figure 8. $\gamma_1$ for a single resonator (E151D-S1) and dual resonator before (E151D) and after (E151D*) laser-trimming.

magnitude of the vibrationally-induced phase fluctuations from S2 will be considerably smaller than S1 at the drive frequency for the dual resonator and thus, S2's contribution to cancelling the voltage produced by phase fluctuations in S1 is negligible.

Three dimensional data for device E151D* is displayed in Figure 9. Also shown is $|\vec{\Gamma}_{3D}|$, as calculated from equation (2). Three dimensional data from a second dual resonator, E151B*, is shown in Figure 10. Both of these devices exhibit extremely low vibration sensitivities in all three axes. The data from both of these dual resonator devices indicate that $\gamma_3$ is now the most sensitive direction to vibration.

It should be pointed out that when making vibration measurements on devices with low microphonic sensitivities, it is important to understand the source

of all possible contributions to phase fluctuations within the measurement apparatus. Of particular concern was the effect of the cables which connected the SAWR device through the OSM connectors to the external circuitry. In order to measure the contribution of these cables to the measured phase fluctuations the SAWR device was removed from the block and replaced with a small piece of semi-rigid coax. A 4 dB pad was then inserted into the loop to simulate the loss of the SAWR device. The Al cube with semi-rigid coax and cables (driven at the same RF frequency as E151B*) were then vibrated in the three orthogonal directions and the voltage fluctuations due to the 10 g peak acceleration were obtained as a function of vibration frequency. Using the loaded Q value from device E151B*, and the measured voltages, we calculated the effective vibration sensitivity of the cables in all three directions for this particular device. This data is presented in Figure 11. As noted, the vibration sensitivity levels are in the mid to high parts in $10^{11}$/g range. Comparing this data with that obtained from the actual dual resonator indicates that the measured sensitivity of the dual resonator devices are on the same order of magnitude as the cable sensitivity. Thus, the data from the laser-trimmed dual resonators are corrupted by the cables, and it is not clear how much better the sensitivity of these dual resonator devices might actually be. The erratic nature of the cables' phase fluctuations may also account for some of the local maxima and minima seen with both the dual and single resonator devices with low vibration sensitivity levels.

Another important feature of the AQP dual resonator is the fact that it is relatively insensitive to the mounting details. Due to differences in the coeffecients of thermal expansion between single crystal quartz and most carriers (such as hybrid packages or PC boards) one usually uses small spots of RTV rubber to hold the SAW device down. RTV has a very low modulus of elasticity and therefore will readily accommodate any strains produced by thermal expansion mismatch. This mounting scheme is radically different from the relatively uniform support afforded by the double sided tape. To determine the sensitivity of the dual resonators to more realistic supports, the following experiment was conducted. Two small strips or "rails" of double sided tape were used to support the far edges of the dual resonator package as seen in Figure 12. The supported region extended 2 mm under each of the two edges. This type of rail support is similar to that used with RTV rubber. It should be noted that this support structure readily allows the substrate to flex in its center which produces large

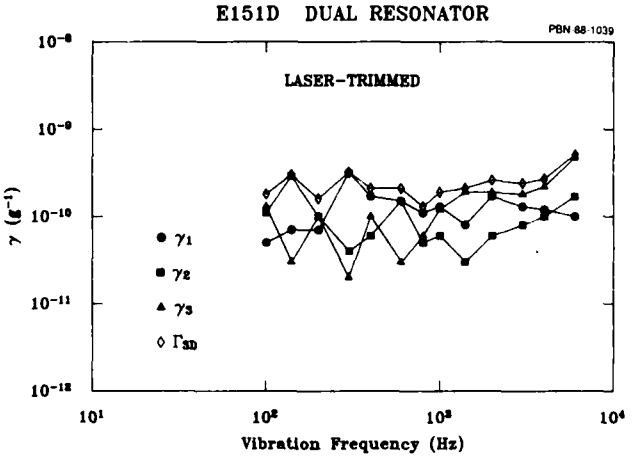



Figure 9. Three dimensional vibration sensitivity for the dual resonator device E151D*.

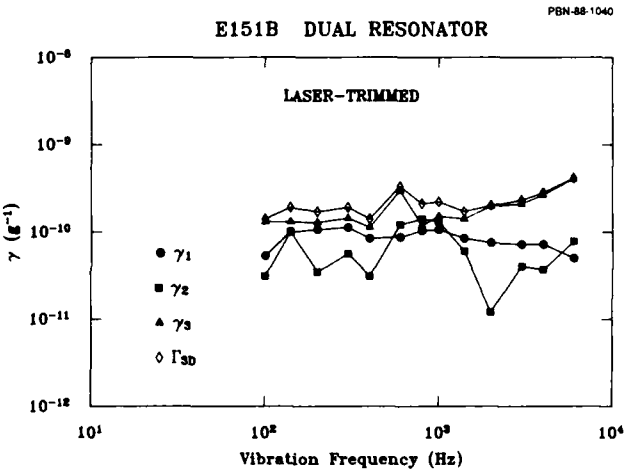



Figure 10. Three dimensional vibration sensitivity for the dual resonator device E151B*.

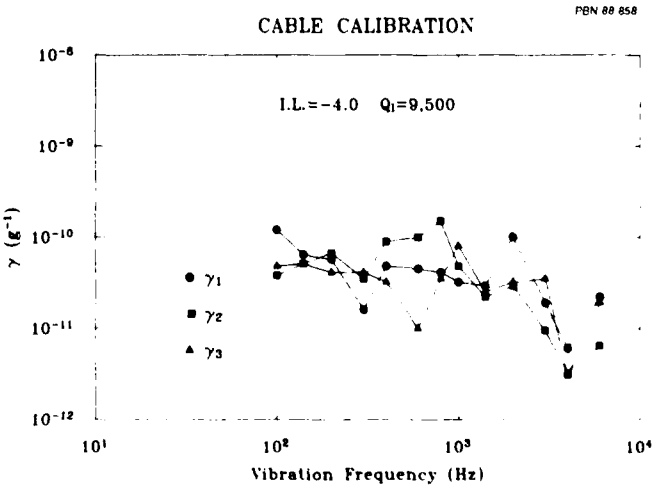



Figure 11. Vibration sensitivities of the cables for a 402 MHz dual resonator.

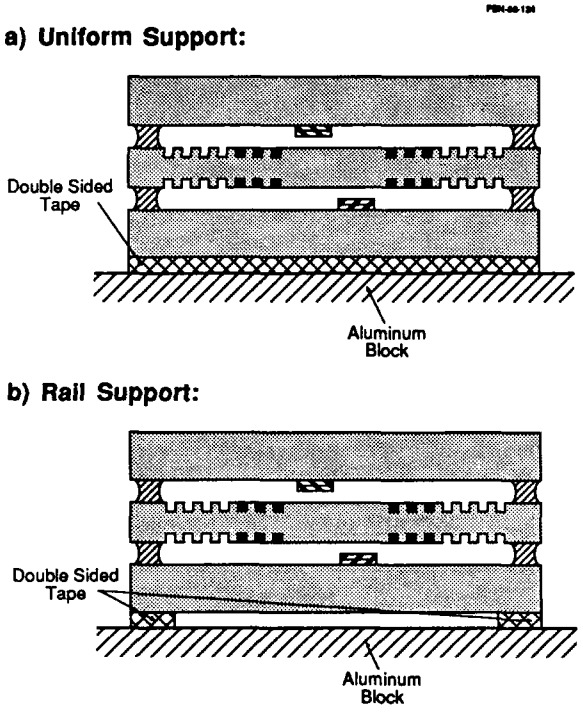



Figure 12. Schematic of a) the dual resonator AQP with uniform support and b) the dual resonator AQP with rail support.

strains on the SAWR device. Data for the normal vibration sensitivity, $\gamma_1$, for E151D* with both uniform and rail supports is plotted in Figure 13. Also plotted is similar data for E151*-S1 with the identical mounts as the dual resonator obtained after breaking the small gold bonding wires as described above. The dual resonator has a vibration sensitivity of about $3X10^{-10}/g$ with the rail support. This is about a factor of three larger than that obtained with the uniform support (neglecting cable effects). However, the single resonator's normal vibration sensitivity jumped from $3X10^{-10}/g$ up to about $5X10^{-9}/g$. These results indicate that even under less than ideal mounting conditions the dual resonator can provide partial cancellation of the normal vibration sensitivity in contrast to a three piece single SAWR AQP device. The in-plane sensitivities of both single and dual resonators are not affected by this type of mounting support.

FRIT GEOMETRY RESULTS: The three orthogonal vibration sensitivities versus vibration frequency for the seven devices listed in Table 1, all uniformly supported with either an USD or RSU mount, are displayed in Figures 14 through 20. Also, $|\vec{\Gamma}_{3D}|$ as calculated from equation (2) is plotted for some of these devices as well.

Figure 14 shows USD data for device E-428A with R=1.0. This two piece device has a square substrate with relatively small overhangs (c=d=2 mm) and the

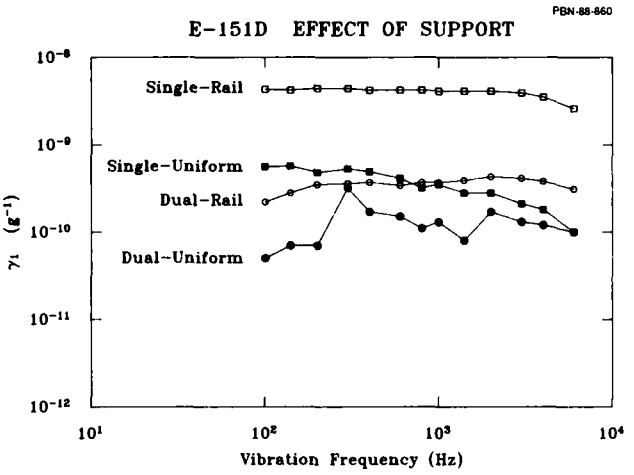



Figure 13. Normal vibration sensitivity of both a three piece single and dual resonator AQP as a function of support.

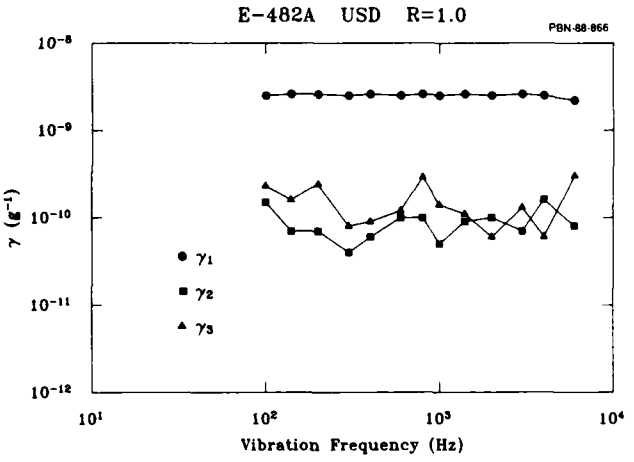



Figure 14. Three dimensional vibration sensitivity of a 2 piece AQP device E-482A with R=1.0, mounted in the USD position.

normal vibration sensitivity for this device was $2.6X10^{-9}/g$. However, the in-plane sensitivities, $\gamma_2$ and $\gamma_3$, are all below about $2X10^{-10}/g$.

Figures 15 and 16 display USD data from identical two piece AQP SAWR devices with a frit ratio R=1.8. Here the vibration sensitivity for the three orthogonal directions for both devices are small. These two Figures demonstrate how reproducible the data is for "identical" devices even with extremely low microphonic sensitivities. Also note that $|\vec{\Gamma}_{3D}|$ (not plotted in Figure 16) for device E-169A does not exceed $2.5X10^{-10}/g$ over the entire vibration range tested. This represents the lowest vibration sensitivity for any single SAW device published to date.

A calculation of the phase fluctuations produced by the cables for the single resonator devices fabricated from the 402 MHz mask was made using a 7 dB pad in place of the SAWR and a loaded Q of 14,000. The data obtained is plotted in Figure 21. Here the phase fluctuations from the cables are in the mid parts in $10^{11}$/g. Thus, the data for these two single resonators with very low vibration sensitivity levels are probably being corrupted by cable effects. The vibration sensitivity contribution from the cables for a single resonator is less than that obtained for dual resonators, as seen in Figures 21 and 11, respectively. This is due to both differences in I.L. (which change the calibration constant C) and loaded Q values for these types of devices. Again, the local maxima and minima in the magnitude of the vibration sensitivity as a function of vibration frequency for devices with low sensitivities are probably caused by cable effects. Note that when the sensitivity of a device is high (mid parts in $10^{10}$/g or above) the data is usually constant or changes monotonically over the full vibration spectrum tested ($\gamma_1$ for device E-482A in Figure 14 for instance).

In principle, the contribution from the cables could be subtracted from the data obtained with a SAWR device. However, this would require extreme care in placing the cables in exactly the same position for each measurement and a detailed knowledge of the relative phase difference between the signals from the SAWR device and cables as a function of vibration frequency would be required.

Data for two 402 MHz devices fabricated on 25.4 mm long by 12.7 mm wide substrates with frit ratios R= 3.0 and 4.5 are shown in Figures 17 and 18, respectively. This data indicates that simultaneously increasing the

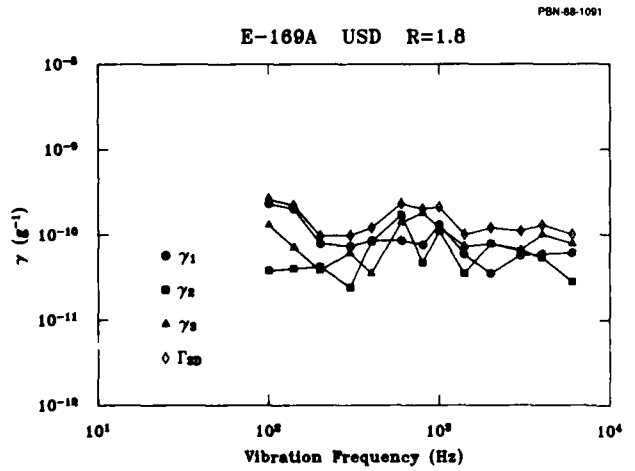



Figure 15. Three dimensional vibration sensitivity of a 2 piece AQP device E-169A with R=1.8, mounted in the USD position.

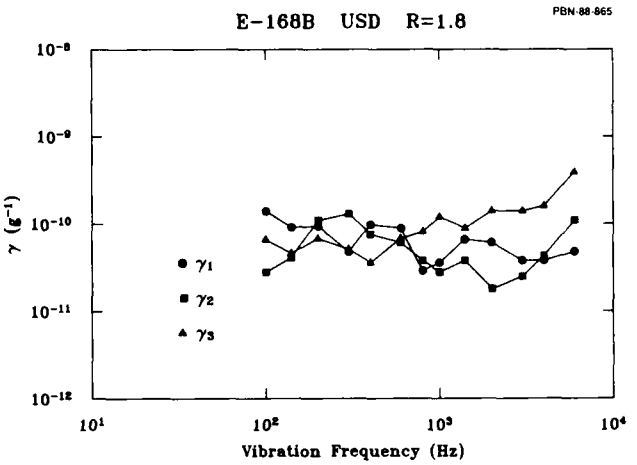



Figure 16. Three dimensional vibration sensitivity of a 2 piece AQP device E-168B with R=1.8, mounted in the USD position.

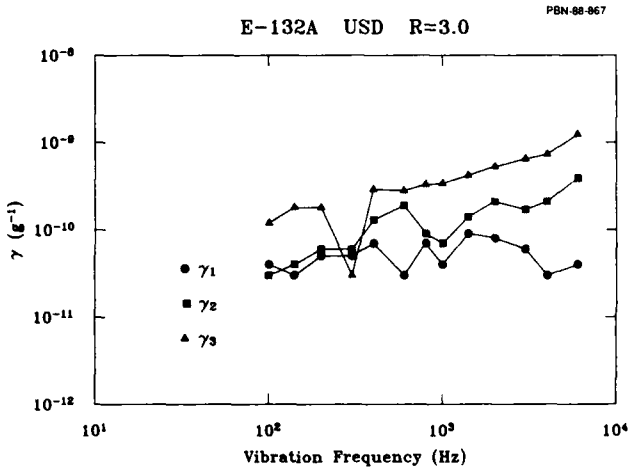



Figure 17. Three dimensional vibration sensitivity of a 2 piece AQP device E-132A with R=3.0, mounted in the USD position.

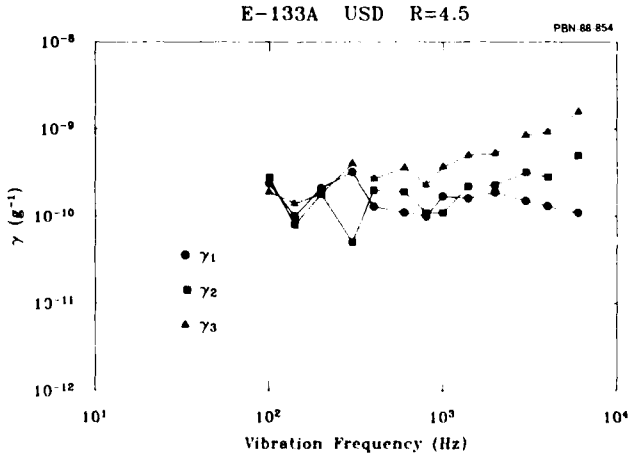



Figure 18. Three dimensional vibration sensitivity of a 2 piece AQP device E-133A with R=4.5, mounted in the USD position.

ratio R and reducing only one over-hang, namely d, has the effect of increasing the normal vibration sensitivity, $\gamma_1$, by about a factor of 2. However, both $\gamma_2$ and $\gamma_3$, are unchanged. The large dip in the curve for $\gamma_3$ occurring at a vibration frequency of 300 Hz in Figure 17 is most likely due to a coherent cancellation of the signal from the cables with that of the SAWR device.

Figures 19 and 20 represent RSU data obtained from two three piece 360 MHz SAWR's with identical substrate sizes and frit ratio's (R=2.0), but differing overhang values (see Table 1). Increasing the overhang dimensions, c and d, had the effect of reducing $\gamma_1$ and $\gamma_2$, while increasing $\gamma_3$ for this particular device and substrate design.

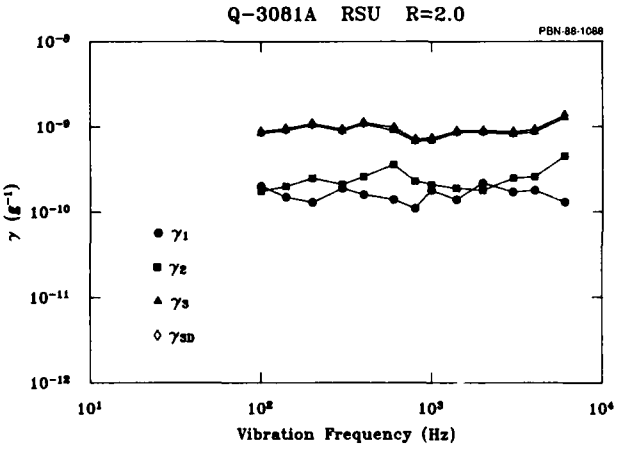



Figure 19. Three dimensional vibration sensitivity of a 3 piece device Q-3081A with R=2.0, mounted in the RSU position.

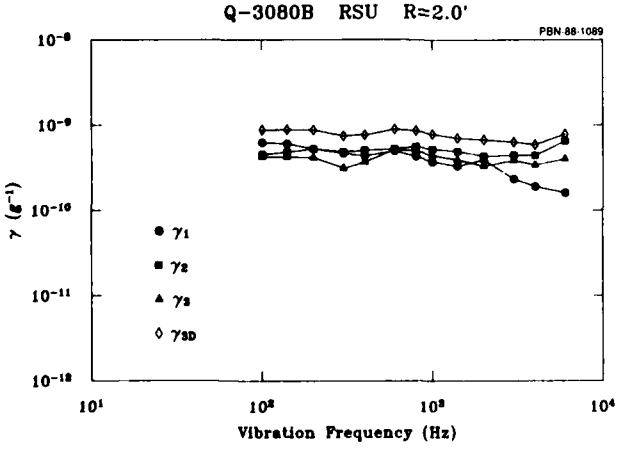



Figure 20. Three dimensional vibration sensitivity of a 3 piece device Q-3080B with R=2.0, mounted in the RSU position. Note that this device has different overhangs than Q-3081A seen in Figure 19.

A comparison of USD versus RSU mounting for a two piece AQP SAWR, E-132A, is shown in Figure 22. The data for the RSU mount is usually higher than that obtained with the USD mount, indicating that a true uniform support is difficult to achieve.

Closed-loop vibration sensitivity measurements were also obtained for some of the devices before dismounting them from the aluminum cube, after the open-loop measurements were completed. During the closed-loop measurement only the SAWR device was vibrated, and the rest of the oscillator electronics (connected by the ruggedized cables) were in a relatively quiescent environment. The normal vibration sensitivity of device E-133A obtained with both the open- and closed-loop technique are shown in Figure 23. Excellent agreement between the two measurements was obtained, with the closed-loop technique giving slightly higher values.

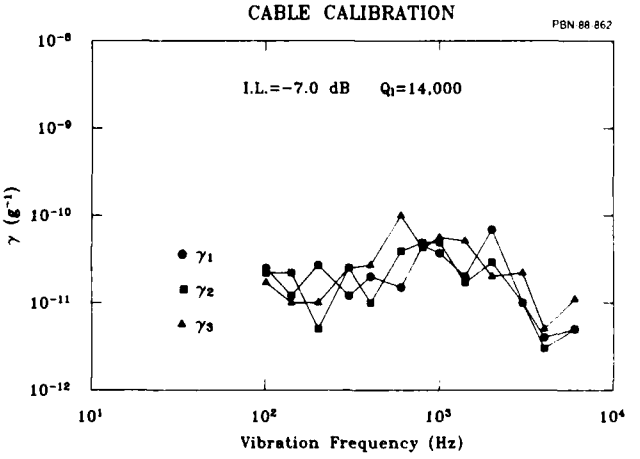



Figure 21. Vibration sensitivities of the cables for a 402 MHz single resonator device.

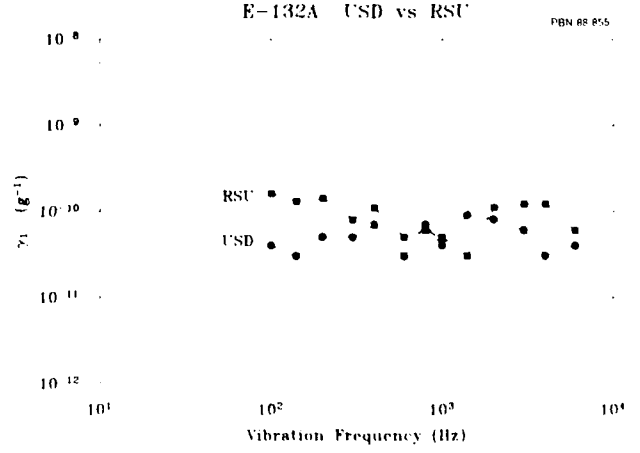



Figure 22. Comparison of RSU and USD mounting for device E-132A.

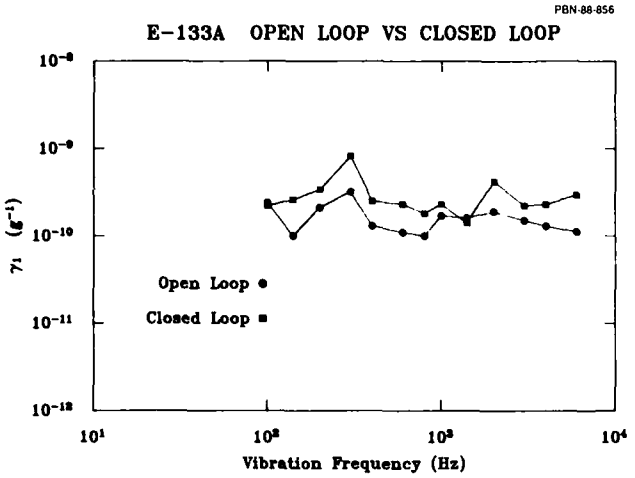



Figure 23. Open- and closed-loop normal vibration sensitivity for device E-133A.

DISCUSSION: The results from these experiments have demonstrated that the vibration sensitivity of SAWR devices can be made comparable with that of bulk wave resonators. However, it should be pointed out that the results reported here were obtained for SAWR devices which were mounted onto a solid aluminum block. The block will provide a more rigid support than that obtained with either a hybrid flat pack or PC board type carrier. On the other hand, the results obtained from the dual resonator device with rail supports, as seen in Figure 12, indicate that this type of device is relatively insensitive to the mounting details. This clearly contrasts with that of the single SAWR device with rail type supports obtained from the same package. It should be pointed out that a two piece AQP SAWR is more flexible than a three piece AQP (such as the single resonator tested in the dual resonator package) and it is likely that these two piece AQPs will be even more sensitive to mounting details than a similar three piece unit. Since an oscillator's vibration sensitivity is primarily driven by the sensitivity of the resonant device (SAWR or BAWR), it seems likely that a dual resonator approach will be required to obtain a complete SAWR oscillator with minimum vibration sensitivity.

Besides having a low sensitivity to the mounting details, the dual resonator has some other practical advantages over single resonators as well. First, by placing two resonators in parallel, the effective acoustic aperture of the device is doubled. Thus, the power handling capability of the dual resonator is increased by a factor of two. Second, the insertion loss of the dual resonator will be smaller than that of an individual resonator (if the acoustic apertures are not to wide) as noted in Table 2. These three-piece dual resonator devices will also be less sensitive to both atmospheric pressure and transient temperature fluctuations than a standard two piece single resonator AQP device. However, a three piece single resonator AQP (base, substrate, and cover) will be less sensitive to these variations as well.

The main problem with the dual resonator design is that difficult fabrication techniques are required. As indicated in Figure 8, the frequency of the two individual resonators need to be well matched, thus post-seal frequency trimming using the laser-based technique is required. Although it has not been demonstrated as yet, it is also likely that both the insertion loss and loaded Q values will need to be matched as well, in order to obtain the maximum amount of cancellation between the two devices.

The results obtained from altering the frit geometry of the AQP, Figures 14 to 20, strongly support the theoretical work of Tiersten and Shick [6],[7]. These results demonstrate that the normal vibration sensitivity of a SAWR device rigidly supported by a glass frit (as it is with an USD mount for the AQP) can be below $1X10^{-10}/g$. Measuring lower values than presented here is not possible at this time, since the current experimental apparatus does not permit sensitivities below about $6X10^{-11}/g$ to be measured (Figures 11 and 21), and further improvements in reducing the vibration sensitivity of the AQP will only be realized by vibrating a full SAWR oscillator, thereby eliminating the cable contributions. These results do indicate that both the frit ratio R, and the overhang dimensions, c and d, can strongly affect all three vibration sensitivities as seen from Figures 14 through 20. However, it is awkward in practice to systematically vary the ratio R without varying the overhangs as well. Thus, a more quantitative agreement between theory and experiment will be cumbersome to produce.

It is interesting to compare the data from a dual AQP resonator with that of a single two piece AQP resonator of the same geometry (device, substrate, and frit geometry) as seen in Figures 10 and 16. This particular package geometry appears to provide excellent sensitivities for both resonator types and as mentioned above the data for these devices are at the limits of the measurement technique. Thus, it can not be determined whether the in-plane sensitivities, $\gamma_2$ and $\gamma_3$, are significantly reduced by the dual resonator approach. This question will be best answered by fabricating a dual resonator with a device and package design similar to that used in Figure 20. It should be

emphasized that even though the single resonator devices of this package design provide equivalent sensitivities as their dual resonator counter-part, the dual resonator device has the clear advantage when mounted with realistic support structures.

The data over a large number of devices (some not presented here) indicates that $\Gamma_2$ (vibration sensitivity parallel to the IDT's) is the least sensitive direction to variations in package design, as its value varied from $5X10^{-10}$/g to below $1X10^{-10}$/g for the different package types tested. In contrast, $\Gamma_1$ could be varied over a wide range ($3X10^{-9}$/g to below $1X10^{-10}$/g) by varying the package type or number of resonators. $\Gamma_3$, is also sensitive to package design, and appears to be the most difficult component of $\Gamma_{3D}$ to suppress. Further reductions in $\Gamma_1$ could probably be obtained by increasing the thickness of the quartz substrate. However, since the vibration sensitivity now appears to be dominated by $\Gamma_3$, there does not appear to be any advantage to increasing substrate thickness.

CONCLUSIONS: Two techniques for reducing vibration sensitivity of the AQP SAWR have been demonstrated. Both techniques provide SAWR devices with vibration sensitivities below $3X10^{-10}$/g and a realization of further improvements in these levels have now been limited by the measurement technique. The sensitivity of these devices represents the lowest values of any published SAW device to date.

The normal vibration sensitivities of AQP SAWR devices are strongly influenced by the frit geometry in good agreement with current theory [7]. The sensitivity of an AQP SAWR can be made extremely low by optimizing the package geometry. However, in-plane sensitivities, especially $\Gamma_3$, are still higher than theoretical predictions [10].

The dual resonator approach, although complicated, also provides extremely low vibration sensitivities when the frequencies of the two resonators are well matched. This type of device design may ultimately be more attractive than a single resonator, due to its relative insensitivity to the mounting details.

SAWR devices have now demonstrated vibration sensitivities equivalent to the best BAW oscillators. However, demonstration of SAWR oscillators with vibration sensitivities compatible with the SAW devices shown here is needed before SAW technology can be incorporated into high performance airborne applications.

ACKNOWLEDGEMENTS: Thanks to Drs. Joe Callerame and Gary Montress for their encouragement, helpful discussions, and manuscript suggestions. Special thanks to Mssrs. Ernest Sabatino III and Patrick J. Harkins for fabricating the SAWR devices and also Drs. H. Tiersten and D. Shick for several interesting discussions and suggestions throughout the course of this work.

REFERENCES:

1. R. L. Filler, "The Acceleration Sensitivity of Quartz Crystal Oscillators: A Review," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Special Issue on Frequency Control (Part II.) Vol. 35, No. 3, pp. 297-305, May 1988.

2. T. E. Parker and J. Callerame, "Sensitivity of SAW Delay Lines and Resonators to Vibration," in Proceedings IEEE Ultrason. Symp., 1981, pp. 129-134.

3. H. F. Tiersten, D. S. Stevens, and R. D. Weglein, "Acceleration Sensitivity of Surface Acoustic Wave Resonators," Appl. Phys. Letts., vol. 48, pp. 97-99, 1986.

4. B. H. Kolner, "Microphonic Sensitivity of SAW Resonators," in IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Special Issue on Frequency Control (Part II.) Vol. 35, No. 3, pp. 365-371, May 1988.

5. T. E. Parker, J. Callerame, and G. K. Montress, "A New All Quartz Package for SAW Devices," in Proceedings of the 39th Annual Symposium on Frequency Control, 1985, pp. 519-526.

6. H. F. Tiersten and D. V. Shick, "An Analysis of the Normal Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators Rigidly Supported Along the Edges," in Proceedings of the 41st Annual Symposium on Frequency Control, 1987, pp. 282-288.

7. H. F. Tiersten and D. V. Shick, "An Analysis of the Normal Acceleration Sensitivity of ST-Cut Quartz Surface Wave Resonators with the Substrate Extending Beyond the Support," to be published in Proceedings of the 2nd European Frequency and Time Forum, Neuchatel France, March 1988.

8. J. A. Greer, T. E. Parker, M. Rothschild, and D. J. Ehrlich, "A Novel Technique for Trimming the Frequency of a Sealed Surface Acoustic Wave Resonator," in Proceedings of the 41st Annual Frequency Control Symposium, 1987, pp. 351-359.

9. W. J. Tanski, "Surface Acoustic Wave Frequency Trimming of Resonant and Traveling-wave Devices on Quartz," Appl. Phys. Lett., Vol. 39, No. 1, pp. 40-42, July 1, 1981.

10. D. V. Shick 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," to be Published in Proceedings of the 42nd Annual Frequency Control Symposium, June 1988, Baltimore, MD.

42nd Annual Frequency Control Symposium - 1988

# SAW TAPPED DELAY LINES FOR NEW POTENTIAL CIRCUIT APPLICATIONS

William J. Skudera, Jr.

U.S. ARMY LABCOM
Electronic Technology & Devices Laboratory
Fort Monmouth, NJ 07703-5000

## SUMMARY

Two new potential applications for SAW tapped delay lines will be discussed in this paper. The first deals with using a SAW differential tapped delay line (TDL) to detect an important Property of a single BPSK type signal, namely its chip rate. For this application, differential delay lines are required since it would be too difficult to build standard type delay lines having delay times of the order of 50 ns due to their feedthrough and triple transit problems. After the BPSK signal is fed into the TDL, the resultant output of the TDL is subsequently fed into either a summer or a mixer device whose output can then be viewed on an oscilloscope. Both the summer and mixer circuits employing the SAW differential TDL are analyzed and results are shown for the summer circuit along with the impulse response for the tapped differential delay line by itself. The summer circuit was found to be superior over the mixer for a specific chip rate primarily because it did not require high input power signal levels, however, it does require precise phase control that apparently can be obtained in this type of SAW device. In addition, chip rate discrimination is briefly discussed and remains to be studied further for both types of circuits.

The second potential application is concerned with using a TDL in an amplifier feedback loop circuit to replace the inverse transform of a Chirped Fourier Transform (CFT) processor for CW input signals. The problem of the inverse transform is that is loses all the processing gain achieved by the forward transform, i.e., of the order of 20 dB can be lost. However, if one uses a TDL design with the proper amount of delay and recirculates the delayed signal in a manner described in this paper, it should then be possible to create a continuous CW signal, and thus improve the CFT processor's sensitivity. This paper also discusses some of the limitations of the TDL method. In addition, a staggered channelized method will be discussed for the SAW TDL so as to ease fabrication and minimize reflections utilizing a minimum number of output taps.

## INTRODUCTION

The circuit shown in Figure 1 has been previously used by J. F. Lutz[1] to detect FM signals, whereby, the delay time of the delay line itself is required to be a nominal $\pi/2$. The questions that this paper has attempted to address are: can this or a similar type of circuit be used to detect the chip rate of BPSK signals, also if the circuit is satisfactory then what is the delay time value required for the delay line, and finally can a SAW device be utilized for the delay element? In an attempt to answer these questions, the circuit was analyzed and a delay time was found that was different from the FM value. This new delay time was considerably greater then the FM value, but considerably short (on the order of 25 to 100 ns depending upon the chip rate) for constructing a practical SAW single channel fixed delay device, since a minimum of 300 ns or more is usually required. Therefore, a differential type of SAW device had to be considered using two channels or two separate output taps for each discrete chip rate of a BPSK signal. In practice, the input signal is split and fed into two distinct delay channels that differ by the desired delay. The utilization of the differential TDL device also resulted in a slight modification in the final circuit, as shown in the following heading. A SAW TDL device was fabricated and evaluated in the modified circuit using a summer or combiner device. These initial results, that apparently showed good differential phase control, led to consideration of the second potential SAW TDL application, namely replacing the inverse transform of a CFT processor.

The CFT processor is generally well known[2] for initially doing a chirp-Z transform of several CW instantaneous input signals, wherein, unwanted signals are gated out and only the desired signal is allowed to be inputted into the inverse transform. The signal coming out of the inverse transform should then be an uncorrupted representation of the desired signal that was originally fed into the input of the CFT. The problem of the inverse transform, as stated above, is that it loses all the processing gain of the forward transform because it effectively performs an expansion of the compressed signals that result after the forward transform. It can be shown, however, that the compressed pulse signal that comes out of the forward transform still contains the input carrier signal or an LO product of the input signal. Therefore, if one uses a TDL with the proper amount of delay and recirculates a delayed version of this compressed pulse via a feedback loop and maintains phase coherency, it should then be possible to fill in enough adjacent delayed pulse signals that overlap in such a manner as to create a continuous CW signal, and thus improve the CFT processor's sensitivity. The phase coherency required is directly related to the differential phase control that apparently can be achieved in the TDL. Also, it is desirable in both proposed applications that the number of taps be kept to a minimum and spaced as far apart as possible in each channel so as to minimize distortion effects caused by reflections. However, wide tap spacings will yield narrow operating bandwidth,[3] and, therefore, to avoid this apparent problem, a multiple differential TDL approach is described in the following section, for this second potential application.

## DISCUSSION AND RESULTS

Figure 1 shows the basic concept for detecting the chip rate of a BPSK signal where $V_1$, $V_2$ and

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

$V_3$ are the inputs to the multiplier or mixer. Table 1 lists the equations into and out of the mixer and low pass filter (LPF) respectively, where $V_1$ is the undelayed BPSK signal and $V_2$ is the delayed BPSK signal. The $V_{out}$ equation represents the resultant signal from the LPF, and shows that if the "V" terms are set equal to one then only a quarter of the amplitude of the initial BPSK signal results because energy goes into harmonics that are filtered out. Table 1, also shows that the $V_{out}$ equation is optimized if $T = t_c/2$ where; T is the device delay time, $t_c$ is the chip rate of the BPSK signal, $W_o$ is the RF carrier, and $W_c$ is the chip frequency.

Figure 2 shows in graphical form what the equations in Table 1 represent. Figure 2(A) shows that if a delay time of 50 ns is used for the delay line of Figure 1, then it should be optimized for a 10 MHz chip rate signal. Figure 2(B) shows the delayed ($V_2$) and undelayed ($V_1$) BPSK signals where; a +V represents a zero phase shift and a -V represents a 180 degree phase shift respectively for the first eleven bits of a BPSK signal. When the delay difference ($V_2$ - $V_1$) equals half a bit or chip time length, then output pulses occur from the mixer or LPF only over the duration where the two signals have dissimilar phase states, and the LPF output is zero everywhere else.

Figure 3 shows a sketch of the type of dual channel (differential) SAW delay line that had to be designed because "$\Delta t$" is too short to utilize a single channel device. Figure 4 is a modified version of Figure 1 so that the SAW TDL can be accommodated. Initially a true multiplier circuit was intended for this measurement, however, this was not readily available. Therefore, a double balanced mixer was considered, but this idea was discarded because the mixer requires at least one port to be at a relatively strong signal level so that its conversion loss can be kept to a minimum. In addition, the output of the LPF reduces the signal amplitude, as stated previously. Thus a summer or combiner device was considered to replace the mixer, and the results along with an explanation of how these results can be interpreted are addressed further on in this paper.

Figure 5 illustrates how a single SAW device could be implemented to detect several distinct chip rate signals. For example, the differential delay between output transducers $D_7$ & $D_1$, $D_5$ & $D_3$, $D_8$ & $D_2$, and $D_6$ & $D_4$ could be used to detect four different chip rates. It is to be noted that up to 24 output tap combinations could be used, however, each would require its own summer or multiplier and half of these combinations would lose half of their output signal amplitude, relative to the other group of combinations, since the input transducers are bidirectional. Also, logic decision circuits would be required to fully implement this approach.

The SAW differential delay line, shown in Figure 6, was designed and fabricated for two different chip rates (10 & 20 MHz) centered at 100 MHz. It operates by feeding a BPSK signal into the top or bottom center connector, and than the resultant signals coming out of the left and right connectors are fed into the summer device. The left and right transducers differ in delay by half the chip rate. The advantage of this center fed method is that it does not require the power divider shown in Figure 4.

Figure 7 shows the left (output 1) and right (output 2) impulse responses of the Figure 6 device where the output 2 transducer is differentially delayed by 50 ns with respect to the output 1 transducer. The responses are reasonably close except for a small spurious signal that appears near the right null of the main lobe. The signal has not been identified, but is believed to be a bulk reflected type. The overall insertion loss of 35 dB was higher than expected, and is believed due to a higher than normal metal film resistance that occurred during metalization, as well as, the result of using a large device fixture that required relatively long wire leads.

Figure 8 shows how the BPSK signal was generated and detected. A BPSK signal was generated by impulsing a 127 bit pseudo-noise (PN) SAW tapped delay line, and then feeding this BPSK signal to the dual differentially tapped delay line (DDL) shown previously in Figure 6. The resultant DDL output signals are then fed into a summer device for reasons stated previously. The top waveform of Figure 9 shows part of the BPSK generated signal, whereby, the nulls indicate the point at which the signal changes phase. The lower part of Figure 9 shows detected pulses that come out of the summer of Figure 8; these pulses are all of constant pulse width and occur at the phase change points of the BPSK signal. It is to be noted that the lower level signal that is seen to appear in the regions between these detected pulses is believed due to a feedthrough signal that occurs from the input to output transducers of the SAW tapped DDL device. It should be possible to remove most of this signal by using better RF shielding between transducers, separating the transducers further apart, and lowering the overall device insertion loss.

Figure 10 illustrates a simplified explanation of how the Figure 8 circuit operates. The Figure 10 concept is based solely upon the phase of the BPSK type signal, and is done for simplicity of understanding since the terms are complex because they do not drop out as in the case of the multiplier circuit of Figure 4. Figure 10(A) & 10(B) represent the two inputs into the summer of Figure 8 that have a differential delay of T'/2, where T' is the length of a chip or bit. Figure 10(C) is the summed resultant of the (A) & (B) signals, whereby, no output should appear except in the region between the phase change of the (A) signal and the (B) signal. The limited bandwidth of the Figure 8 circuit tends to yield triangular pulses instead of trapezoid type pulses. The real part of the sum terms are the same as $V_1$ & $V_2$ of Table 1 and again appear to be an optimum at $t_c/2$, however, a preliminary check of this expression indicates that the amplitude appears to fall off more like a Sinc function and could effect the unambiguous value of the chip rate; further study of this expression is required.

To achieve a high signal to noise result as indicated in Figure 10(C), requires not only that the feedthrough, as mentioned above, be reduced, but also that the output tap transducers be precisely located within an assumed +/- ten degrees or less of their designed relative locations. It is also known that electron beam photomask machines are capable of "butting" accuracies to a 0.1 micron or less for each 3.2 mm square area, and accuracies to better than 0.05 micron within this area.[4] A simple calculation, then yields relative tap placement accuracy for a SAW mask designed for 100 MHz center frequency to be within this required tolerance. In addition, if one can achieve good differential phase control, then a second potential application

is postulated and described below.

Figure 11 is a block diagram of a standard chirp-Z forward transform followed by a modulator (mod) used for excision of signals, and whose output is fed into a reverse transform to restore the desired signal.[2] In Figure 11, the compressed signal $F_2$ is allowed to be passed through the modulator such that a sine wave signal of $F_2$ emerges after the reverse chirp-Z transform. The main problem of this standard processor is that all the processing gain of the forward transform is lost via the reverse transform as explained previously. The top result of Figure 12 illustrates what the output from the forward transform of Figure 11 looks like for four CW input signals, and the bottom result of Figure 12 shows the result after the modulator removes the unwanted signals and allows a desired signal $F_2$ to pass to the inverse transform. It is obvious, then that $F_2$ would have a poor S/N after the reverse transform. Figure 13 is a closer look at a compressed transform signal, and it shows that it contains RF cycles that are equal to or related to the original signal depending upon the SLO used.

Figure 14(A) illustrates how a TDL could be used to possibly replace the inverse transform of Figure 11. In this figure the forward transform and modulator are the same as in Figure 11, however, the TDL with a feedback loop is used to replace the inverse transform. The TDL taps have to be designed such that when a compressed pulse enters the TDL then it outputs a series of three or four replica pulses that are overlapped at the half amplitude values and recirculated such that a CW type signal appears at the output, as shown in Figure 14(C) & (D). This recirculation continues until a second compressed pulse is outputted from the modulator, as shown in Figure 15(A) & (B) occurring at a period T equal to the forward transform SLO period. The combined operation of the switch (S) and the delay line (DL) are used to interrupt and update the TDL when required. The updating of the TDL is described in the following paragraph. However, the success of Figure 14(A) circuit depends upon the phase coherency of the overlapped pulses, and this in turn is dependent upon the relative tap placement accuracy as previously discussed. Also, it is to be noted that the pulses that occur in Figure 14(B) possesses frequency sidelobes,[2,5] not shown, that will cause distortions in the Figure 14(D) output. However, the implementation of a weighting function[6] that can reduce the sidelobe level to -20 dB or more is a fairly standard technique used in most chirp-Z transforms. Therefore, this low sidelobe level should cause only a 10% or less distortion in the recreated signal, and the distortion could be reduced further with the use of a low pass filter, should it become necessary. The LPF could also be useful in reducing any phase distortions, caused by any tap placement error, in the TDL output signals since both this and the sidelobe distortion can be viewed as high frequency type distortion effects.

Figure 15 illustrates a concept of using four parallel TDL channels whose outputs are connected. This design permits each channel to have only a few widely spaced taps and thus, reduces tap reflection problems. The number 1a, 2a, 3a, and 4a, etc., taps are slightly offset from each other such that the output pulses appear as indicated in Figure 15, and recirculated as before. After, for example, ten recirculations of the group of 12 overlapped pulses (1a to 4c) a second compressed pulse would appear after the modulator, and the switch has to stop the recirculation routine long enough for the TDL memory to be updated. This update operation is graphically described in (B), (C), and (D) of Figure 15. The update operation is automatically accomplished by using a direct line from the modulator to immediately turn the switch OFF and a DL with a fixed delay, equal to the TDL memory time plus any small circuit delay, to turn the switch ON, such that there is no interruption in the Figure 15 output signal. In addition, it should be mentioned that each time the group of pulses is recirculated, the signal to noise ratio (S/N) is expected to be reduced after the first recirculation by an additional 0.1 dB or more per recirculation. The consequence of this S/N problem is discussed further in the conclusions.

Figure 16 illustrates a more compact way of putting the four channels of Figure 15 on one SAW substrate using a wide input transducer. This figure is self explanatory and again uses a minimum of relatively widely spaced taps to minimize reflection effects. The primary disadvantage of this arrangement over that of Figure 15, is that it requires one additional external summer device that adds to the complexity of the overall circuit, and also it could have increased differential phase control error because half of the taps are separated further apart.

## CONCLUSION

Two potential applications for SAW differentially tapped delay lines have been discussed in this paper. The first application utilizes a TDL SAW device in a circuit to detect the chip rate of BPSK type signals, wherein, the outputs of the SAW device is subsequently fed into either a summer or mixer device. The analysis performed indicated first; that the optimum differential delay for the TDL is half the chip rate, and second, that the summer detection circuit is more efficient over the mixer circuit, as would be expected, however, the unambiguous value of the chip rate may not be as good, relative to the mixer circuit. More importantly is the fact that the initial results indicate that the SAW TDL device can be fabricated with reasonably accurate differential phase control. It is to be noted, however, that additional effort is required, as previously indicated, in order to determine the full potential of this phase control. Also, the TDL- summer detection method is believed to be limited to an upper center frequency of approximately 200 MHz since, the relative tap placement accuracy will diminish as the wavelength decreases.

A second potential application, namely using a similar TDL device with feedback to replace the inverse chirp-Z transform, has also been discussed in this paper. This application is postulated, based again upon achieving the required differential tap phase control, so as to achieve coherent phase, in adjacent replicated pulses. The number of recirculations appear to be limited since the signal to noise ratio will degrade slightly for each recirculation, however, this method appears to be useful for replacing moderate (up to 20 dB) time-bandwidth inverse transforms. In addition, a compact TDL structure was indicated that had widely separated taps to reduce reflection problems, while it operated as if it effectively had closely spaced taps for wideband operation.

## REFERENCES


1. J. F. Lutz, "Synchronous Delay-Line Detector Provides Wideband Performance," Microwaves & RF, pp. 71-79, November 1982.

2. R. M. Hays, et al, "Surface Wave Transform Adaptable Processor System," 1975 Ultrasonics Symposium Proceedings, IEEE Pub. # 75 CHO 994-4SU, pp. 363-370, September 1975.

3. H. Matthews, editor, "Surface Wave Filters: Design, Construction, and Use," John Wiley & Son, Inc., pp. 319-320, New York, 1977.

4. W. Goodreau, of USALABCOM, Private Communication.

5. C. E. Cook and M. Bernfeld, "Radar Signals," Academic Press, pp. 130-172, New York, 1967.

6. C. Lardat, "Improved SAW Spectrum Analyzer with 80 dB Dynamic Range," 1978 Ultrasonics Symposium Proceedings, IEEE Pub. # 78 CH 1344-1 SU, pp. 518-521, September 1978.


TAPPED DELAY LINE-MULTIPLIER DETECTION EQUATIONS

$$V_1 = \frac{V(t)}{V} A\cos(\omega_0 t)\cos(\omega_c t)$$

$$V_2 = \frac{V(t-T)}{V} A\cos(\omega_0 (t-T))\cos(\omega_c (t-T))$$

$$V_3 = \frac{V(t)V(t-T)}{V} A^2(\cos(\omega_0 T)+\cos(2\omega_0 (t-T/2)))(\cos(\omega_c t)+\cos(2\omega_c (t-T/2)))$$

$$V_{out} = \frac{V(t)V(t-T)}{4V^2} A^2\cos(\omega_0 T)\cos(\omega_c T)$$

"T" is selected so that the cosine terms equal 1

Therefore: $\omega_c T = \pi$ or $T = t_c/2$

Where: V(t) = +V or -V

TABLE 1. Lists the equations into and out from the multiplier of Fig. 1 for a BPSK input signal.

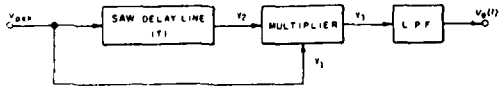



Fig. 1. A simplified block diagram showing the concept of a BPSK transition detection circuit using a SAW delay line.

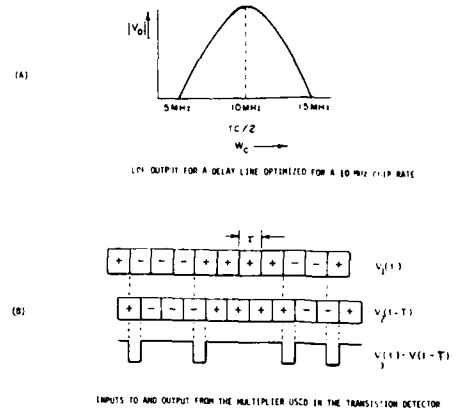



Fig. 2. Explains in graphical form the Table 1 equations. (A) is the Fig. 1 output response for a delay line of 50 ns, and (B) shows how the delayed and undelayed bits are multiplied.

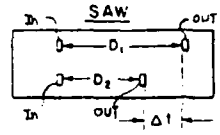



Fig. 3. Illustrates a dual channel SAW differential delay line design.

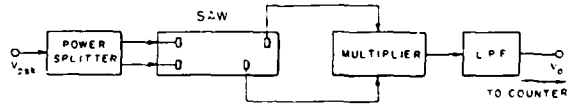



Fig. 4. A block diagram of a SAW transition detector for a given chip rate.

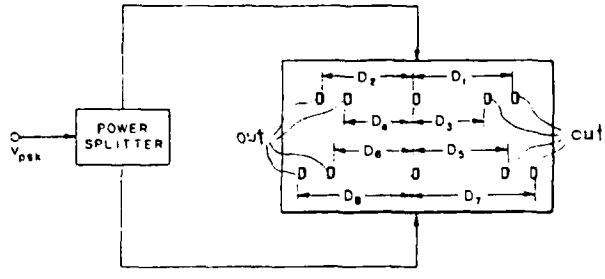



Fig. 5. Illustrates a bank of dual channel SAW differential delay lines, each set to a different Δt value so that BPSK signals of different chip rates can be detected.

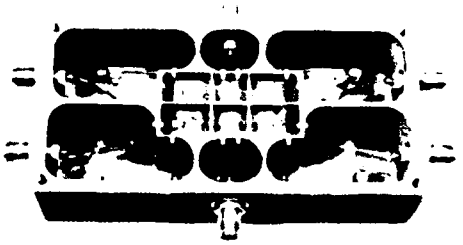


Fig. 6. A photograph of a two channel SAW differential delay line.

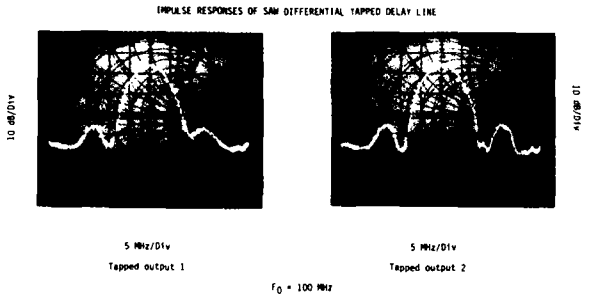



Fig. 7. Shows the impulse response for the Fig. 6 device.

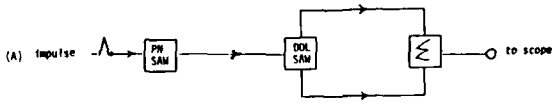



Fig. 8. Shows the simplified circuit used to both generate and detect a 10 MHz chip rate BPSK signal.

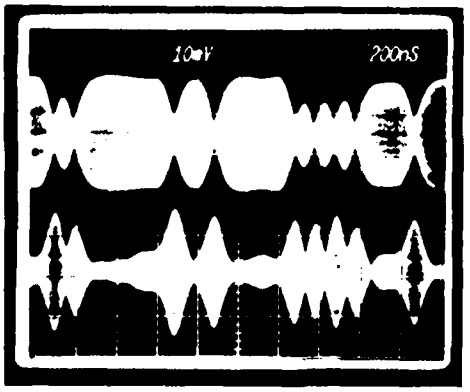



Fig. 9. Waveform generated (Top) and waveform detected (Bottom) using the circuit of Fig. 8.

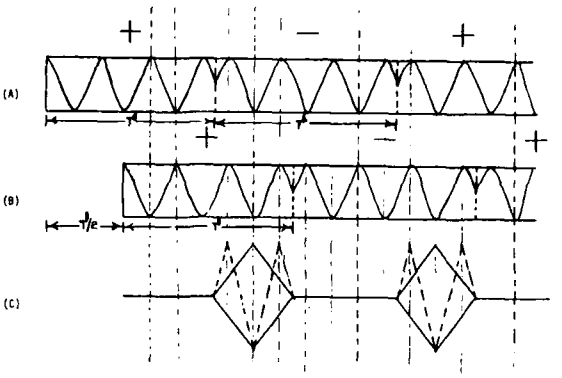



Fig. 10. Illustrates the concept of how the summer circuit of Fig. 8 operates.

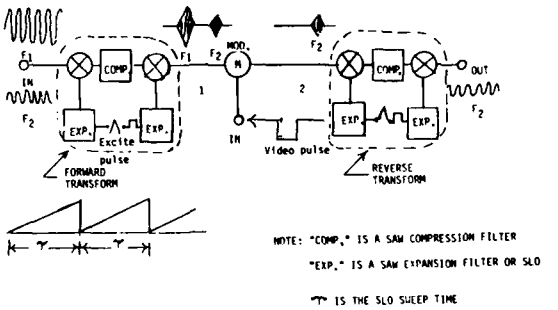



Fig. 11. A block diagram of a standard chirp-Z forward transform followed by a modulator (used for excision of signals) and, in turn, followed by a reverse transform.

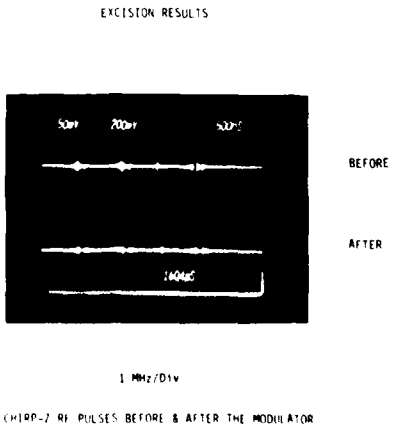



Fig. 12. A photograph showing excision results utilizing the modulator as located in Fig. 11.

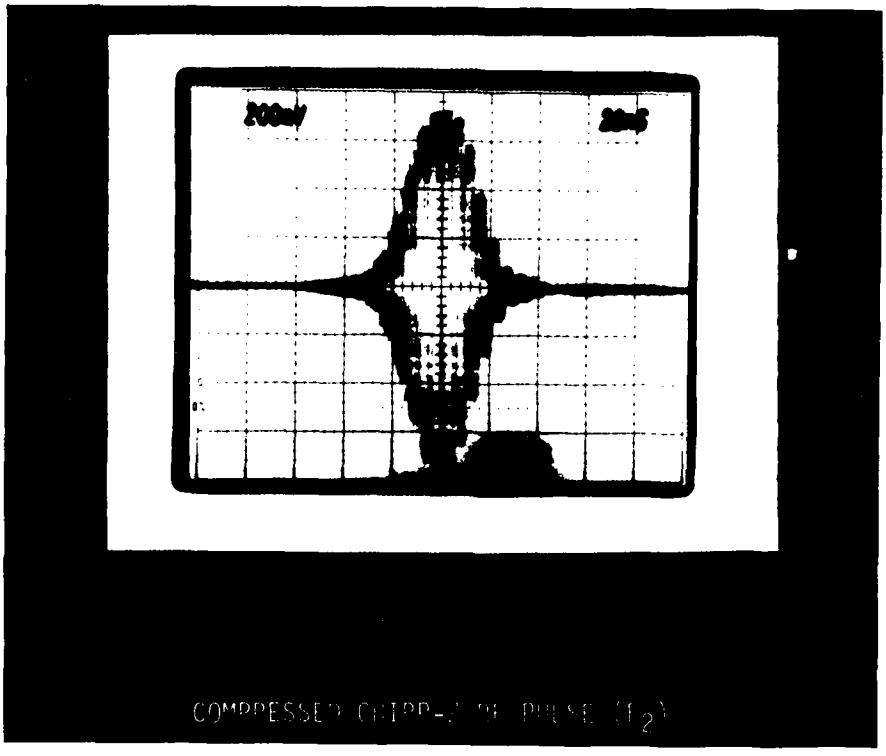


Fig. 13. Shows the RF cycles of the compressed pulse located after the modulator of Fig. 11.

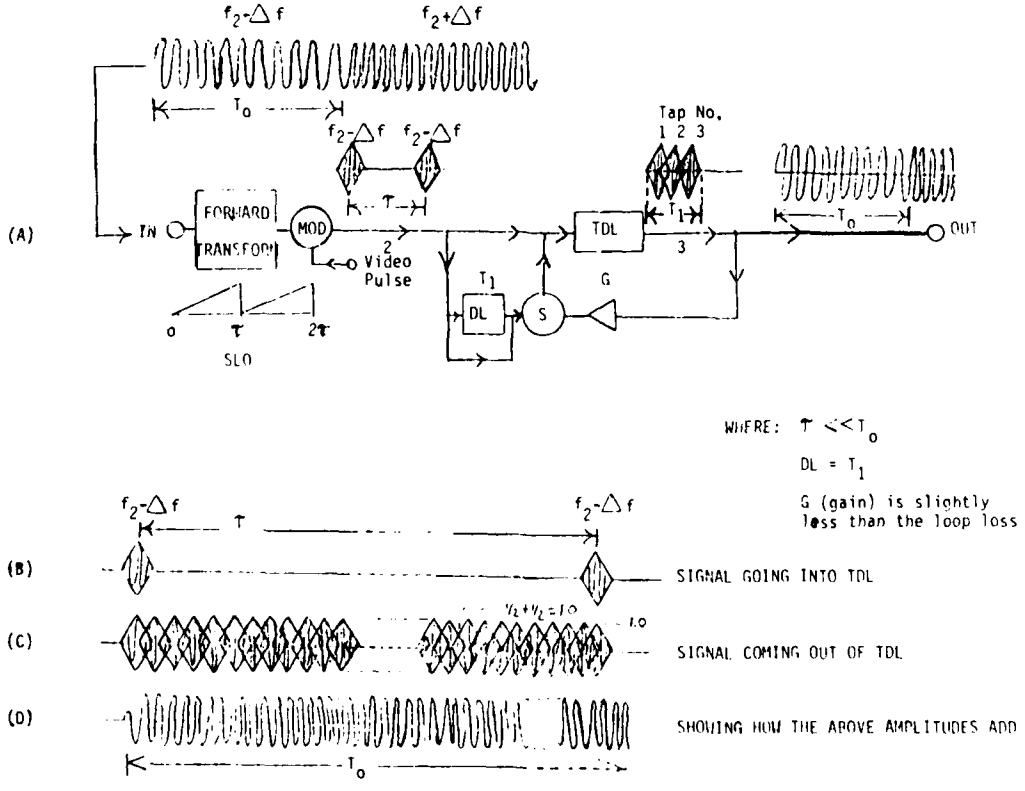



Fig. 14. Shows the tapped delay line (TDL) circuit (A) that is intended to replace the inverse transform shown in Fig. 11. (B), (C), & (D) illustrate how the TDL circuit operates.

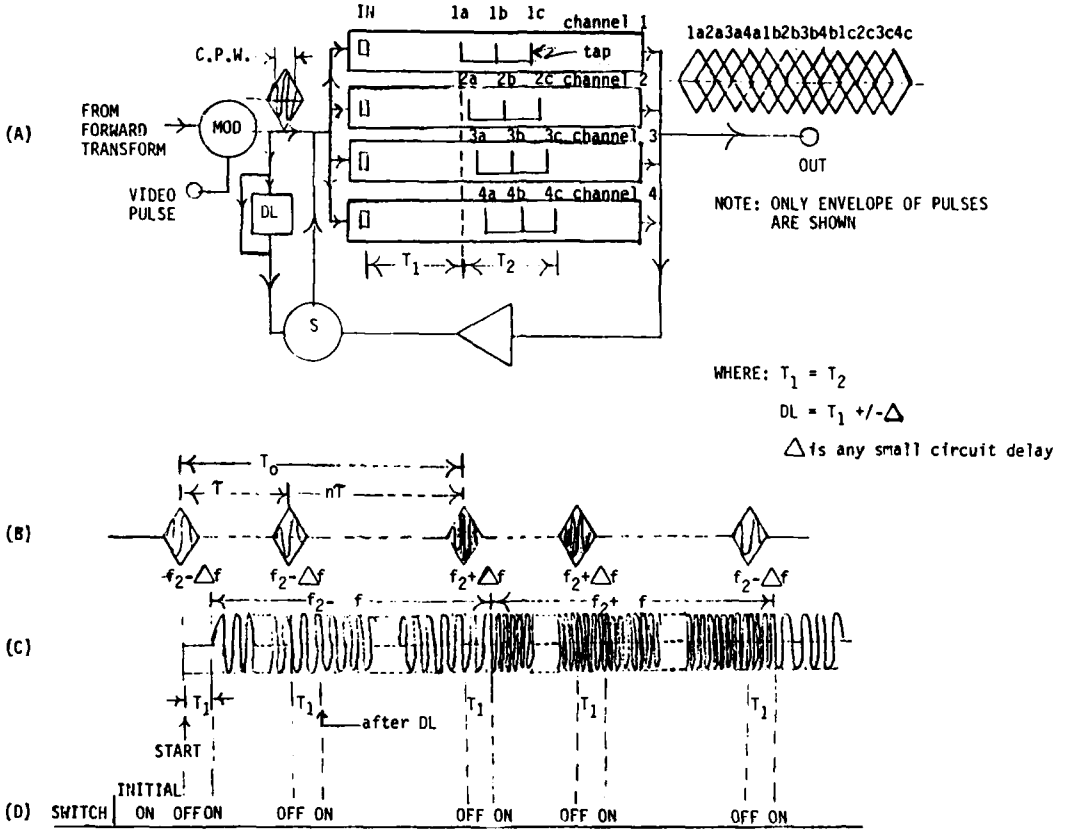
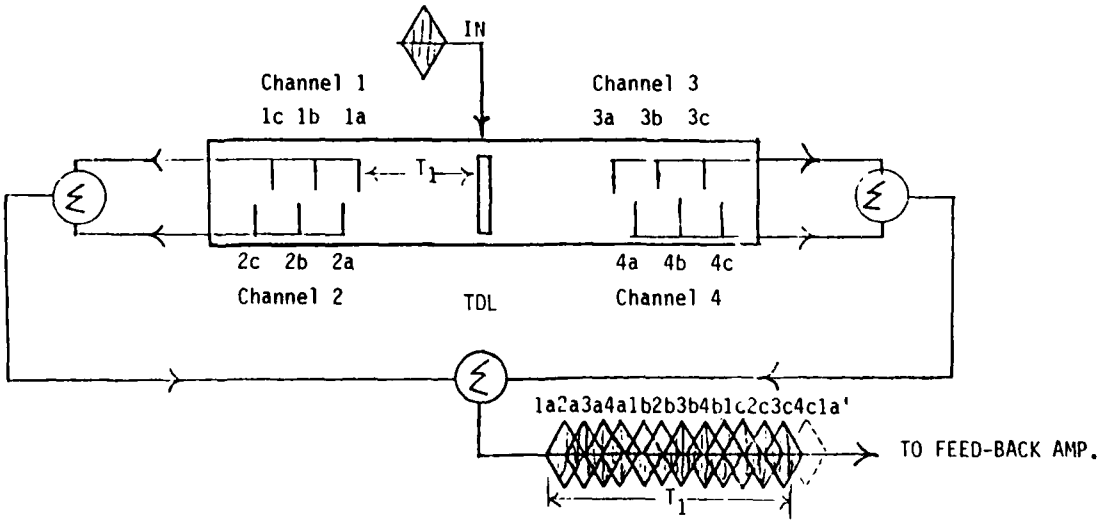



Fig. 15. (A) shows four TDL channels that can be utilized in Fig. 14(A) to ease the fabrication problem of a single TDL channel. (B), (C), & (D) illustrate operation of the Fig. (A) circuit.




NOTE: A GROUPING OF 12 PULSES COME OUT OF THE TDL FOR A SINGLE INPUT PULSE, $T_1$ (FIXED DELAY) IS CHOSEN SUCH THAT WHEN THIS GROUPING IS FEED BACK, EACH SUBSEQUENT GROUPING OVERLAPS THE PREVIOUS GROUP, SO AS TO APPEAR AS A CONTINUOUS CHAIN OF OVERLAPPED PULSES.

Fig. 16. Illustrates a more compact four channel TDL device.

42nd Annual Frequency Control Symposium - 1988

# DIELECTRIC RESONATORS

*Pierre GUILLON*

*U.F.R. des Sciences - I.R.C.O.M. - U.A. CNRS 356*

*123 Avenue Albert Thomas - 87060 LIMOGES Cédex FRANCE*

## SUMMARY

*Microwave passive and active devices containing dielectric resonators and design procedures used to characterize these components are reviewed. The applications of dielectric resonators in millimeter frequency wave band are also discussed.*

## INTRODUCTION

*A recent advance in miniaturization of microwave circuits has been the appearance of low loss temperature stable dielectric resonators [1]→[4]. These dielectric resonators are used to replace waveguide filters [5] in particular in satellite applications [6], [7], [8] where microstrip and strip line resonators cannot be used because of their high losses.*

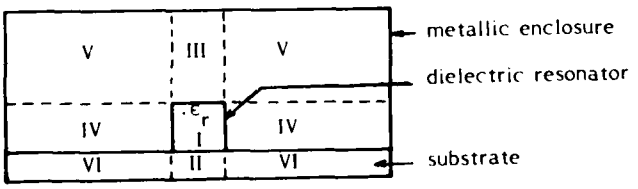
*Furthermore microwave oscillators and power combiners which use dielectric resonators [9] [10] can equal the characteristics of conventional microwave resonant cavities machined from invar.*

*Before we present the dielectric resonators filter's and oscillator's applications we review the theoretical electromagnetic and electric procedures available to design dielectric resonators parameters.*

## ELECTROMAGNETIC ANALYSIS

*The accurate mathematical description of the electromagnetic field in a dielectric resonator is very more complicated than the field description in a metallic waveguide resonator. It is often of great practical interest to approach the solution of the electromagnetic field in the dielectric resonator in some simplified ways. These simple models are useful to understand the principles of operating of dielectric resonators but they fail to provide the high accuracy usually required today in microwave circuit design. As a result more sophisticated rigourous techniques have been developed. They allow to compute both the resonant frequency, the field distribution and the Q factors of both TE, TM and hybrid modes to any desired accuracy.*

*We only considered circular cylindrical resonators since these are the shapes most often encountered in practice Also we give most attention to the analysis of shielded dielectric resonators i.e resonators placed in metallic enclosures as illustrated in figure 1.*




*- Figure 1 -*

*Simple models are capable to give results which are not too far from the exact values. The most popular of them is known as the Itoh and Rudoka's model [11] which consists to insure the continuity conditions of the EM tangential fields between regions (1) and (4) and between regions (1) and (2), (3) - the boundary conditions of EM fields on the metallic surfaces. Two trancendatal equations are then obtained and solved simultaneously to obtain resonant frequencies of dielectric resonator modes in particular of the dipolar $TE_{01\delta}$ one's.*

*An improvement of this model can be achieved by using the variational method [12] which permits to take into account the expansion of the EM field in regions (5) and (6).*

*Rigourous techniques [13] can be classified in four main categories : the mode matching method, the integral equations method, the perturbational and the numerical one's.*

*In the mode matching method [14] the structure (figure 1) is divided into different zones in which the field are presented in terms of series of wave guide modes. The electromagnetic parameters are then evaluated by searching the zeros of the determinant obtained when the fields of each zone are matched to satisfy the continuity conditions.*

*In the integral equation method [15] a surface integral equation formulation is used. It permits to obtain an homogeneous system of equations in which the unknown are the coefficients of expansion of E and H fields used to represent the equivalence surface currents.*

*The perturbational method [16] is based on the expansion of the field into an asymptotic series of inverse power of $\sqrt{\epsilon_r}$ and it is only exact in the limit of $\epsilon_r \to \infty$. But for values of $\epsilon_r$ commonly used in dielectric resonators only two or three terms of the asymptotic serie are acceptable.*

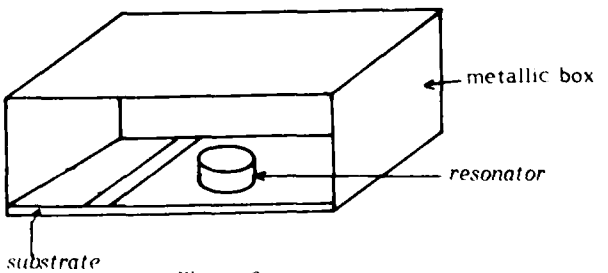
*To complete the review of rigourous techniques computation we add the numerical finite difference [17] and finite element method [18]. The last one is the most promising since it will permit in a future development to take into account any elements (transmission line, discontinuities...) inserted near the dielectric resonator.*

## ELECTRICAL ANALYSIS

*The use of dielectric resonators in microwave passive and active devices necessit the computation of the coupling between the resonator and an external circuit (microstrip line, waveguide...) and also the computation of the coupling between adjacent dielectric resonators.*

*The theoretical procedure usually used consists to establish a low frequency equivalent network of the system and then combining the electric and electromagnetic analysis to evaluate the corresponding coupling coefficients and scattering parameters as a function of all the other parameters (distances, impedances...) of the considered structure.*

*The simplest way to incorporate the dielectric resonator into a microwave network is placing it in top of a microstrip substrate as shown in figure 2.*




*- Figure 2 -*

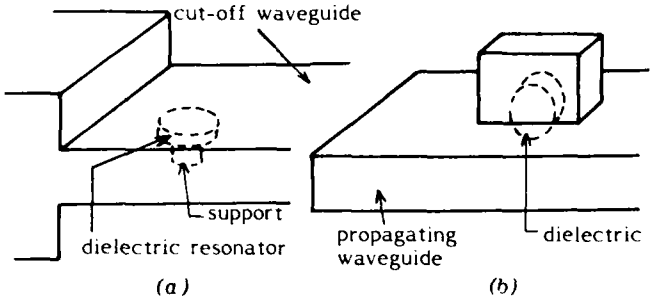
CH2588-2/88/0000-259 $1.00 C 1988 IEEE

The lateral distance between the resonator and the microstrip conductor determine the amount of coupling between the dipolar $TE_{01\delta}$ mode and the microstrip line.

For the excitation of the $TE_{01\delta}$ mode of a dielectric resonator by using a metallic waveguide two configurations are usually considered.

In the first one used for bandpass response [19] the resonator is housed in an evanescent waveguide, the end of which being connected to a waveguide above cut off (figure 3a).

In the second configuration proposed by Ren [21] and used for band stop response the dielectric resonator is isolated into its own box and coupled to the main waveguide through an aperture on the metallic wall (figure 3b).




- Figure 3 -

In these two cases the $TE_{01\delta}$ dielectric resonator mode is excited by the magnetic field of the dominant mode of the rectangular waveguide and the coupling is a function of the penetration of the resonator into the propagating waveguide.

Coupling between adjacent dielectric resonators is essential in the design of compact microwave filters. Two dielectric resonators placed close each other can be coupled directly by the evanescent fields respectively of the dielectric resonator and that of the mode excited in the metallic shield by the dielectric resonator [5]. An indirect coupling between two dielectric resonators can also be achieved by means of a transmission line [6]. In this case, the coupling coefficient is a function of the electric length of the line and of the distance between the line and the resonator. For satellite applications and to reduce weight and size of microwave filters the hybrid $HEM_{11\delta}$ mode which shows strong ressemblance with the $TE_{111}$ mode of the metallic circular cavity is usually used [3]. The coupling of the two orthogonal polarizations is achieved via a metallic screw [21] while the inter $HEM_{11\delta}$ dielectric resonators modes is provided by selective cross slots of different dimensions placed in the common wall of the two shields [7]. New realizations using dual mode dielectric resonators in a simple tubular enclosure have been recently described [8]. The new configurations do not require iris to achieve the coupling among the resonators, thus eliminating expansively machined parts with tight tolerances.

## FILTERS APPLICATIONS

There are filters types for almost every usages and dielectric resonators filters cover approximatively the band from 900 MHz to 35 GHz.

The dielectric filters design is similar to the conventional cavities and transmission line filters designs. The coupling coefficients are computed accurately from the filter specifications and are then related to dielectric resonator data to obtain physical dimensions of the filter.

Butterworth, Tchebyscheff band pass and band stop filters are defined as a section of evanescent or propagating waveguides in which the dielectric resonators are inserted. Transverse and longitudinal orientations can be considered in rectangular or circular waveguides : the excitation of the first (last) dielectric resonator can be provided by a line, a loop, or a waveguide.

Elliptic band pass filters are potentially useful for satellite applications in particular in input multiplexers where losses are tolerable. The filter configurations given are of the canonical form and dielectric resonators elliptic filter can be realized using $TE_{01\delta}$ [13] or $HEM_{11\delta}$ modes [7] [8]. At about 900 MHz the physical dimensions for filters using conventional dielectric resonator modes become too large for radio link communications. Some authors [22] [23] [24] have developped a metallized dielectric TEM resonator for miniaturization, but its intrinsic Q factor is degradate.

The insertion losses of dielectric resonator filters are greater than the same conventional waveguides filters. They depend on the number of resonators and on the surrounding wall positions, the dielectric fixtures and the metallic tuning screws which are generally close to the resonators and which induce additional losses.

One of the major problem in dielectric resonator filters is caused by the presence of spurious responses which interfere with the filter's performances. These spurious modes can be those of dielectric resonators, those of the metallic shield excited by the resonator or by the input/output excitations. Some solutions have been proposed to eliminate them which consist to place suppressing elements between adjacent resonators or to select appropriate choice of resonators and shielded dimensions [25].

## OSCILLATORS APPLICATIONS

Microwave dielectric resonators oscillators form an important part of all microwave systems. The emphasis has been focused in low noise, small size, low cost and high temperature stability. Dielectric resonators have been first used with diode oscillators [26] and combiners and now transistors dielectric resonators present an interesting solution as a quality oscillator for fixed frequency and narrow band tunable oscillators [10] [13].

To realize a stable MIC source, dielectric resonators can be used in two different ways : - as a passive stabilization element coupled to a free running oscillator or as a circuit element to determine the oscillation frequency. In these circuits dielectric resonators can be used as a series of parallel feedback elements, the pushing, pulling stabilization and power being controlled by the coupling between the resonator and the microstrip line [27]. We can also note that cylindrical dielectric resonators have been recently used in two new applications, one concerning a frequency discriminator [28] and the other a fast switching x and Ku Band multi frequency oscillator using a single Ga As FET [29]. Temperature stabilization can be achieved by using a high stable temperature material. In some cases digital compensation technique will permit to improve this temperature stabilization.

The frequency tuning can be accomplished mechanically, electrically or optically. This tuning is almost very low and is smaller than 1 %.

## MILLIMETER WAVELENGTH DEVICES APPLICATIONS

Improvements in ceramic materials of high permittivity and small losses will permit to dielectric resonators to become major components in communication and radar sys-

*tems. Since dielectric resonators have no physical limitations their use in millimeter structures promises to extend their utilization beyond 100 GHz. At this frequency the size of cylindrical resonator is very impractivally small. Spherical resonators and whispering gallery modes of cylindrical resonators are attractive solutions for solving this problem. They have already been used in some millimeter wave length devices in particular in a diode oscillator stabilized by a spherical resonator [38] and in a filter [31] and an oscillator [32] using whispering gallery modes of a cylindrical dielectric resonator*

## CONCLUSION

*Dielectric resonators are now commonly used in many microwave active and passive devices acting form about 850 MHz to 100 GHz.*

*They will be also main parts in some future superconducting circuits where they could be used in two different ways : - to characterize superconducting materials - to permit the realization of high microwave performances superconducting resonators.*

## REFERENCES


[1] **J.K. PLOURDE et Al**
$Ba_2$ $Tl_9$ $O_{20}$ as a microwave dielectric resonator
Journ. Amer. Ceramic Society, vol.58, pp.418-420, sept-oct. 1975

[2] **W. SIMONET, J.C. MAGE**
(Zr) SN TI $O_4$ dielectric materials with asjustable temperature coefficient and low microwave loss
Proceed of the Am. Ceramic Society meeting - 1980 - CHICAGO

[3] Microwave dielectric technical briefs, n° 831, Trans. Tech. Inc. Adamotown MD (incluted)

[4] RESOMICS - Microwave ceramics - Design manual n°58-04 MURATA ERIE - MARIETTA - Georgia (undated)

[5] **S.B. COHN**
Microwave bandpass filters containing high Q dielectric resonators
IEEE Trans. Microwave theory and techniques, vol. MTT-16, pp. 218-222, April 1968

[6] **A.E. ATIA, R. BONETTI**
Generalized dielectric resonator filters. Comsat technical review, vol. 11, pp. 321-343, Full 1981

[7] **S.J. FIEDZIUSKO**
Dual mode dielectric resonator loaded cavity filters
IEEE Trans. on microwave theory and techniques, vol. MTT-30, pp. 1311-1314, september 1982

[8] **K.A. ZAKI, C. CHEN, A.E. ATIA**
Dual mode dielectric resonator filters without iris
Proceedings of the International Microwave Symposium MTTS, p,.141, June 1987, Las Vegas

[9] **N.K. AGARWAL**
Dielectric resonator oscillators using GA AS/GA AL AS heterojonction bipolar transistors
IEEE MTTS International microwave symposium digest, pp. 95-98, Baltimore June 1986

[10] **H. ABE et al**
A high stabilized low noise GA AS FET integrated oscillator with a dielectric resonator in the C band
IEEE Trans. microwave theory techniques, vol. MTT-26, pp. 156-162, march 1978

[11] **T. ITOH, R.S. RUDOKAS**
New method for computing the resonant frequencies of dielectric resonators
IEEE Trans. Microwave Theory and Techniques, vol. MTT-25, pp. 52-54, January 1977

[12] **L. BERMUDEZ, P. GUILLON**
Applications of variational principle for calculation of resonant frequencies of cylindrical dielectric resonators
Electronic Letters, vol. 22, n°1, pp.31-33, 2 January 1986

[13] **D. KAJFEZ, P. GUILLON**
Dielectric resonators
Artech House, Inc ISBN 0 89006-201-3, 1986

[14] **K.A. ZAKI, C. CHEN**
New results in dielectric loaded resonators
IEEE Trans. Microwave Theory Techniques, vol. MTT-34, pp. 815-824, July 1986

[15] **A.W. GLISSON, D. KAJFEZ, J. JAMES**
Evaluation of modes in dielectric resonators using a surface integral equation formulation
IEEE Trans. Microwave Theory Techniques, vol. MTT-31, pp. 1023-1029, june 1984

[16] **J. VAN BLADEL**
On the resonances of a dielectric resonator of very high permittivity
IEEE Trans. Microwave Theory and Techniques, vol. MTT-23, pp. 199-208, February 1975

[17] **P. GUILLON, D. ROUSSET, Y. GARAULT**
A new determination for the resonant frequencies and field patterns of dielectric resonators

[18] **P.S. KODI, M.S. LEONG**
Finite element analysis of the shielded cylindrical dielectric resonator
Proc. IEEE part H, vol. 132, pp. 7-16, february 1985

[19] **Y. KONISHI**
External Q of a $TE_{01\delta}$ dielectric resonator in a $TE_{1,0}$ waveguide
Electronics and communications in JAPAN, vol. E-59, pp. 143-146, january 1976

[20] **C. LI REN**
Waveguides bandstop filter utilizing $Ba_2$ $Tl_9$ 020 dielectric resonators
IEEE MTT-S International microwave symposium digest, pp. 170-173, june 1978

[21] **J.P. ASTIER, P. GUILLON**
Elliptic microwave filter using dual modes of dielectric resonators
European microwave conference digest, pp. 335-337, PARIS, sept. 1985

[22] **K. WAKINO**
Quarter wave dielectric transmission line diplexer for band mobile communications
IEEE MTT-S International microwave symposium digest, pp. 278-281, june 1979

[23] **E. KAMPPINEN and al**
Realization of integrated miniature ceramic filters for 800 MHz cellular mobile radio applications
Proceedings of the European Microwave Conference, p. 175, Rome, 1987

[24] **T. NISHIKAWA et al**
800 MHZ high power bandpass filter using $TM_{110}$ mode dielectric resonators for cellular base stations
Proceedings of the International Microwave Sumposium MTT-S. p. 519, june 1988, New-York

[25] L. REN
*Mode suppresor for dielectric resonators filters*
*IEEE MTT-S International Microwave Symposium digest, pp. 389-391, june 1982*

[26] W.R. DAY
*MIC diode oscillator stabilized by a dielectric resonator*
*Proced 3 rd biennal Cornell : Electrical engeneering conference, august 1971*

[27] A. PODCAMENI, L.F.M. CONRADO
*Design of microwave oscillators and filters using transmission mode*
*IEEE Trans. Microwave Theory and Techniques, vol. MTT-13, pp. 1325-1333, december 1985*

[28] S.J. FIEDZIUSKO
*High φ dielectric resonator frequency discriminator*
*Proceedings of the International Microwave Symposium MTTS, p. 175, june 1987, Las Vegas*

[29] A.P.S. KHANNA, R. SOO HOO
*Fast switching α and Ku Band multi frequency dielectric resonators oscillator using a single GA AS FET*
*Proceedings of the International Microwave Symposium MTTS, p. 189, june 1987, Las Vegas*

[30] N. MAMODALY, A. BERT, J. OBREGON, F. FARZANEH, P. GUILLON
*A fondamental mode INP Gunn dielectric resonator oscillator at 94 GHz*
*European microwave symposium digest, pp. 170-177, PARIS, sept. 1985*

[31] X.H. JIAO, P. GUILLON, L.A. BERMUDES, P. AUXEMERY
*Whispering gallery modes of dielectric structures : applications to millimeter wave bandstop filters*
*IEEE Trans. on microwave theory and techniques, vol. MTT-35, n°12, décembre 1987, p. 1169*

[32] L.A. BERMUDES, P. GUILLON, A. BERT, J. OBREGON
*A 94 GHz low noise GA AS MESFET oscillator using whispering gallery dielectric resonator modes and a new push push configuration reducing 1/f converter*
*Proceedings of the International microwave symposium, june 1988, New-York*

42nd Annual Frequency Control Symposium - 1988

# TEMPERATURE SENSITIVITY OF DIELECTRIC RESONATORS AND DIELECTRIC RESONATOR OSCILLATORS

M.J. LOBODA, T.E. PARKER and G.K. MONTRESS
Raytheon Research Division
131 Spring Street
Lexington, MA 02173

## *ABSTRACT*

The specification of an oscillator's frequency stability over a temperature range is a common requirement for many high-stability oscillator applications. While the mechanisms which contribute to the frequency stability over temperature have been extensively studied for quartz based frequency sources such as surface acoustic wave (SAW) and bulk acoustic wave (BAW) oscillators, little information exists in regard to understanding and estimating the frequency vs. temperature characteristics of dielectric resonator based oscillators (DRO). Experiments were performed on L-Band and S-Band DRs and DROs to investigate the details of their behavior over temperature. Measurements were performed using commercially available Zr/Sn titanate and barium tetratitanate dielectric resonator materials enclosed in metal cavities. The results indicate that the "temperature coefficient of the resonant frequency, $\tau_f$" supplied by vendors of the high dielectric constant ceramics is not sufficient to accurately predict a DRO's total frequency excursion over a given temperature range. Significant, second order contributions to the frequency vs. temperature characteristics were observed, with values ranging from -27 to +20 ppb/°C$^2$. Also the cavity material used to enclose the DR can significantly alter the DR's temperature characteristic. Procedures have been developed to aid in the estimation of a DRO's characteristic temperature sensitivity, allowing one to adjust the DRO design to minimize the total frequency excursion over a desired temperature range.

## *INTRODUCTION*

Of the many requirements imposed on the high-stability oscillators used in today's military radar and communication systems, one usually involves the specification of the oscillator's frequency stability over a given temperature range. For quartz based frequency sources such as surface acoustic wave (SAW) and bulk acoustic wave (BAW) oscillators, the mechanisms which contribute to their frequency stability over temperature are well documented, and accurate techniques for predicting the frequency stability over temperature are available (see [1],[2] and their references). However, while the dielectric resonator oscillator (DRO) has been shown to be a low-noise signal source [3],[4], little information exists in regard to understanding and accurately predicting its frequency vs. temperature characteristic. Data sheets provided by the vendors of the high dielectric constant, ceramic materials suggest that it is relatively easy to predict and control a DRO's frequency deviation over temperature, by selecting an appropriate DR from the large range of linear sensitivities provided. The sensitivity is described by the vendors as a "temperature coefficient of the resonant frequency", and is expressed in parts per million per degree centigrade. This single parameter implies a linear behavior in the change of the resonant frequency of the dielectric resonator puck over temperature. Detailed experiments have been performed to investigate the precise temperature behavior of these materials when assembled into cavity resonators and also as free running oscillators. The motivation behind these experiments was to develop a technique for predicting the frequency versus temperature characteristics of DROs which would be similar to that already established for quartz based frequency sources. This information would be of use when evaluating DROs for applications currently addressed by SAW and BAW based microwave sources where it might be more advantageous to use a higher fundamental frequency, low-noise signal source. The experiments were performed on L-Band and S-Band DRs and DROs constructed using commercially available Zr/Sn titanate and barium tetratitanate dielectric resonator materials enclosed in metal cavities. The results indicate that a non-linear frequency vs. temperature characteristic exists for these resonator materials, and that the magnitude of the non-linearity can be comparable to that measured for quartz resonators. Consequently, the temperature coefficient data supplied by the vendors of the high dielectric constant ceramics will not provide sufficient information to accurately predict a DRO's total frequency excursion over a given temperature range. Also, other essential components such as the metal cavity and the oscillator electronics will contribute additional linear terms to the overall DRO frequency vs. temperature characteristic.

To determine to what extent each component of a DRO will contribute to the overall temperature sensitivity, the oscillator can be divided into two parts. The first part includes those components which make up the dielectric resonator assembly (DRA), i.e., the high dielectric constant ceramic puck, the puck support material, and the cavity (enclosure) material. The second part includes the contributions made by the oscillator electronics, e.g., bipolar transistor or FET amplifiers, power dividers, couplers, etc.

Within the DRA, a significant factor inducing frequency change is thermal expansion of the assembly components. As the DR puck changes size, the resonant frequency of the DRA changes. Also, as the proximity between the cavity walls and the puck changes, the frequency will change due to perturbations of the electromagnetic field. Thus the thermal expansion coefficients of the materials comprising the DRA will play a significant role in determining the frequency vs. temperature characteristic. At the same time, the

CH2588-2/88/0000-263 $1.00 C 1988 IEEE

dielectric constant of the ceramic puck changes as a function of temperature, also inducing a resonant frequency shift. Therefore, the temperature dependence of the resonant frequency in a DRA is the result of complicated physical interactions of the material properties of its components. This is much different than that found for a quartz resonator, where it is the cut angle of the quartz itself which essentially determines the temperature characteristic.

As for the oscillator electronics, the most significant effect causing a frequency shift would be temperature dependent phase shifts thru the electronic components, which alter the phase around the oscillator loop, and thus the oscillator's frequency. The results presented in this paper indicate that for a two-port resonator, used in a basic feedback oscillator design, the oscillator electronics can play a minor role in determining the DRO's temperature characteristic.

On a vendors' specification sheet describing the temperature sensitivity of a dielectric resonator puck, one finds only a single numerical factor describing the frequency change. Information concerning higher order temperature coefficients, the cavity and support materials used during the characterization, or a temperature range over which the numbers are valid, is not included in the catalogs. When inquiries were made as to how these measurements are performed, it was found that the techniques differed between material vendors, indicating that there is not yet a standard definition applied to this parameter associated with a DR puck. One vendor performs the measurements using cavity and support materials with the same thermal expansion coefficient as the dielectric puck material, with the inside of the cavity plated with a high conductivity metal. This will allow the ratio of the dimensions of the cavity and puck to remain constant over temperature [5]. Another vendor uses an oversized, gold plated aluminum cavity and a quartz support. Both techniques attempt to minimize the influence of the cavity walls on the temperature sensitivity of the DRA. Two or three measurements of the DRA resonant frequency are made at different temperatures. The temperature sensitivity coefficient is calculated from the fractional frequency change which occurred over the temperature range defined by the highest and lowest temperature measurement points. The **spacing** between these points ranged from 30 to 85 °C depending on the vendor. Thus the coefficient provided is simply a linear term (given as a fractional frequency change per degree centigrade) averaged over a relatively large temperature range. But more importantly, if this characterized puck is used in a cavity which is in any way different from the one used to measure the vendor's coefficient, the number provided will not be applicable to this new DRA configuration.

## *EXPERIMENTS*

The standard DRA used in this investigation of temperature sensitivity was constructed from a brass cavity whose inner diameter was 1.6 times the diameter of the DR, and whose height was 4.5 times the height of the DR. A two-port design was chosen for its ease in separating the testing of the resonator from the associated oscillator electronics. Energy was coupled in and out of the cavity using straight, open-circuit probes, which were placed at the bottom of the cavity. The resonators were designed to operate in the $TE_{01\delta}$ mode. The puck and its fused quartz support were secured by a nylon screw or by low-loss epoxy, while the probes were secured at their tips by a low-loss, non-conductive epoxy. Unless otherwise specified, the DR was centered within the cavity. A basic feedback loop configuration was used for the oscillators which were designed within a 50 ohm characteristic impedance environment. Commercial wideband amplifiers and power dividers were used for the oscillator electronics [4].

A brief description of the response of an oscillator to a change in ambient temperature will help to introduce some of the problems involved with this aspect of DRO characterization. Figure 1 illustrates the fractional frequency change for a 1.0 GHz barium tetratitanate DRO in response to a temperature step. The shape and physical description of the response is common to most oscillators

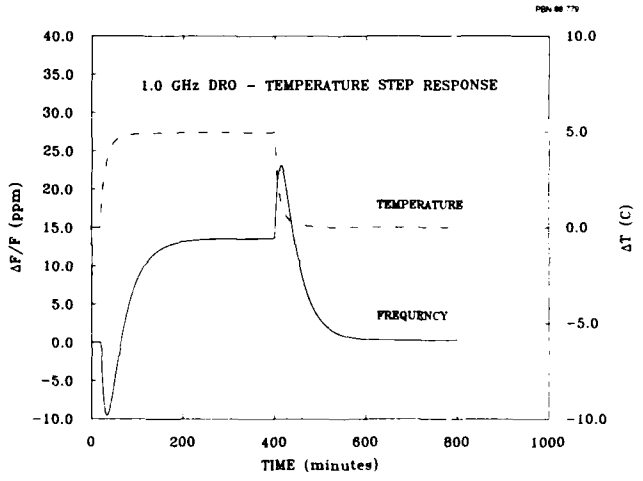



Figure 1. Fractional frequency change of a free-running DRO in response to a +5 °C temperature step, illustrating both the dynamic and static effects.

and can be separated into two parts. First, at the onset of the step there is a frequency transient, indicative of a dynamic response which is proportional to the time rate of change of temperature [4]. This is followed by a steady-state frequency shift whose magnitude is proportional to the amount of the temperature change. This frequency change can be described in terms of a static temperature coefficient, usually expressed in ppm/°C. While it is the latter part of the response for which this work (and almost all work) is concerned, it is important to notice the large magnitude and duration of the dynamic response, as well as a large thermal time constant for these oscillators. The thermal response of a DRO makes the accurate measurement of its static temperature characteristic a very slow process, in order to allow the entire oscillator to reach thermal equilibrium after each temperature change [4]. The temperature characteristic curves shown later in this paper were made over a 120 °C temperature range and typically took 4-5 days to complete!

Initial frequency vs. temperature measurements were performed on a 1.0 GHz DRO design which used a barium tetratitanate DR manufactured by Trans-Tech, Inc. (part number D-8512). The oscillator electronics consisted of a Watkins-Johnson A19-1 silicon bipolar transistor amplifier, and an Anzac DS-313 power divider. The static temperature coefficient was measured from the response to a 5°C temperature step. The experiment was repeated at various points over the temperature range -40 to 85 °C, resulting in a curve of static temperature coefficients as a function of temperature, and this data is shown in Fig. 2. The barium tetratitanate DR was listed by the vendor as having a temperature coefficient of 4±0.5 ppm/°C, and this value was only observed at the high end of the temperature range for the measurement.

To determine the contribution to the oscillator's temperature characteristic due to the loop electronics, a phase locked loop (PLL) temperature measurement system was developed so that the temperature sensitivity of the DRA alone could be measured. The PLL is shown as part of Fig. 3. The measurements of the static temperature coefficients were repeated on the same resonator used in the oscillator, and the results are also shown in Fig. 2. This data indicates that the oscillator electronics made a negligible contribution to the DRO's temperature sensitivity characteristic.

Noting that the curve representing the static temperature coefficient versus temperature is the derivative of the fractional frequency change versus temperature, the results of these measurements indicate that a parabolic frequency vs. temperature characteristic exists for this oscillator. A **positive** second order coefficient somewhat smaller in magnitude and opposite in sign than commonly found for ST cut quartz SAW resonators was observed. From Fig. 2, a value of approximately +20 ppb/°C$^2$ was calculated.

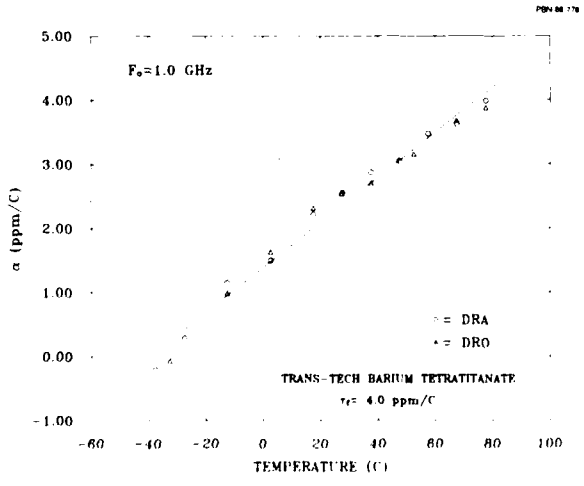



Figure 2. The static temperature coefficient, α, versus temperature, measured on a 1.0 GHz DRA constructed with a barium tetratitanate DR (h=2.3 cm, r=2.85 cm), fused quartz support, brass cavity, and probe coupling. Also shown is the data for the same measurement performed on the DRO configured using this DRA.

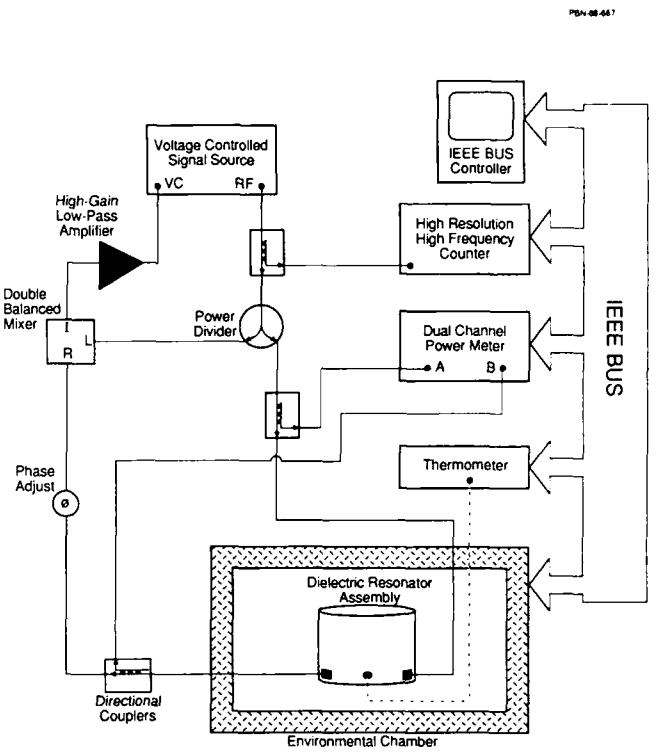



Figure 3. Block diagram of the automated, phase locked loop system for the measurement of a dielectric resonator assembly's frequency vs. temperature characteristic.

Knowledge of an oscillator's characteristic frequency excursion over temperature is best obtained using data collected from **direct** measurements of total frequency change versus temperature. An automated measurement system was designed for this task. This system controls the cycling of the temperature in an environmental chamber over a specified range, stepping the temperature in given increments, and at the same time records the frequency and insertion loss (or oscillator power) of the device under test. With the ability to continuously monitor the device parameters as a function of temperature, over the nominal five days which is required for a complete temperature cycle, we could now construct a detailed table of the data for later analysis. A schematic of this system is shown in Fig. 3. For the data presented in this work, the standard measurement using this system was to cycle the temperature from 25 °C up to 85 °C, down to -40 °C and back to 25 °C. Every two hours a data point was recorded, and then the temperature was changed five degrees, resulting in approximately 50 points per run.

A large set of measurements were performed on a number of 2.0 GHz DRAs using the automated system and the phased lock loop. All of the DRAs were constructed using the same brass cavity and fused quartz support, and all of the DRs were made from Zr/Sn titanate ceramic materials. The DRs were purchased from two different vendors and with various vendor specified "temperature coefficients, $\tau_f$", ranging from 0 to +5.2 ppm/°C.

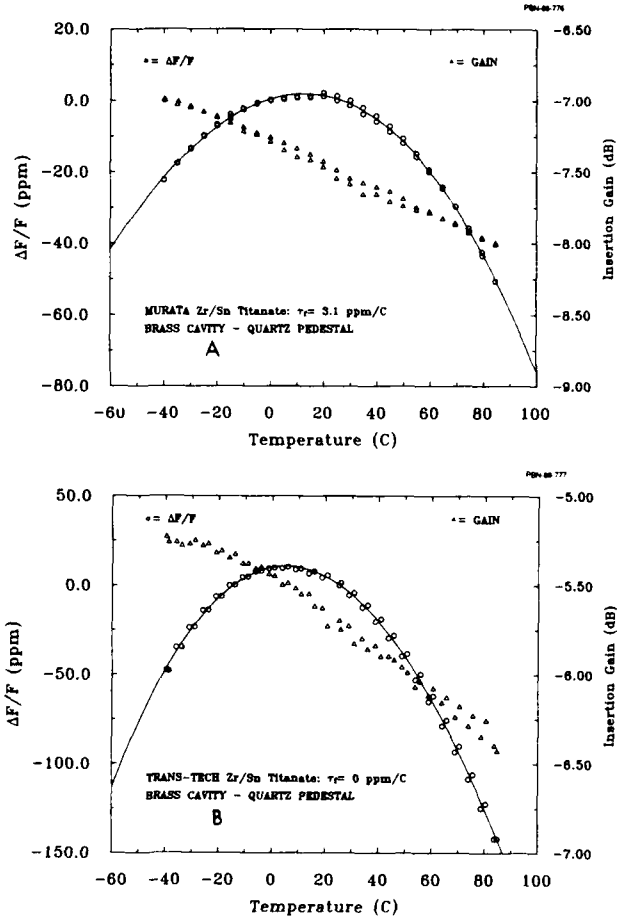



Figure 4. Fractional frequency change and insertion gain vs. temperature measured on Zr/Sn titanate based DRAs. This shows the parabolic temperature dependence, as well as the significant difference in total frequency excursion between a Murata DR (4a) and a Trans-Tech DR (4b).

Figure 4a shows the characteristic fractional frequency and insertion loss change versus temperature for a DRA built with a Murata/Erie Series "U" DR having $\tau_f$=3.1 ppm/°C, a radius of 1.33 cm, height of 1.15 cm and a center hole of 0.6 cm diameter. Figure 4b shows a similar measurement on a DRA using a Trans-Tech DR with $\tau_f$=0 ppm/°C, a radius of 1.43 cm, height of 1.14 cm, and a center hole of 0.3 cm diameter. $\tau_f$ for a DR is determined by Murata from the measured change in frequency between 25 and 55 °C. Using this same procedure on the data in Fig. 4a we calculate a value of -0.52 ppm/°C rather than 3.1 ppm/°C. Trans-Tech determines its value of $\tau_f$ in a similar fashion but uses the frequency change between -25 and 60 °C. For the data in Fig. 4b we find a value of -0.56 ppm/°C, rather than their nominal value of 0±.5 ppm/°C. (We believe that these are the procedures used by these two DR material vendors to calculate the values of $\tau_f$, but it has been difficult to obtain accurate information.) Notice the difference in the range of the fractional frequency axis for the two plots. Both of these DRAs exhibit parabolic frequency versus temperature characteristics, with computed second order polynomial coefficients for the Murata and Trans-Tech resonators of approximately -8.8 ppb/°C$^2$ and -27 ppb/°C$^2$, respectively. For comparison, a SAW resonator fabricated on ST quartz exhibits a second order coefficient of nominally -35 ppb/°C$^2$. It is important to note that the two "Zr/Sn titanate" DRs from the two different vendors exhibit significantly different second order temperature coefficients. It is also worth noting that in both measurements the insertion loss varied approximately 1 dB over the temperature measurement range, which is comparable to what we observe in similar measurements made on quartz SAW resonators.

Measurements of the same type described above were performed on eight other Murata DRs with $\tau_f$'s of 1.4, 3.1, and 5.2 ppm/°C. Figure 5 compares the typical frequency versus temperature characteristics for the three different $\tau_f$'s evaluated in our cavity. All of the Murata Zr/Sn titanate DRs measured exhibited the same type of parabolic temperature sensitivity, with an average second order polynomial coefficient of approximately -9 ppb/°C$^2$. The dominant effect of the different values of $\tau_f$ is simply to shift the peaks of the parabolas from low to high temperatures for increasing values of $\tau_f$.

In analogy with quartz SAW resonators, it is useful to define the "turnover temperature" of the fractional frequency change versus temperature curve as the temperature where the slope of the curve equals zero. This concept can be very useful when comparing the differences between temperature characteristics. In terms of the family of curves for the various $\tau_f$'s just discussed, it can then be shown that each value of $\tau_f$ corresponds to a particular turnover temperature. Mathematically, the frequency versus temperature characteristic can be described by the relation:

$$\Delta F/F(ppm) \approx C_0 + C_1T + C_2T^2 + C_3T^3... \qquad (1)$$

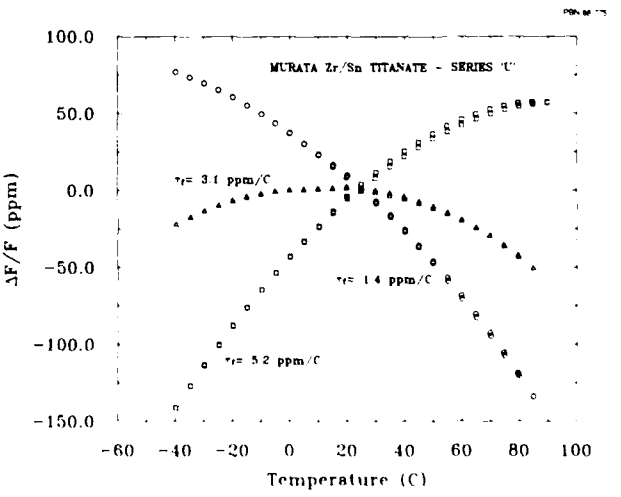



Figure 5. Frequency vs. temperature characteristics as a function of $\tau_f$. These measurements were performed on 2.0 GHz Zr/Sn titanate based DRAs, constructed using Murata Series "U" DR materials, a brass cavity, fused quartz support, and probe coupling.

where the $C_i$'s are the coefficients of a polynomial fit performed on the measured data and T is the temperature in degrees centigrade. The second order temperature coefficient $C_2$ has already been discussed. It was found to have different values depending on the type and source of the DR material, specifically negative values for the Zr/Sn titanates (nominally -9 ppb/°C$^2$ for the Murata material and -27 ppb/°C$^2$ for the Trans-Tech material) and a positive value for barium tetratitanate. If the third and higher order terms are neglected, a simple approximate expression for the turnover temperature, $T_0$, can be obtained by taking the derivative of Eq. 1 with respect to T and determining the value of T for which the derivative goes to zero.

$$T_0 \approx -C_1/2C_2 \qquad (2)$$

For the data in Figs. 4 and 5, third and higher order terms can be safely neglected. In Fig. 5, the nominal values of the calculated DRA turnover temperatures are -85, 12, and 105 °C for the vendor temperature coefficients 1.4, 3.1, and 5.2 ppm/°C, respectively. Table 1 summarizes the coefficients and turnover temperatures for the various resonators we examined from Murata. As related to Fig. 2, the turnover temperature corresponds to the value of T where the static temperature coefficient equals zero. For the barium tetratitanate resonator measured, the turnover temperature was nominally -40 °C.

From the data shown in Fig. 5, it is evident that the different values of $\tau_f$ for a DR material can be modelled by appropriately modifying the value of the linear coefficient, $C_1$. By writing $C_1$ in terms of $\tau_f$ as

$$C_1 \approx \tau_f + \eta_E \qquad (3)$$

the new term $\eta_E$ represents any additional linear shifts in the frequency versus temperature characteristic due to those components external to the DR, such as the cavity, support, oscillator electronics, etc. Substituting Eq. 3 into the expressions for turnover temperature (Eq. 2) and fractional frequency versus temperature (Eq. 1), these equations can be written in terms of the properties of the components comprising the DRO:

$$T_0 \approx -(\tau_f + \eta_E)/2C_2 \qquad (4)$$

$$\Delta F/F(\text{ppm}) \approx C_0 + (\tau_f + \eta_E)T + C_2T^2 \qquad (5)$$

By performing Murata's calculation for $\tau_f$ on each of the curves in Fig. 5, we found that the addition of those external components necessary to construct the DRA caused a downward shift in the value of $\tau_f$. Based on the values for $C_1$ found in Table 1, this shift in $\tau_f$ corresponds to an average value for $\eta_E$ of -2.90±0.22 ppm/°C. It is now possible to write an approximate expression for the total frequency excursion of our DRA in terms of the Murata definition of $\tau_f$, and the properties of their DR material:

MURATA:

$$\Delta F/F(\text{ppm}) \approx C_0 + (\tau_f - 2.9)T - 0.009T^2 \qquad (6)$$

where $C_0$ is an arbitrary constant related to the reference frequency used to define the fractional frequency change. A similar expression can be derived for the Trans-Tech Zr/Sn titanate material in our DRA design using their **different** definition of $\tau_f$:

TRANS-TECH:

$$\Delta F/F(\text{ppm}) \approx C_0 + (\tau_f + 0.29)T - 0.027T^2 \qquad (7)$$

These approximate expressions can be used to estimate the fractional frequency change for a DR placed in a brass cavity of our design on a quartz support. They are approximate in the sense that they do not attempt to account for factors such as the limited accuracy for the values of $\tau_f$, or the small dependence of $C_2$ on the value of $\tau_f$, as evident in Table 1.

The reproducibility of temperature characteristics between DRs made in the same and/or different batches is very important in order to be able to consistently estimate the sensitivity of a particular DRA design. In the set of eight Murata Zr/Sn titanate DRs, pairs of DRs were ordered at each value of $\tau_f$, and were all tested in the group of measurements described earlier. For each pair of DRs that came from the same batch, we found that the turnover point for a given value of $\tau_f$ reproduced nominally within ±1°C, and that the second order coefficient $C_2$ reproduced nominally within less than ±5%. A second set of DRs identical to those just discussed was ordered from Murata a year later, and no degradation in the reproducibility of the temperature characteristics has been observed. It should be pointed out that the biggest source of lack of reproducibility will be the tolerance to which the material vendors will guarantee $\tau_f$, where the best tolerance is ±0.5 ppm/°C. The implication of this point with regard to controlling the temperature characteristic will be discussed later. Also regarding the subject of reproducible characteristics, the effects of small perturbations in the positioning of the DR and support relative to the sidewalls of the cavity were studied by disassembling and rebuilding many of the DRAs, and then re-measuring their temperature characteristics. The same degree of reproducibility (as discussed above) was observed. Table 1 summarizes the typical results. It is important to emphasize once again that for all of the 2.0 GHz measurements, the only different component used was the DR itself; all of the DRAs shared the same brass cavity, coupling structure and quartz pedestal. If all of the components of the DRA were to be replaced, the degree of reproducibility as defined here would not apply, and, depending on the "equivalence" of like components, a degraded degree of reproducibility should be expected.

Using a 2.0 GHz Murata Zr/Sn titanate DR, with $\tau_f$ =3.1 ppm/°C, measurements were performed to investigate the change in the temperature characteristic as a function of the height of the DR within the cavity. Keeping in mind that the two-port coupling structure (probes) was located near the bottom of the cavity (h=0), Table 1 also shows the change in the characteristic frequency vs.

*TABLE I*

**Dielectric Resonator Temperature Study Results - Murata Series "U"**

***Frequency= 2.0 GHz DR: RADIUS=1.33 cm. HEIGHT=1.15 cm.***

**CHARACTERISTIC REPRODUCIBILITY**

*POLYNOMIAL COEFFICIENTS*

| $\tau_f$ (ppm/°C) | DR # | $C_3(10^{-5})$ | $C_2(10^{-3})$ | $C_1$ | $C_0$ | Turnover Point (°C) |
|---|---|---|---|---|---|---|
| 1.4 | 1 | -1.42 | -7.77 | -1.33 | 40.77 | -85.59 |
| | 2 | -1.36 | -7.60 | -1.28 | 37.64 | -84.21 |
| 3.1 | 1 | -1.16 | -8.83 | .21 | 1.70 | 11.66 |
| | 1 (repeat)[‡] | -.83 | -8.97 | .21 | 3.14 | 11.93 |
| | 1 (repeat) | -1.39 | -8.63 | .25 | -.19 | 14.65 |
| 3.1 | 2 | -.99 | -9.03 | .25 | .74 | 13.95 |
| 3.1 | 3 | -1.23 | -8.67 | .22 | .40 | 13.03 |
| 3.1 [**] | 3 | -1.68 | -8.77 | .23 | 1.27 | 13.18 |
| 5.2 | 1 | -1.19 | -9.83 | 2.05 | -44.18 | 104.27 |
| | 2 | -.92 | -9.89 | 2.08 | -42.98 | 105.16 |

**TEMPERATURE CHARACTERISTICS vs. QUARTZ SUPPORT HEIGHT**

$\tau_f$ (ppm/°C)=3.1

| Height (cm)[*] | DR # | $C_3(10^{-5})$ | $C_2(10^{-3})$ | $C_1$ | $C_0$ | Turnover Point (°C) |
|---|---|---|---|---|---|---|
| 1.112 | 1 | -7.07 | -6.18 | .73 | -14.04 | 36.41 |
| 1.112 | 1 (repeat) | -6.33 | -6.40 | .72 | -12.81 | 36.50 |
| 1.924 | 1 | -1.16 | -8.83 | .21 | 1.70 | 11.66 |
| 1.924 | 1 (repeat) | -.83 | -8.97 | .21 | 3.14 | 11.93 |
| 1.924 | 1 (repeat) | -1.39 | -8.63 | .25 | -.19 | 14.65 |
| 2.406 | 1 | -0.76 | -8.92 | .05 | 5.95 | 2.69 |

[‡] Repeat indicates that the DRA was completely taken apart, re-assembled, then re-measured.

[*] Fused quartz support height in centimeters. Cavity height = 5.090 cm.

[**] Measurement repeated in a silver plated brass cavity, using same DR and support.

temperature curve for three different support heights. Listed are the coefficients of the polynomial fit of the data, and the resulting turnover temperature. An overall shift in turnover temperature of approximately 34 °C was observed in moving the DR from the bottom to the top of the cavity. Third order contributions to the characteristic become more prominent when the puck is located very close to the bottom of the cavity, and $C_3$ decreases in magnitude as the DR is moved away from the bottom. In addition there is a decrease in the coefficient $C_1$ as the puck is moved from bottom to top, all of which point out that the influence of the components external to the DR is changing as the puck moves along the cavity height axis. The larger magnitude of $C_3$ observed at the cavity bottom might be explained by the fact that when the DR is closer to the coupling structure, the fields just outside the DR will be slightly higher, and this, in combination with the closer proximity to a cavity wall or floor will combine to form a higher frequency sensitivity to small changes in component dimensions, which occur during thermal expansion and contraction.

To confirm that the DR material (and not the brass cavity) was the major source of the parabolic temperature characteristic, a KOVAR cavity was constructed identical to the brass cavity design, including an identical coupling arrangement. A previously characterized DR ($\tau_f$ =3.1 ppm/°C) was placed in the KOVAR cavity along with the quartz support, the temperature characteristic was re-measured and the results are shown in Fig. 6. A value for $C_2$ of approximately -7 ppb/°C$^2$ was observed, and this is not too different from the value measured in the brass cavity (-8.8 ppb/°C$^2$). The cavity material's influence on the DRA frequency vs. temperature characteristic can be estimated by comparison with the case of an empty (unloaded) cavity resonator. Here there is a one to one correspondence between the change in the cavity dimensions due to thermal expansion and the resulting change in the resonant frequency. For example, a 0.1% increase in the cavity size will result in a 0.1% decrease in the resonant frequency. If one were to construct cavity resonators using different cavity materials, it would be found that the difference in the coefficients $C_i$, of the frequency vs. temperature characteristics for each material would correspond to the difference in the thermal expansion coefficients of the material for the like order terms. Using the linear thermal expansion terms, for brass ($\alpha_1 \approx 20$ ppm/°C) and KOVAR ($\alpha_1 \approx 6$ ppm/°C) the difference is ≈14 ppm/°C. The degree of the cavity influence can be described by a weighting factor, so that for the case of an unloaded cavity, this factor would equal one. But in the DRA, where the majority of the stored energy is located within the DR puck, the influence of the cavity is decreased, thus implying a weighting factor less than one. The degree of influence in this case can be estimated by comparing the change in the linear coefficient $C_1$ in going from a brass cavity to a KOVAR cavity. The ratio of this change to the change just discussed for the unloaded cavities is the same as the ratio of the weighting factors for the loaded and unloaded cavities. From the measured data, the values of the DRA coefficients $C_1$ were found

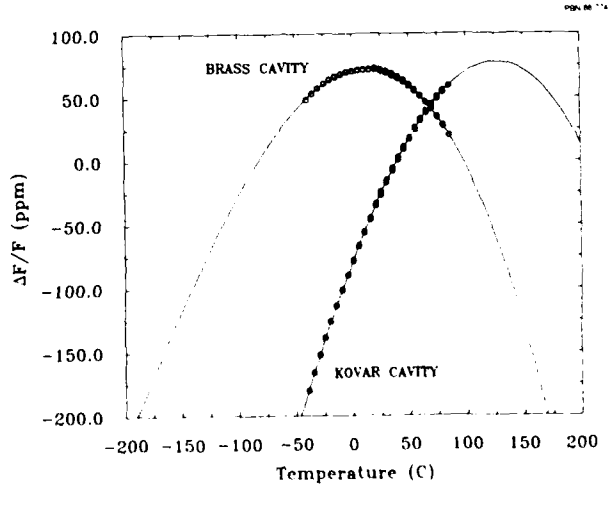



Figure 6. The dependence of a DRA's temperature characteristic on the cavity material, as measured using a Murata Zr/Sn titanate DR ($\tau_f$ =3.1 ppm/°C) in a brass cavity and KOVAR cavity. The parabolic curves fitted to the data are only to help illustrate the linear shift induced by the cavity, and are not necessarily indicative of the frequency behavior outside the temperature range of -40 to 85 °C.

to be 0.2 and 2.3 for brass and KOVAR, respectively. The influence of the cavity on the DRA frequency is then estimated to be:

$$w_{DRA} \approx [\,|\Delta C_{1DRA}| / |\Delta C_{1EMPTY}|\,] w_{EMPTY} \qquad (8)$$

$$w_{DRA} \approx [2.1 / 14] \approx 0.15$$

which indicates for the DRA that the cavity is 85% less influential than in the case of the unloaded cavity. Assuming that this weighting factor does not change significantly for the next couple of higher order terms, we can use this value to estimate the influence of the cavity's higher order thermal expansion characteristics. From Ref. [6] it was found that the second order thermal expansion coefficient for brass was nominally 6 ppb/°C$^2$, but its influence on the second order frequency change is reduced by a factor of 0.15 to approximately 0.9 ppb/°C$^2$ . By comparing this with the -8.8 ppb/°C$^2$ value of $C_2$ calculated for the Murata Zr/Sn titanate material with $\tau_f$ =3.1 ppm/°C, the results indicate that it is the dielectric resonator material which is the dominant factor in determining the second order coefficient in the DRA's temperature characteristic. For the equivalent case of the Trans-Tech material, with $\tau_f$ =0 ppm/°C and $C_2$= -27 ppb/°C$^2$ , the dielectric material is by far the dominant source of the second order coefficient.

Nishikawa, et. al. [5] have suggested that if the cavity and support material have the same coefficient of expansion as the dielectric material, the DRA will have the same frequency vs. temperature characteristic as the DR. KOVAR is a good thermal match to the Murata material (5-6 ppm/°C vs. 6–7 ppm/°C) and indeed we did observe that $\eta_E$ decreased from -2.9 ppm/°C in the brass cavity to –0.7 ppm/°C in the KOVAR cavity. However, $\eta_E$ did not go to zero and this may be due to the fact that our support material was fused quartz, which has a coefficient of expansion that is nearly zero. Also, our cavity design and probe structure was different from that used by Nishikawa.

Using a model for an inhomogeneously loaded DRA developed by Bonetti and Atia [7], an attempt was made to simulate the influence of the cavity walls on the resonant frequency. This was done by perturbing the dimensions of all of the components comprising the DRA to simulate thermal expansion, and then calculating the resulting frequency change. While the calculated value of the resonant frequency of the unperturbed DRA was within the ±1% accuracy stated in Ref. [7], the sensitivity predicted by this model was more than an order of magnitude smaller than that which was observed experimentally. At this time, we have not been able to account for this discrepancy. It should be noted that this particular model does not include the RF coupling structure, which may also contribute to the DRA temperature sensitivity.

Before discussing the oscillator measurements, it is worth summarizing the important points resulting from the DRA temperature characterization experiments.

1. An L-Band dielectric resonator enclosed in a metal cavity results in a device with a long thermal time constant when compared with quartz based resonators.

2. The frequency change versus temperature characteristic of the dielectric resonator materials is non-linear in nature. The characteristics observed for the titanate materials discussed herein are parabolic and can have a positive (barium tetratitanate) or negative (Zr/Sn titanate) second order coefficient with values ranging from -27 to +20 ppb/°C$^2$. In comparison, ST quartz exhibits a second order temperature coefficient of -35 ppb/°C$^2$.

3. The magnitude of the DR materials' second order temperature coefficient is only weakly dependent on the vendor specified various values of $\tau_f$, the vendors' "temperature coefficient of the resonant frequency".

4. For equivalent types of dielectric resonator materials, the second order temperature coefficient of DRs purchased from **different** vendors can vary significantly. For Zr/Sn titanate, Murata material was typically -9 ppb/°C$^2$, while the Trans-Tech material was -27 ppb/°C$^2$.

5. Metal cavities will tend to shift a DR's turnover temperature toward lower temperatures. The results using our cavity designs indicate that the linear shift induced by a metal cavity is directly proportional to the linear thermal expansion coefficient.

Before concluding the discussion of DRAs, we note that measurements are currently underway on the frequency vs. temperature characteristics for a series of 3.0 GHz DRA's constructed similar to the 2.0 GHz DRAs, and using the Murata Series "U" Zr/Sn titanate material. The observed influence of the cavity material on the DR's temperature sensitivity is in excellent agreement with our previous results, also the value of $C_2$ is approximately the same as that observed during the 2.0 GHz measurements. Both of these results suggest that these characteristics are not significantly dependent on the value of the resonant frequency. However, we note that a higher value of $C_3$ has been observed, and this might be accounted for in the slightly different probe structure and the larger diameter quartz support that was used in the DRA.

The final justification for the approach of characterizing only the DRA's temperature sensitivity is presented in the following oscillator measurements. Figures 7 and 8 show the difference in frequency vs. temperature characteristics for a DRA and its corresponding DRO, where the DRA was incorporated as the stabilization element in a basic feedback oscillator. The oscillator electronics were comprised of a Watkins-Johnson A36-2 silicon bipolar transistor amplifier and an Anzac DS-313 power divider. Figure 7 shows the results using a Murata Zr/Sn titanate DR with $\tau_f$=3.1 ppm/°C, and second order temperature coefficient of -8.8 ppb/°C$^2$. Figure 8 shows the results using a Trans-Tech Zr/Sn titanate DR with $\tau_f$=0.0 ppm/°C, and second order temperature coefficient of -27.4 ppb/°C$^2$. The most important aspect of both

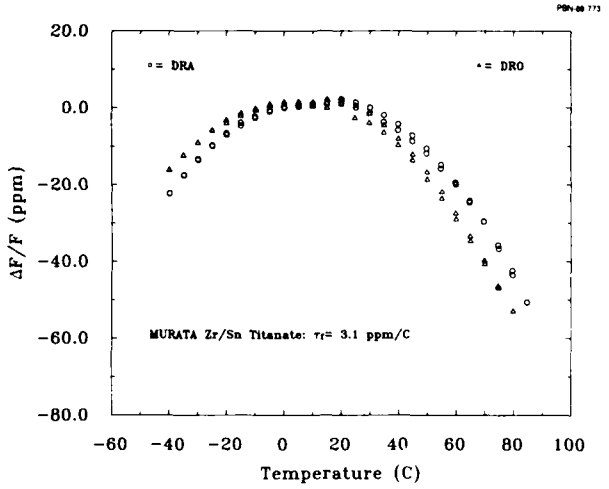



Figure 7. The influence of the oscillator electronics is shown in this comparison of the frequency vs. temperature characteristics between the DRA and DRO for the measurement described in Fig. 4a. For this Murata DR, the value of $\eta_E$ was shifted by -0.10 ppm/°C, and the turnover temperature was shifted by -7.5 °C.

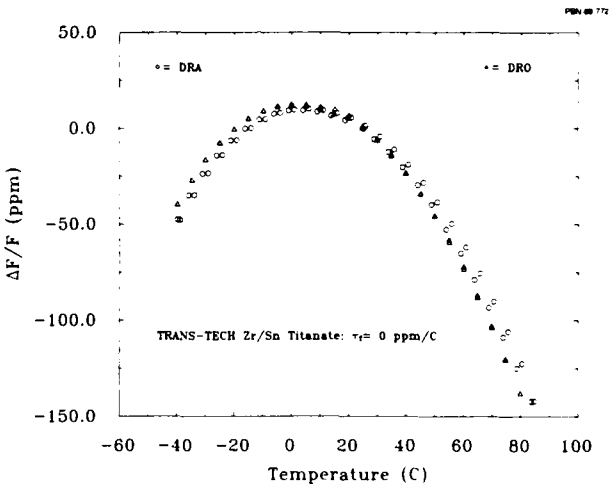



Figure 8. The influence of the oscillator electronics is shown in this comparison of the frequency vs. temperature characteristics between the DRA and DRO for the measurement described in Fig. 4b. For this Trans-Tech DR, the value of $\eta_E$ was shifted by -0.14 ppm/°C, and the turnover temperature was shifted by -2.7 °C.

figures is that the addition of the oscillator electronics has not caused a significant shift in the DRA's frequency vs. temperature characteristic. Again, this is the same result observed in the 1.0 GHz barium tetratitanate DRA/DRO discussed earlier. In terms of the equation for the fractional frequency change versus temperature, the oscillator electronics have shifted the value of $\eta_E$ by -0.10 ppm/°C in Fig. 7 and by -0.14 ppm/°C in Fig. 8. This downward shift is consistent with what we generally observe on similar measurements involving quartz SAW resonators. This corresponds to shifts in the turnover temperature of -7.5 °C and -2.7 °C for Figs. 7 and 8, respectively. The larger turnover temperature shift observed with the Murata DR is a direct result of this DRA's lower second order coefficient, thus making the contribution of the electronics' temperature sensitivity more significant. Therefore, it is clear that a larger downward temperature shift will occur for those DR materials with a smaller second order temperature coefficient. On the other hand, one finds a larger total frequency excursion for those materials having a larger second order coefficient. The consideration of these two properties is most important when choosing components which will minimize a DRO's frequency excursion over temperature.

## *SUMMARY*

Detailed experiments were performed on dielectric resonators placed in metal cavities to determine the exact temperature sensitivity of the resonant frequency for the $TE_{01\delta}$ mode. The results show that the frequency versus temperature characteristic is a parabolic function of temperature, with a parabolicity similar to or smaller than that observed on quartz SAW resonators. It was also found that the commercial wideband electronic components used to construct a truely low-noise DRO do not contribute significantly to the frequency versus temperature characteristic.

Our results indicate that the DR vendors' specification of the "temperature coefficient of the resonant frequency" does not provide the necessary information to accurately predict the temperature sensitivity of a DRO's operating frequency at a given temperature, or the total frequency excursion over a given temperature range. The physical properties of the metal cavity used to house the dielectric resonator are such that it induces significant linear shifts in the DR's temperature characteristic, with the magnitude of this shift dependent on the type of cavity material and the cavity design. In calibrating a cavity and/or oscillator design with any DR (with a known value of $\tau_f$), a measurement of the frequency change versus temperature will provide the value of $\eta_E$, which characterizes the shift in the DR's temperature sensitivity caused by the components external to the DR. Based on our observation that the second order coefficient, $C_2$, of the frequency vs. temperature characteristic is primarily dependent on the source or type of the DR material, the results presented in this paper could be used as a means to estimate the total frequency excursion for a similar DRA/DRO design. Adjustment of the temperature sensitivity is accomplished by picking a new DR with a $\tau_f$ such that its value differs from the calibration DR's value by an opposite amount equal in magnitude to the shift induced by the cavity, plus any additional shift necessary to place the turnover point of the response in the center of the temperature range. The centering of the turnover point will then minimize the total temperature induced frequency excursion. Table 2 summarizes the properties of the materials investigated in this work. Based on our observations for Zr/Sn titanate, one should excercise considerable caution in purchasing DRs from different vendors.

Based on the measurements of the parabolicity of the DR

*TABLE 2*
***SUMMARY of DIELECTRIC RESONATOR MATERIAL SECOND ORDER TEMPERATURE CHARACTERISTICS***

| VENDOR | MATERIAL | $C_2$(ppb/°C²) |
|---|---|---|
| Murata/Erie | Zr/Sn Titanate | -9* |
| Trans-Tech, Inc. | Zr/Sn Titanate | -27** |
| Trans-Tech, Inc. | Barium Tetratitanate | +20 |

*Series "U" material, average value for all $\tau_f$'s measured.
** $\tau_f$=0±0.5 ppm/°C

materials' temperature characteristics, the minimum frequency excursion attainable appears to be with the Murata Series "U" Zr/Sn titanate. Using the nominal -9 ppb/°C² measured, and assuming a turnover point centered over the temperature range of -55 to 85 °C, a minimum frequency excursion of 44 ppm could be possible. In comparison, an ST-cut quartz SAW resonator oscillator, with a -35 ppb/°C² coefficient, would result in a frequency excursion of 172 ppm. Ideally the temperature sensitivity of this DR material is four times better than this commonly used SAW quartz cut. But the vendors of the DR materials generally specify a tolerance of ±0.5 ppm/°C on the temperature coeffficient $\tau_f$. This corresponds to an ability to specify the turnover point of a DR material to within ±31°C, whereas for SAW resonators fabricated on ST cut quartz, the uncertainty on the turnover point is generally ±5°C. Therefore, in the worst case situation, the same DR discussed above could result in a frequency excursion of 92 ppm, while the worst case for quartz would be 197 ppm, which then makes the DR only 2 times better. To take advantage of the best temperature stability currently obtainable using a dielectric resonator, a much tighter tolerance on the value of $\tau_f$ would be required.

## REFERENCES


[1]. T.R. Meeker, "Theory and Properties of Piezoelectric Resonators and Waves," in *Precision Frequency Control: Acoustic Resonators and Filters,* Vol. 1, E.A. Gerber and A. Ballato, Eds. New York: Academic Press, 1985, pp. 48-146

[2]. T.E. Parker and G.K. Montress, "Precision Surface Acoustic Wave (SAW) Oscillators," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control,* Vol. UFFC-35, pp. 342-364, May, 1988.

[3]. G.D. Alley and H.-C. Wang, "An Ultra-Low Noise Microwave Synthesizer," *IEEE Transactions on Microwave Theory And Techniques,* Vol. MTT-27, No.12, pp. 969-974, December, 1979.

[4]. M.J. Loboda, T.E. Parker, and G.K. Montress, "Frequency Stability of L-Band, Two-Port Dielectric Resonator Oscillators," *IEEE Transactions on Microwave Theory and Techniques,* Vol. MTT-35, No.12, pp. 1334-1339, December, 1987.

[5]. T. Nishikawa, et. al., "Precise Measurement Method For Temperature Coefficient of Microwave Dielectric Resonator Material," *1987 IEEE MTT-S International Microwave Symposium Digest,* Vol. 1, pp. 277-280.

[6]. American Society for Metals, *Metals Handbook*, 8th Edition, Vol. 1, Novelty, Ohio, 1961, pp. 1018.

[7]. R.R. Bonetti and A.E. Atia, "Design of Cylindrical Dielectric Resonators in Inhomogeneous Media,"*IEEE Transactions on Microwave Theory and Techniques,* Vol. MTT-29, No.4, pp. 323-326, April, 1981.

42nd Annual Frequency Control Symposium - 1988

# NON-LINEAR MODELING AND PERFORMANCE OF OSCILLATORS USING THIN-FILM BULK-ACOUSTIC WAVE DEVICES

S.G. Burns, P.H. Thompson, and G.R. Kline
Microelectronics Research Center
Iowa State University
1925 Scholl Road
Ames, Iowa 50011

## Abstract

Thin-film bulk-acoustic-wave resonators (TFR) and TFR-based filters have been used in single-line spectra UHF and L-band oscillators in both hybrid and monolithic integrated versions. Common to the design of these circuits is the use of linear circuit models along with the Barkhausen criteria to predict the frequency at the onset of oscillation. However, when used as the gain element in an oscillator circuit, the amplifier operates non-linearly. This non-linearity contributes to the spectral content away from the predicted frequency of oscillation. Two methods are proposed for modeling the behavior of the amplifier in a SPICE simulation. The first employs a non-linear controlled generator derived from measured gain compression data. The second involves a circuit model of the amplifier used in a closed-loop simulation. The method of deriving the amplifier models is presented along with experimental results of the actual oscillator circuits. The oscillators, using commercial hybrid Darlington amplifiers as the gain element, are evaluated with fundamental mode frequencies in the area of 1GHz.

## Introduction

Thin film bulk-acoustic-wave resonators (TFR) and TFR-based filters have been used in single-line spectra UHF and L-band oscillators in both hybrid [1-3] and monolithic integrated[4] versions. Using similar TFR techniques and topologies, semiconductor delay lines using either ZnO or AlN for piezoelectric transduction have also been used as the frequency controlling element of comb oscillators at L-band.[5]

Common to the design of these circuits is the use of linear circuit models, Fig. 1, along with the Barkhausen criteria

$$|G|+|H|=1 \quad (1a)$$
$$\phi_G+\phi_H+\phi_S=0 \quad (1b)$$

to predict the output frequency at the onset of oscillation. However, once oscillation commences, the forward amplifier, G, operates in a non-linear regime. Consequently, linear analysis does not predict saturation effects such as output power at the fundamental frequency and harmonics. Spectral broadening from non-linear behavior also tends to degrade the single-sideband phase noise performance.

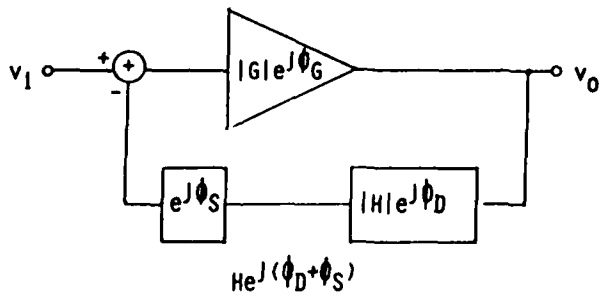



Fig. 1 Basic feedback oscillator. $\phi_S$ represents both parasitic phase shift and phase shift deliberately introduced for tuning along the $\phi_D(f)$ characteristic.

This paper presents a design approach using measured non-linear S-parameter transfer functions for either individual active elements or composite two-port amplifiers and incorporating these transfer functions in closed-loop SPICE simulations. Feedback frequency control utilizes a double-stacked crystal filter (DSCF) synthesized using microelectronics technology in a multilayer structure with AlN as the piezoelectric.[6] The single stacked-crystal filter (SCF) topology was originally proposed using quartz technology.[7]

## Modeling Approaches

One approach used for large-signal circuit characterization consists of the design of microwave FET oscillators by using the large-signal S-parameters to optimize output power.[8] We extended this technique by first measuring S-parameters as a function of input drive level for a commercial 10dB BJT amplifier. These parameters then serve as the basis for a two-port model to which we add the frequency selective feedback network. For example, Fig. 2 is a plot of $V_{out}(P_{in})$ vs. $V_{in}$ derived from measured $|S_{21}|(P_{in})$ to obtain the expected amplifier gain compression characteristic. The basic non-linearity results from the inherent exponential behavior of the BJT transfer function. The amplifier gain may be represented by an $n^{th}$ order polynomial which is of the form $V_{out}(P_{in})=V_o+AV_{in}+BV_{in}^2+\ldots$ in general. The coefficients are obtained by fitting the polynomial to the measured gain compression data. $V_{out}(P_{in})$ is to be included as a non-linear dependent source, E1=f(VX), in a SPICE two-port model, Fig. 3, for the amplifier. $C_F$ and the T-line parameters were selected to obtain the desired small-signal dominant-pole frequency response and linear phase characteristic. In principle, the two-port should include an equivalent non-linear controlled generator derived from power-level dependent $S_{12}$ measurements, however it was observed that measured values of $S_{12}$=-16.7dB 7° made this unnecessary for most calculations. Similarly, $|S_{11}|$ and $|S_{22}|$=-20dB in the frequency range of interest which allows using 50Ω amplifier input and output impedances in the model. We then use this model with a DSCF derived circuit model which will be presented in the next section.

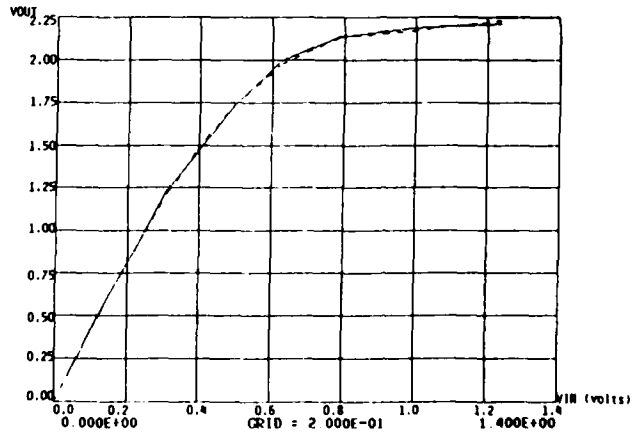



Fig. 2 Forward gain characteristic. $V_{out}$ vs. $V_{in}$.
——— measured
——— polynomial fit
$E1(VX)=0+4.0904(VX)+0.7554(VX)^2-(5.0616)(VX)^3+(2.3723)(VX)^4+(0.0125)(VX)^5$

CH2588-2/88/0000-272 $1.00 C 1988 IEEE

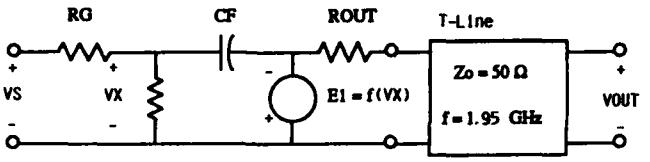



Fig. 3 Two-port model with a non-linear generator to simulate $|S_{21}|$ characteristics. Amplifier T-line designed to match the overall amplifier phase characteristics.

The amplifier non-linearity and gain compression can be modeled directly and compared to measured S-parameter data by deriving a SPICE model directly from the circuit. As shown in Fig. 4a, the gain element consists of a direct-coupled Darlington pair with series feedback, $R_{E2}$ to establish gain and shunt feedback, and $R_F$ for bias stabilization. An external load resistor, $R_C$, and dc blocking capacitors must be used. Fig. 4b illustrates the simulated non-linear transfer function for $V_{CC}$=10 volts at 1093MHz.

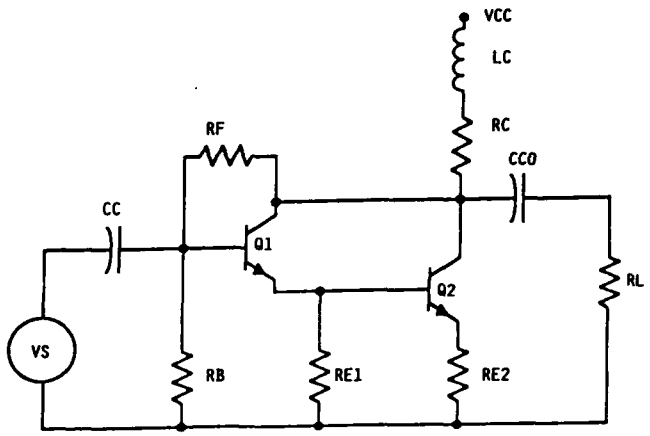



Fig. 4a Basic Darlington pair. 10dB Amplifier.
RF=300Ω RE1=375Ω RC=100Ω
RB=115Ω RE2=10Ω LC=50μH
CC=CCO=1μF RL=50Ω
Q1=Q2 [β=100, IKF=25mA, RC=5Ω, RE=1.2
TF=10.8pS, CJC=0.224pF, RB=8.1
CJE=4.8pF]

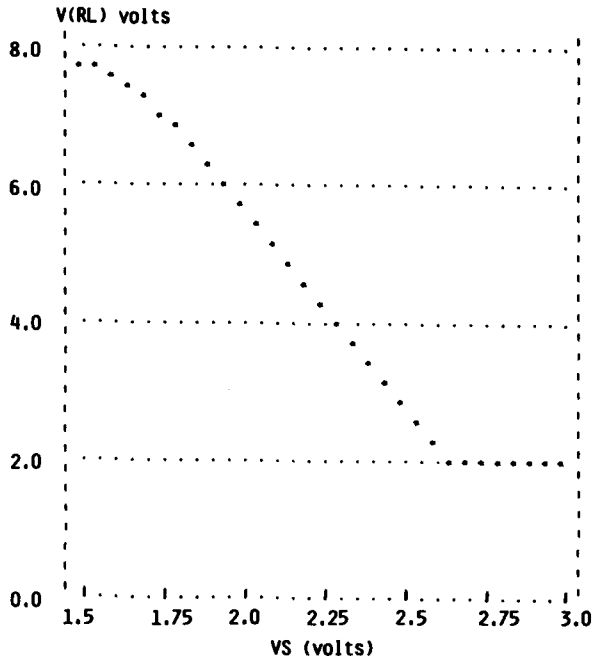



Fig. 4b Non-linear output voltage transfer function for the Darlington. Derived from the Gummel-Poon model.

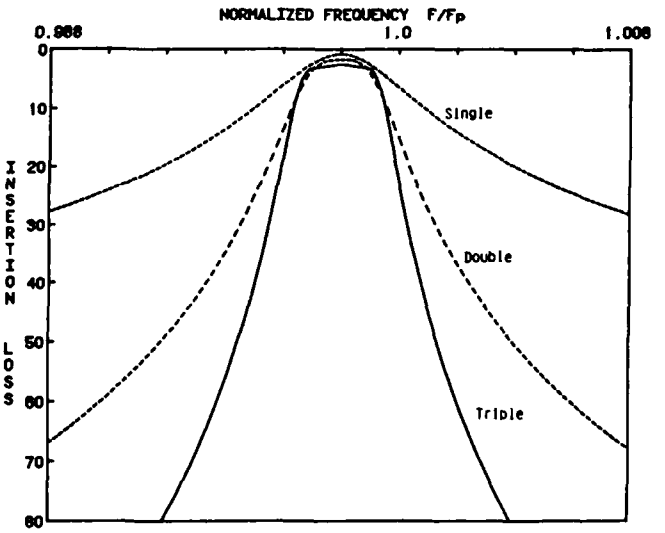



Fig. 5 Analytical results for a single, double, and triple SCF using Mason's model and ABCD two-port network theory.

## Equivalent Circuit of DSCF

Each section of the DSCF is modeled as a five layer structure composed of three finite-thickness metal (Al) films sandwiching two piezoelectric (AlN) layers. The diameter-to-thickness ratio for a DSCF is in excess 50 so that acoustically the structure can be considered as one-dimensional. Each section can be represented as a transmission line with the overall cascade response computed using ABCD two-port network theory.[9] Furthermore, the cross-sectional area can be tailored to obtain 50Ω input and output impedances. Analytical results for a single-, double-, and triple-stack are shown in Fig. 5. The one-dimensional wave propagation and boundary conditions means that the DSCF can be modeled as a cascade of two Pi-networks shown as the shaded portion in Fig. 6. LS, CS, and CO are the basic Butterworth-van Dyke model parameters. RS represents loss in the piezoelectric and is approximated by

$$RS = \frac{2}{\omega(CO)k^2Q} \qquad (2)$$

where $k^2$ is the coupling coefficient and $Q = \frac{f}{2}\frac{d\phi_H}{df}$.

To improve the accuracy of the model when used as the design macro in a SPICE simulation, we incorporate the following parasitic elements as shown in Fig. 6. LWB1 and LWB2 are the bonding wire inductances. RB1 and RB2 are included to account for bonding wire and metallization losses. LG models a measured common inductance in the die-to-package ground. CP1 and CP2 are parasitic input and output capacitances and are slightly asymmetrical due to package layout.

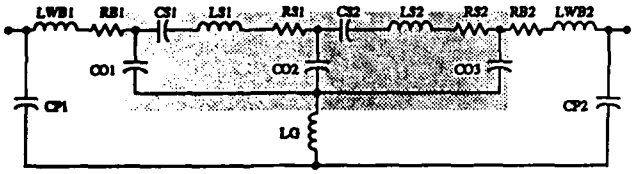



Fig. 6 DSCF model (shaded) with parasitics optimized using TOUCHSTONE.

LWB1=2.198nH CP1=0.0538pF
LWB2=2.195nH CP2=0.0540pF
RB1=RB2=1.5Ω LG=0.0817nH
CO1=CO3=$\frac{CO2}{2}$=3.835pF CS1=CS2=0.07785pF=$k^2$(CO)
LS1=LS2=278.2nH RS1=RS2=2.765Ω
LS=$\frac{A}{(2\pi f_o)^2 CS}$, $f_o$ center frequency
A optimization variable
$\varepsilon_r$=10 to compute CP

Frequency swept S-parameter data were collected using an automated HP 8505 network analyzer. Equivalent circuit element values were obtained from these data using optimization routines available on TOUCHSTONE.[10] Initial values were obtained from basic Pi-network mathematical relationships with results scaled by the driving-point 50Ω impedances. Equivalent circuit results from a SPICE 2G.6[11] simulation along with the measured filter characteristic are shown in Figs. 7a and 7b respectively. The TOUCHSTONE optimization was then transferred to SPICE to take advantage of the extensive active device model library.[12]

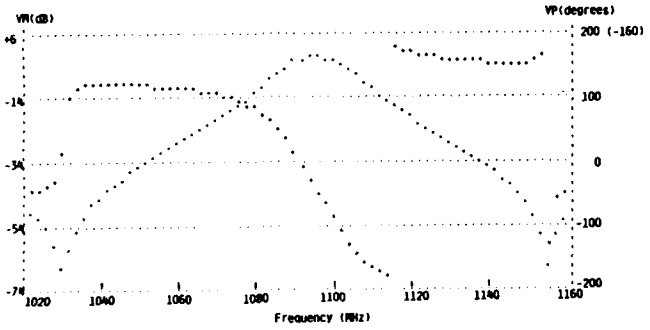



Fig. 7a DSCF SPICE simulation.

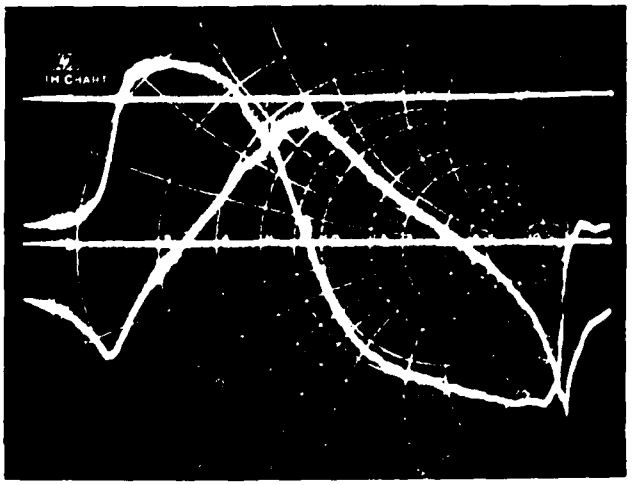


Fig. 7b Magnitude and phase.
DSCF packaged 1094MHz ± 75MHz (15MHz/div)
IL~4dB
10dB/div and 45°/div

## Model and Experimental Results

The closed-loop gain and phase simulations for the DSCF, Fig. 8, are included as the feedback elements in an oscillator using the Darlington pair for the gain element. We include a variable frequency voltage generator in series with the loop and observe that at the onset of oscillation, based upon the Barkhausen criteria, the loop phase shift is 0° and the loop gain exceeds unity at nominally 1088MHz. Large signal behavior may now be studied by performing a SPICE.TRAN (transient) analysis with a variable amplitude excitation at the oscillator fundamental output frequency followed by a .FOUR (9 term FFT) computation to obtain an estimate of the relative harmonic amplitudes at various levels of bias current. Using the non-linear controlled generator model, the simulation predicted qualitatively the odd harmonic content of the frequency spectrum but failed to predict even harmonic generation due to the inherent assymmetry of the amplifier's voltage transfer function. The second model of the amplifier gives a more realistic prediction of the spectral content of the actual oscillator circuit. Figs. 9a and 9b summarize the measured results over the range from maximum power output with 6dB of excess gain to oscillator operating threshhold respectively. The fundamental output power decreases by 15.2dB while the second and third harmonics decrease 22.2dB and 30.0dB respectively. These data show, as expected, that harmonic content increases with additional excess loop gain pointing to the need for AGC in these circuits. Qualitatively, phase noise performance degrades from spectral broadening. Although specific phase instability in the amplifier cannot be modeled directly on SPICE, overall phase-slope sensitivity can be studied by using the .TRAN simulation with single-frequency FM excitation given by

VS=V0+VA sin(2π(FC)t)+MDI sin(2π(FS)t)=
SSFM(V0 VA FC MDI FS) (3)

FC is set to 1090GHz and FS is the carrier offset for which 1kHz or 10kHz are typically used although to trace the characteristic curve requires a series of calculations, each at a different FS. Satisfying resolution requirements for calculating these curves requires excessively long run times. Better results can be obtained by using loaded Q values and amplifier performance parameters in Leeson's noise equation.[13]

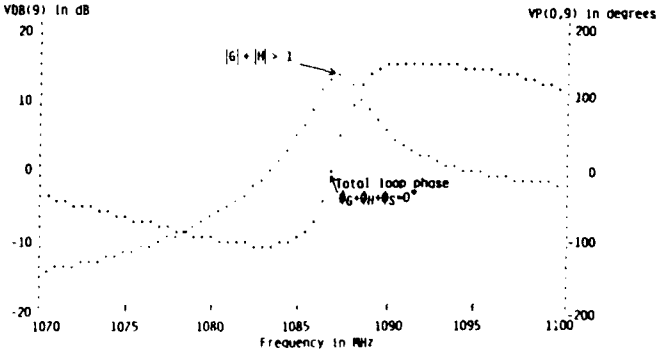



Fig. 8 Darlington pair closed-loop response with DSCF+Barkhausen criteria satisfied.

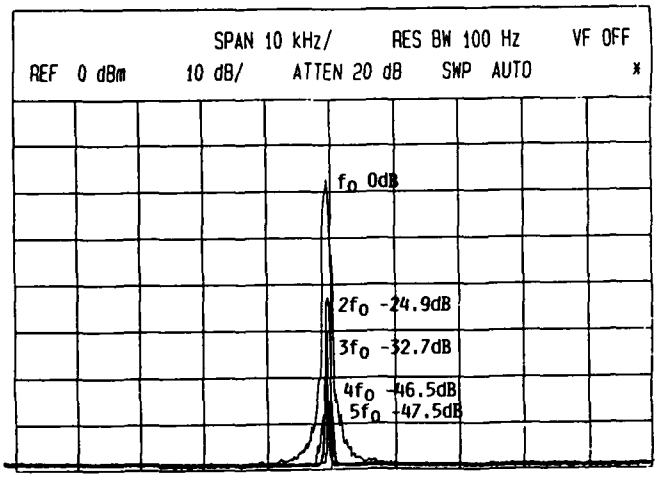



Fig. 9a Bias 37mA maximum output at $f_o$, 22.7dBm. Overlay of first 5 harmonics where $f_o$=1090MHz.

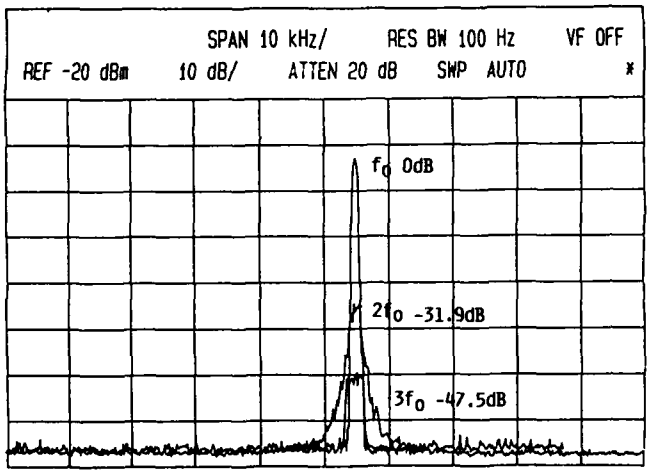



Fig. 9b Bias 13.5mA output at $f_o$, 7.5dBm. Overlay of first 3 harmonics where $f_o$=1090MHz.

## Conclusions and Direction of Future Work

A DSCF has been employed as the frequency control element in an L-band delay-line oscillator. Linear circuit models, based upon one-dimensional microwave acoustic behavior, were used in a TOUCHSTONE™ optimization of swept-frequency S-parameter data. Packaging and device parasitics were included. Two basic large-signal models were employed for the forward amplifier. One approach used a two-port with an S-parameter derived polynomial transfer function for the voltage-controlled generator. The second approach, which was somewhat more accurate in harmonic content prediction, used a complete Gummel-Poon integral charge control model for the BJTs in the Darlington amplifier. A linear closed-loop SPICE simulation predicted the onset of oscillation. As expected, an increase in excess gain increased the harmonic content pointing to the need for AGC control or a matched bandpass output filter. An FFT SPICE simulation of the large-signal output waveform also was used to study harmonic content.

Work continues in exploiting the TFR technology for oscillator design as well as their deployment in more complex circuits and systems, both hybrid and monolithic. A key design criteria is assessing their potential for realization as a monolithic circuit. In particular, VCO operation using the DSCF as well as overmoded comb operation is being pursued. VCO operation in conjunction with temperature coefficient measurements would yield TCXO behavior. More advanced filter models would employ provisions for including the effects of multi-mode harmonic and anharmonic behavior. Because the ultimate phase noise performance of these circuits depends not only on the frequency-controlling Q but on amplifier non-linearity, work continues on non-linear modeling and the incorporation of amplitude-leveling feedback.

## REFERENCES


1. R.A. Moore, J.T. Haynes, B.R. McAvoy, "High Overtone Bulk Resonator Stabilized Microwave Sources", IEEE Proceedings of the 1981 Ultrasonics Symposium, pp. 414-424.

2. S.G. Burns, G.R. Kline, K.M. Lakin, "UHF Oscillator Performance Using Thin-Film Resonator-Based Topologies", IEEE Proceedings of the 41st Annual Symposium on Frequency Control, May 1987, pp. 382-387.

3. M.M. Driscoll, S.V. Krishnaswamy, R.A. Moore, J.R. Szedon, "UHF Film Resonator Evaluation and Resonator-Controlled Oscillator and Filter Design Using Computer-Aided Design Techniques", 1985 IEEE MTT-International Symposium.

4. W.A. Burkland, A.R. Landin, G.R. Kline, R.S. Ketcham, "A Thin-Film Bulk-Acoustic-Wave Resonator-Controlled Oscillator on Silicon", IEEE Electron Device Letters, November 1987.

5. S.G. Burns, G.R. Kline, K.M. Lakin, "Design and Performance of Oscillators Using Semiconductor Delay Lines," IEEE Proceedings of the 1987 Ultrasonics Symposium, October 1987, pp. 369-372.

6. G.R. Kline, R.S. Ketcham, K.M. Lakin, "Low Insertion Loss Filters Synthesized with Thin-Film Resonators", IEEE Proceedings of the 1987 Ultrasonics Symposium, October 1987, pp.375-380.

7. A. Ballato and T. Lukaszek, "A Novel Frequency Selective Device: The Stacked Crystal Filter," Proceedings of the 27th Annual Frequency Control Symposium, June 1973, pp. 262-269.

8. R.J. Gilmore and F.J. Rosenbaum, "An Analytic Approach to Optimum Oscillator Design Using S-Parameters," IEEE MTT, Vol. MTT-31, No. 8, pp. 633-639, August 1983.

9. K.M. Lakin, "Equivalent Circuit Modeling of Stacked Crystal Filters," Proceedings of the 35th Annual Frequency Controlled Symposium, 1981, pp. 257-262.

10. J. Guild and S. Rein, EEsof TOUCHSTONE™ USER'S MANUAL, TS1-4B-060386, All Versions, June 1986.

11. A. Vladimirescu, Kaihe Zhang, A.R. Newton, D.O. Pederson, A. Sangiovanni-Vincentelli, "SPICE Version 2G User's Guide," University of California, Berkeley. August 1981.

12. P. Antognetti and G. Massobrio, Editors, Semiconductor Device Modelling with SPICE, McGraw Hill, 1988.

13. E.A. Gerber and A. Ballato, Editors, Precision Frequency Control Oscillators and Standards, Vol. 2, pp. 68-75, Academic Press, 1985.

42nd Annual Frequency Control Symposium - 1988

# A WIDE FREQUENCY RANGE, SURFACE MOUNTABLE, VOLTAGE CONTROLLABLE CRYSTAL OSCILLATOR FAMILY

S. M. Logan, D. M. Embree, R. E. Sheehey, D. S. Stevens
AT&T Bell Laboratories, No Andover, MA 01845

## 1. Introduction

The increasing level of integration in electronic systems has lead to demand for compact, surface mountable quartz crystal oscillators and voltage controllable quartz crystal oscillators. Other features of interest to system designers include the ability to set the oscillator output into a high impedance state for board testing and TTL/CMOS compatibility. The following paper discusses an oscillator family based on two versions of a 1.5μ CMOS technology oscillator and a rectangular AT cut quartz resonator hermetically sealed in a 6 lead ceramic package. Section 2 discusses the CMOS device and quartz resonator used in this family of quartz based oscillators. Section 3 describes the physical design of the oscillator family.

## 2. 1.5μ CMOS Oscillator and Quartz Resonator

### 2.1 Oscillator Architecture

To provide a wide frequency range using a fundamental mode AT cut rectangular resonator in a compact package, two versions of a CMOS design are used. A low frequency device contains 16 flip-flop stages and an encoder to provide an output frequency which is a binary submultiple of the resonator frequency. Four digital inputs are used to select one of sixteen binary submultiples between $2^1$ and $2^{16}$. For resonator frequencies between 10 Mhz and 20 Mhz, the resulting output frequency lies between 10 Mhz and just under 200 Hz.

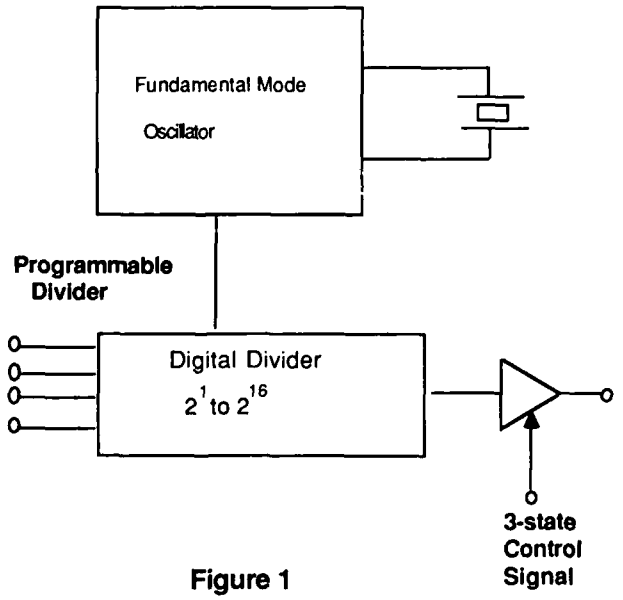



**Figure 1**

**Low Frequency Crystal Oscillator**

A high frequency version of the CMOS device is used for output frequencies above 10 Mhz. The maximum output frequency of the CMOS device in the ceramic package is limited by the frequency of rectangular resonators presently available and is presently about 35 Mhz.

This version of the chip has an additional input which is used for adjusting the logic threshold to permit driving both TTL and CMOS inputs with good transient performance. The user may apply a logic input to center the symmetry of the waveform for driving logic thresholds of 1.40 V (TTL) or $V_{DD}/2$ (CMOS). Figures 1 and 2 illustrate the block diagram of the low and high frequency versions of the CMOS devices used in the oscillator. Both versions of the oscillator chip may be set to a high impedance state to facilitate circuit board testing.

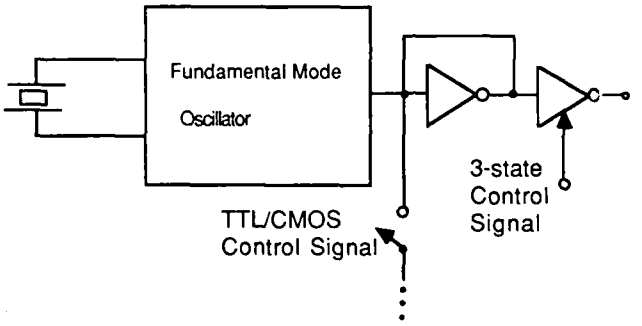



**Figure 2**

**High Frequency Crystal Oscillator**

2.1.1 Internal Varactor An additional feature of both versions of the CMOS oscillator devices is an internal varactor. MOS technology possesses a capacitance-voltage characteristic first described analytically by Goeztberger in 1966[1]. By optimizing the overall topology, a high Q varactor with a sufficient capacitance range is achieved to meet total frequency deviation requirements. The use of the varactor is optional and hence the low and high frequency versions of the oscillator are suited for VCXO or non-VCXO operation. Figure 3 shows the resulting output frequency deviation with control voltage for the resonator parameters shown. The total frequency deviation is directly related to the $C_o/C_1$ of the resonator as well as any added capacitance across the MOS varactor. This latter consideration will be further discussed in Section 3.

CH2588-2/88/0000-276 $1.00 C 1988

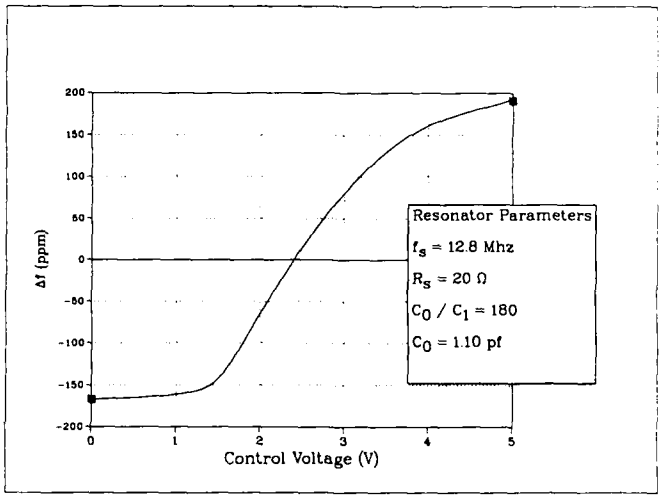



**Figure 3**

**Output Frequency versus Control Voltage**

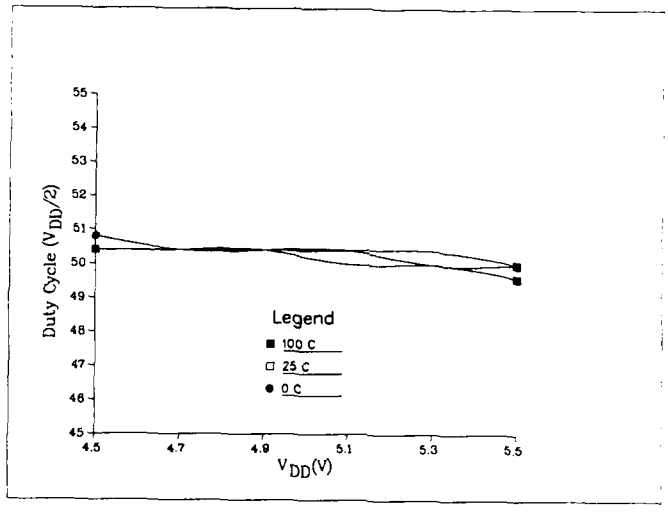



**Figure 4**

**Duty Cycle versus Supply Volltage**

($f_S$= 20 Mhz, $C_L$= 15pf)

2.1.2 Output Waveform Symmetry An additional parameter of interest is the symmetry of the output waveform. The divider version of the CMOS oscillator always divides the fundamental resonator frequency by at least two. Since the difference in propagation delays of the logic gates involved are small compare to the period of the waveform for reasonable values of load capacitance (≤50pf), the output waveform symmetry is within 5% of 50% at TTL and CMOS logic thresholds.

The high frequency version of the CMOS oscillator provides an output frequency whose period may not be large with respect to the difference in gate propagation delays. To provide good output waveform symmetry, a limiting arrangement at the actual resonator node is used. The limiting devices prevent the main oscillating device from entering its non-linear region of operation. Since the AC current through the crystal unit is sinusoidal, the waveform at the output of the oscillating transistor is basically sinusoidal. By coupling this waveform into an amplifier which is biased at exactly its threshold voltage, the output of this amplifier (which presumably swings between the supply voltages) will be very symmetric about the amplifiers' switching threshold. If the switching threshold is to be centered at an arbitrary level, it is only necessary to offset the switching threshold of the stage. This technique is used to provide an output waveform symmetry for the other logic families. By injecting an appropriate signal into the amplifier bias network, the switching threshold is moved to provide output waveform symmetry for the TTL logic threshold of 1.40 V. In the case of CMOS technology where the switching threshold is defined to be one half the supply voltage, no adjustment is necessary to the amplifier threshold. The output waveform symmetry is dependent on the load capacitance and since propagation delay differences are magnified as the load capacitance increases, the output waveform symmetry is degraded. Figure 4 illustrates the stability of the duty cycle with supply voltage at 0°C and 100°C when driving a 15pf CMOS load.

2.2 Quartz Resonator

To provide a substantial size reduction in overall oscillator size, the design of a new quartz resonator and mounting arrangement was required.

The resulting resonator design has a length of 0.200" and a width of approximately 0.090". The gold electrode pattern provides resonator terminals at one end of the plate to provide the compact packaging arrangement shown in Figure 5. The resonator plate is cantilevered to one of the two ceramic shelves in the ceramic package with electrical connectivity established by silver filled conductive epoxy and a wire bond. When mounted in the ceramic package, the series resistance is less than 30 ohms for frequencies above 10 Mhz. The $C_0/C_1$ ratio of the resonator is less than 200. This mounting arrangement is discussed in more detail in Section 3.

## 3.Physical design

To meet size objectives, a custom ceramic package to house the quartz resonator and CMOS silicon chip was designed. The six pin, three layer ceramic package is shown in Figure 5. The 0.300" x 0.340" x 0.120" package has six side brazed leads and is designed to be machine insertable. The top package layer has a gold seal ring to contact the solder preform on the metal cover. When sealed, the cover is connected to the oscillator circuit ground.

The silicon die is bonded to the ceramic substrate with conductive epoxy. The quartz resonator, as shown, is epoxied at one end to a ceramic shelf inside the package. The conductive epoxy provides one electrical connection to the crystal electrode via a thick film gold path on the shelf.

As was mentioned in Section 2, the overall frequency pull range of a voltage controllable crystal oscillator is dependent on parasitic capacitance across the varactor terminals. To minimize the effect of the package capacitance, the second resonator electrode is connected directly to the silicon varactor with a gold wire bond. In essence, this takes maximum advantage of the low $C_o/C_1$ ratio inherent in the resonator design for VCXO applications. Once assembled, the oscillator is tuned to the required frequency using standard Au evaporation techniques.

The package and metal cover are then solder sealed in an $N_2$ atmosphere. In surface mount applications, the package leads are formed in a gull, inverted L or butt lead configuration and then solder dipped.

Using an impulse duration of 0.3 ms, the oscillator package is shock resistant to over 3000g when tested in each of the three axes. Vibration testing over a frequency range of 600 Hz to 2.0 Khz shows an acceleration immunity of $\geq$ 100g for each axis.

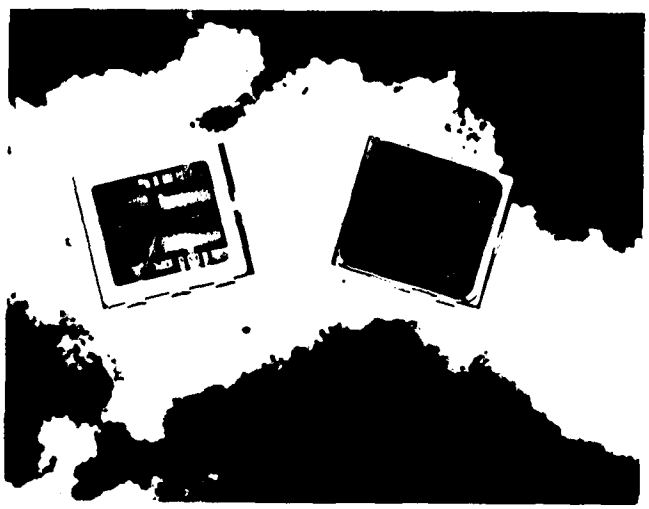


**Figure 5**

**Photograph of 6 pin AT&T Voltage Controllable Crystal Oscillator Package**

## 4. Conclusion

A family of quartz crystal based oscillators using a 1.5μ CMOS technology oscillator chip and rectangular AT cut resonator is described. Its' unique compactness is achieved by the use of a custom ceramic package.

A voltage controllable crystal oscillator requires no additional components since the varactor is integrated on the CMOS chip. This arrangement provides a highly reliable device and is the smallest VCXO package available to date.

## References

1. Goetzberger, A. "Ideal MOS Curves for Silicon", Bell System Technical Journal, September 1966, pp 1097-1122.

42nd Annual Frequency Control Symposium - 1988

# THE INFLUENCE OF PRESSURE AND HUMIDITY ON THE MEDIUM AND LONG-TERM FREQUENCY STABILITY OF QUARTZ OSCILLATORS

by
Fred L. Walls
Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

## Summary

The medium-and long-term frequency stability of most quartz-crystal-controlled oscillators is degraded by various environmental effects. The most important of these are acceleration, temperature, load change, humidity and possibly pressure. In this paper we show data which indicate that the medium-and long-term frequency stability of some oscillators can be improved by controlling the moisture and pressure around the oscillator. Measurements on four different quartz-crystal-controlled oscillators of three different designs yielded improvements of 2 to 5 in frequency stability for measurement times of 1 to 11 days. The frequency stability of one oscillator, with very low drift, improved to $3 \pm 1 \times 10^{-13}$ for measurement times from 0.03 s to 21 days. Supplemental experiments indicate that the probable cause for these improvements is the stabilization of frequency changes due to moisture that corresponds to a fractional change in frequency of about $10^{-9}$ for a 100% change in the relative humidity at room temperature. If these improvements can be routinely obtained in other precision quartz-crystal-controlled oscillators, then they may become useful for some applications generally thought to require atomic standards.

## Introduction

Most precision, quartz-crystal-controlled oscillators exhibit a flicker-of-frequency floor which is approximately given by the phenomenological equation $\sigma_y(\tau) = 2.8 \times 10^{-7}/Q$ for measurement times of order $Q/\nu$. Here, $\sigma_y(\tau)$ is the fractional frequency stability, $\nu$ is the oscillation frequency, and Q is the unloaded quality factor of the resonator [1-7]. For times much longer than $Q/\nu$, the frequency stability generally is degraded by various environmental effects. The predominant processes in the long term appear to be random-walk-frequency modulation and frequency drift. The most important systematic effects are acceleration, temperature, load change, moisture (humidity) and possibly pressure. Some oscillators may also show a sensitivity to magnetic field probably caused by the electronics. The level of flicker frequency modulation appears to be very closely tied to the inverse of the fourth power of the unloaded quality factor, $Q^{-4}$. Gagnepain has shown that the predominant contribution to the flicker level is acoustic scattering losses [6].

In this paper we show experimentally that when the obvious environmental effects such as acceleration, temperature, pressure, moisture, and load changes are controlled, the frequency stability (of at least one crystal controlled oscillator) is constant within a factor of 2 for measurement times from 0.03 s to at least $1.8 \times 10^6$ s (21 days). The random walk process normally observed in quartz-crystal-controlled oscillators appears to be eliminated, and only flicker and frequency drift are apparent in the long term. This indicates that the observed flicker-frequency spectrum is not just a superposition of several noise processes over a narrow range of times, but rather represents a fundamental noise process within the oscillator (quite likely due to the quartz resonator) which extends over 7 orders of magnitude in Fourier frequency or averaging time. If such performance can be duplicated in other oscillators, then the areas in which quartz-crystal-controlled oscillators can be used may potentially be expanded.

Measurements on three other oscillators also show improvements of a factor of 2 to 5 in performance when the humidity and pressure are held approximately constant. These oscillators do not, however, maintain the exceptionally high level of performance demonstrated by the first oscillator. Several experiments show that the most probable cause for these improvements is the stabilization of moisture induced changes in frequency which at room temperature can approach a few parts in $10^9$.

## Measurements of Medium-and Long-Term Frequency Stability

We have recently made measurements on the medium-and long-term frequency stability of several bulk-wave, quartz-resonator-controlled oscillators. Oscillators #1 and #2 use a traditional fifth-overtone, 5 MHz, AT-cut resonator. These oscillators are of different design and from different manufacturers. Oscillators #3 and #4 use a fifth-overtone, 5 MHz, AT-cut resonator of the BVA design [8]. All are precision oscillators exhibiting very good frequency stability for measurement times from 0.03 s to 100 s. For these tests the changes in acceleration, vibration, and load impedance were minimized. The initial tests were performed in a room with a temperature of about $24 \pm 1°C$.

Figures 1 and 2 show the response of oscillators #1 and #3 to a step change in relative humidity from about 20% to 100% created by loosely enclosing them in aluminum foil with a wet sponge. The magnitude of the observed change in frequency is enormous for oscillators with a flicker frequency level of approximately $3 \times 10^{-13}$. The time constant in both cases was several days. The excess moisture was then removed and the oscillators run for several days. The frequency of the oscillators was then recorded while they were rotated about five of the six faces. These data are shown in figures 3 and 4. The steady state levels are primarily due to the acceleration sensitivity. It is interesting to see that for oscillator #1 there is a transient change in frequency of order 1 part in $10^{10}$ shortly after the oscillator is rotated. Quite likely this transient effect is due to a change in the convection cell within the oven, which changes the

Contribution of the U.S. Government, not subject to copyright.

thermal gradients. The changes in frequency for the five different orientations of oscillator #3 were about four times smaller due to a lower acceleration sensitivity. The thermal transients are not noticeable except when the oven is tipped up on end. Measuring the frequency transients under rotation appears to be a simple test to determine the relative importance of thermal gradients in oscillators. The data of figures 3 and 4 show that the orientation (averaged over the measurement time) of oscillator #1 needs to be stable to $10^{-4}$ radians and the orientation of oscillator #3 needs to be stable to 5 x $10^{-4}$ radians in order not to compromise frequency stability measurements at a level of 3 x $10^{-13}$.

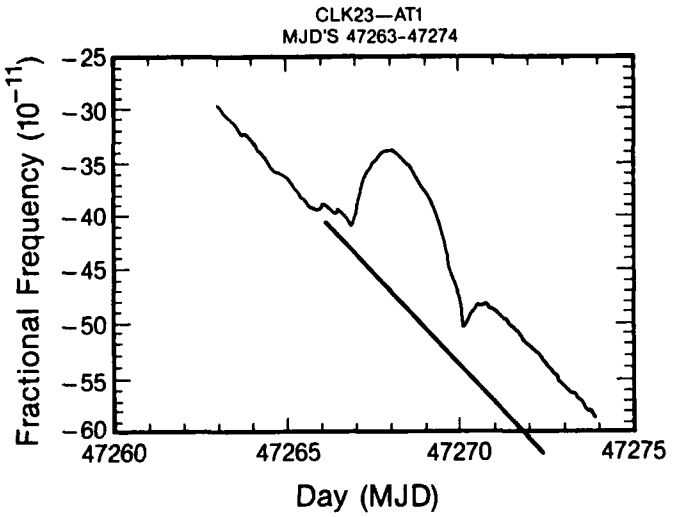



Figure 1. Fractional frequency offset from the NBS time scale of oscillator #1 for a change in relative humidity from approximately 20% to 100% at a nominal temperature of 24°C. This oscillator used a fifth-overtone, 5 MHz, AT-cut resonator of traditional design .

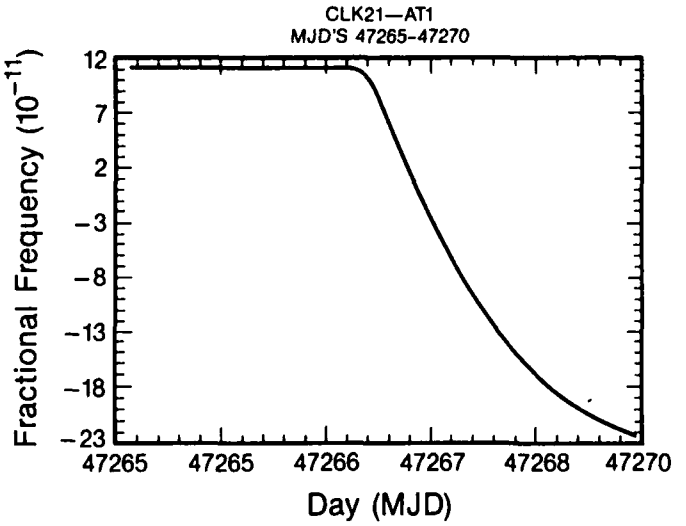



Figure 2. Fractional frequency offset from the NBS time scale of oscillator #3 for a change of relative humidity from approximately 20% to 100% at a nominal temperature of 24°C, with the frequency drift removed. This oscillator used a fifth-overtone, MHz, AT cut resonator of the BVA design [8].

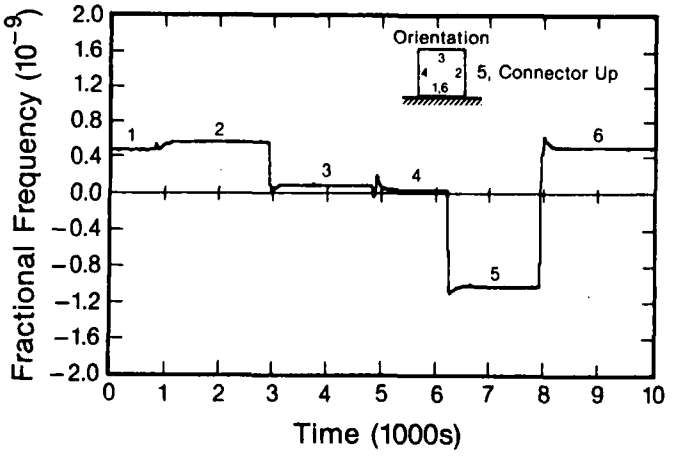



Figure 3. Fractional frequency offset from the NBS time scale of oscillator #1 for 6 different orientations. The last orientation is the same as the first. Note the frequency transient after each rotation.

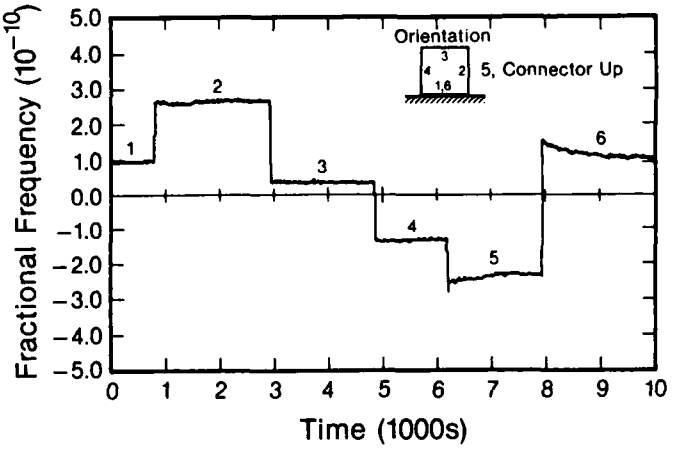



Figure 4. Fractional frequency offset from the NBS time scale of oscillator #3 for 6 different orientations. The last orientation is the same as the first.

Curve A of figure 5 shows the frequency stability of oscillator #1 (with a traditional fifth-overtone, 5 MHz resonator) under conditions of nominally constant power supply voltage, constant load, constant average acceleration, and a temperature of 24 ± 1°C. The humidity and atmospheric pressure were not controlled. Curve B of figure 5 shows the performance of the same oscillator under similar

conditions except that the unit was sealed inside a sturdy metal container to hold the moisture and pressure approximately constant. The noise in curve A is predominantly random-walk frequency modulation, while that of curve B tends to be more like flicker-frequency modulation in the long term.

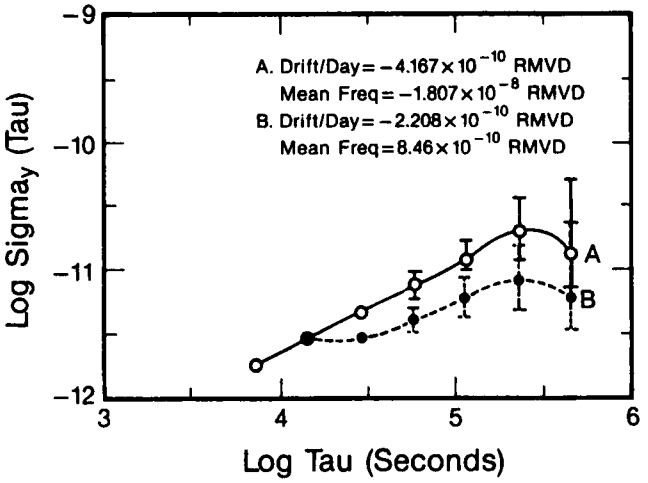



Figure 5. Curve A shows the fractional frequency stability, $\sigma_y(\tau)$, of oscillator #1 as a function of measurement time at a temperature of 24 ± 1°C and normal laboratory humidity and pressure. A second difference has been removed as an estimate of frequency drift. Curve B shows the same clock with the moisture and pressure around the oscillator stabilized and all other conditions similar to curve A. This oscillator used a traditional fifth-overtone, 5 MHz, AT cut resonator.

Curve A of figure 6 shows the performance of oscillator #2 (also with a traditional fifth-overtone AT cut resonator) exposed to room pressure and humidity, while curve B shows the same oscillator with the moisture and pressure held approximately constant. Oscillator #2 also shows some indication of improved performance when the moisture is held constant.

Curve A of figure 7 shows the performance of oscillator #3 (with a fifth-overtone, 5 MHz, AT-cut resonator of the BVA design) exposed to room humidity and pressure, while curve B shows the same oscillator after holding the moisture and pressure approximately constant. The noise in curve A is predominantly random-walk-frequency modulation while that of curve B is more like flicker-frequency modulation in the long term.

Curve A of figure 8 shows the performance of oscillator #4 (with a fifth-overtone, 5 MHz AT cut resonator of the BVA design) exposed to room humidity and pressure, while curve B shows the performance of the same oscillator when the moisture and pressure are held approximately constant. The temperature stability was improved to approximately ± 0.1°C for curve B. Degrading the temperature stability to ± 1°C degrades the frequency stability approximately 50%. We have also included the short-term-stability performance for a measurement bandwidth of 100 Hz.

The 90% error bars on the medium-and-long term data have been calculated for flicker noise. The change in frequency of approximately 1 x $10^{-11}$ between curves A and B appears to be due a change in the load impedance from changing the cable length. This is a typical sensitivity to load changes for precision quartz-crystal-controlled oscillators.

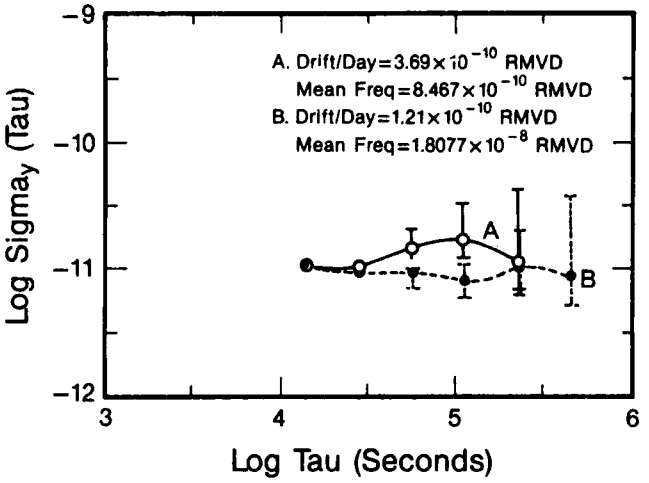



Figure 6. Curve A shows the fractional frequency stability, $\sigma_y(\tau)$, of oscillator #2 as a function of measurement time at a temperature of 24 ± 1°C and normal laboratory humidity and pressure. A second difference has been removed as an estimate of frequency drift. Curve B shows the same oscillator with the moisture and pressure around the oscillator stabilized and all other conditions similar to curve A. This oscillator used a traditional fifth-overtone, 5 MHz, AT cut resonator.

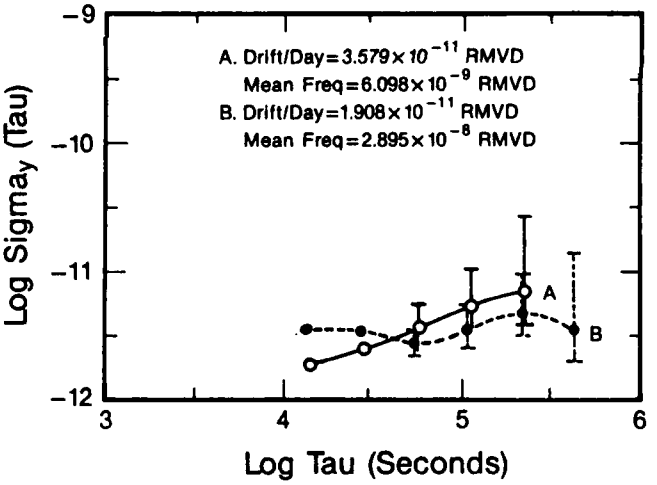



Figure 7. Curve A shows the fractional frequency stability, $\sigma_y(\tau)$, of oscillator #3 as a function of measurement time at a temperature of 24 ± 1°C and normal laboratory humidity and pressure. A second difference has been removed as an estimate of frequency drift. Curve B shows the same oscillator with the moisture and pressure around the oscillator stabilized and all other conditions similar to curve A. This oscillator used a fifth-overtone, 5 MHz, AT cut resonator of the BVA design.

Figure 9 shows the time residuals of oscillator #4 after the second difference was removed as an estimate to frequency drift.

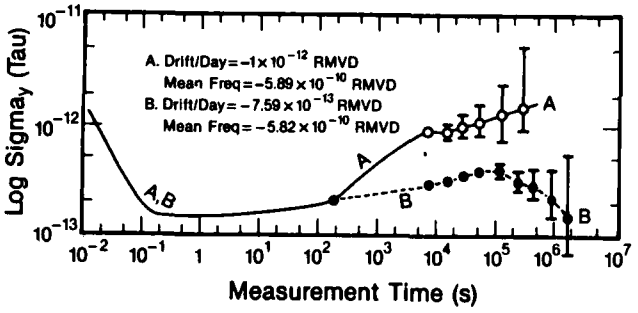



Figure 8. Curve A shows the fractional frequency stability, $\sigma_y(\tau)$, of oscillator #4 as a function of measurement time at a temperature of 24 ± 1°C and normal laboratory humidity and pressure. A second difference has been removed as an estimate of frequency drift. Curve B shows the same oscillator with the moisture and pressure around the oscillator stabilized and the temperature stabilized to approximately 0.1 C. This oscillator used a fifth-overtone, 5 MHz, AT cut resonator of the BVA design.

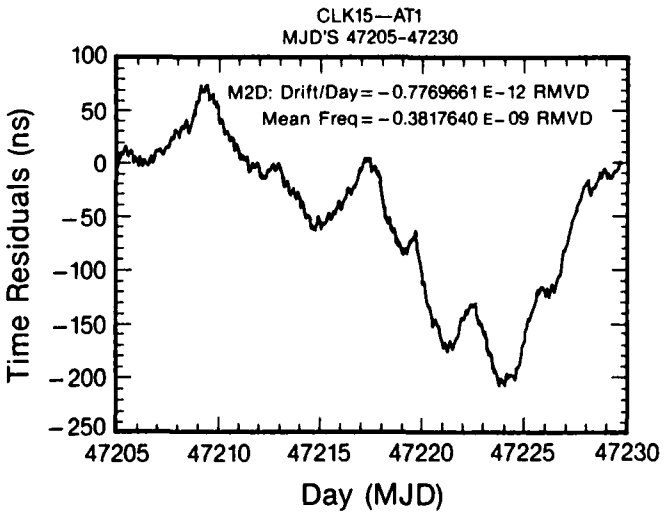



Figure 9. Time residuals of oscillator #4 after a 2nd difference was removed as an approximation to the frequency drift under the conditions of curve B of figure 8.

In all the oscillators examined, we achieved an improvement in the long-term performance by controlling the humidity and pressure. The random-walk-frequency-modulation component of the long term-frequency stability is reduced. The data suggests that the most important effect is the changes in humidity. It is interesting to speculate on the difference between oscillators #3 and #4. The drift of oscillator #3 is about 2 x $10^{-11}$/day. In order to observe a flicker level of 3 x $10^{-13}$ at $10^6$ s, the drift level would have to be constant to 0.3%. Whereas the drift only needs to be stable to 6% to achieve the same flicker level in oscillator #4.

## Conclusion

We have shown that the medium-term frequency stability of four different 5 MHz oscillators controlled by two different types of AT-cut resonators are sensitive to changes in atmospheric moisture and possibly pressure. Stabilizing the pressure and humidity, in addition to the normal parameters in an otherwise controlled laboratory setting, improves the frequency stability for measurement times in the region of hours to days. The mechanism by which fluctuations in moisture and possibly pressure transduce frequency changes are open to speculation. Perhaps the effect is due to changes in the thermal gradients or perhaps to changes in the dielectric constants or residual leakage between critical circuit elements. We also introduced a simple test to determine the relative importance of thermal gradients within air-enclosed ovens.

We have also shown that the frequency stability of one BVA-controlled oscillator was $\sigma_y(\tau)$ = 3 (± 1) x $10^{-13}$ for measurement times from 0.03 s to 21 days (measurement bandwidth of 100 Hz). These data and those on three other quartz-crystal-controlled oscillators indicate that the random-walk-frequency-modulation normally seen in the long-term-frequency stability data of quartz-crystal-controlled oscillators is not necessarily fundamental to the quartz, but is probably due to environmental influences. It also shows that the flicker-frequency modulation observed in quartz-crystal-controlled oscillators is not just a superposition of several noise processes over a narrow range of times, but rather represents a fundamental noise process within the oscillator. The flicker frequency level is most likely due to the quartz resonator. These data give an indication that quartz-crystal-controlled oscillators engineered to minimize the effects of changes in external parameters such as temperature, moisture, pressure, and load, may some day be used in areas which have traditionally been dominated by atomic frequency standards.

## Acknowledgements

I am very grateful to many colleagues over the years for enlightening discussions on the topic. I would especially like to mention Dr. J. A. Barnes, Dr. R. Besson, Dr. J. J. Gagnepain, Dr. J. Vig, C. Stone, and S. Bass, who suggested the wet sponge experiment and along with J. M. Cloeren very kindly shared some unpublished data on this topic.

## 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 Transactions on Instrumentation and Measurement, IM-24, 15-20 (1975).

2. J.J. Gagnepain, J. Ubersfeld, G. Goujon and P. Handel, "Relation Between 1/f Noise and Q Factor in Quartz Resonators at Room and Low Temperatures, First Theoretical Interpretation," In Proc. of the 35th Annual Symposium on Frequency Control, 476-483 (1981).

3. J.J. Gagnepain, "Fundamental Noise Studies of Quartz Crystal Resonators, Proc. of the 30th Annual Symposium on Frequency Control, 84-92 (1976).

4. J.J. Gagnepain, M. Oliver and F.L. Walls, "Excess Noise in Quartz Crystal Resonators," in Proc. of the 37th Annual Symposium on Frequency Control, 218-225 (1983).

5. Y. Noguchi, Y. Teramachi and T. Musha, "1/f Frequency Fluctuation of a Quartz Crystal Oscillator, Proc. of the 35th Annual Symposium on Frequency Control, 484-491 (1981).

6. M. Planat and J.J. Gagnepain, "Temperature Dependence of 1/f Noise in Quartz Resonators in Relation with Acoustic Attenuation," Noise in Physical Systems and 1/f Noise, (Edited by A. D'Amico and P. Mazzetti, Elsevier Science Publishers B.V., 1986), pp. 323-326.

7. T.E. Parker, "Characteristics and Sources of Phase Noise in Stable Oscillators," Proc. of the 41st Annual Symposium on Frequency Control, 99-110 (1987).

8. R. Besson, "A New Piezoelectric Resonator Design," Proc. 30th Symposium on Frequency Control, 78-83 (1976).

42nd Annual Frequency Control Symposium - 1988

# TEMPERATURE COMPENSATION OF SAW OSCILLATORS

M P Cracknell
A M Harrison
D J Sharpe

STC QUARTZ CRYSTAL DIVISION

Harlow, Essex, 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 compensation of AT-cut bulk crystal oscillators. The benefits that result for SAW oscillators are high compensation performance, low power consumption, low complexity and relatively low cost. Compensation is achieved by the use of a custom bipolar integrated circuit which generates a power series representation of the control voltage, of a voltage controlled SAW oscillator. A prototype oscillator has been built at 600 MHz and shows a compensated performance of ±4ppm over a 120 deg.C operating temperature range. The potential to improve the performance to ±2ppm exists.

## Introduction

Over the years there has been a continuing demand for oscillators of increased stability, higher frequency, reduced size, lower power consumption and good phase noise. To meet the demand for high frequencies, surface acoustic wave (SAW) resonators and delay lines have been used to produce oscillators to well over 1GHz. 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 device can be optimised by the selection of angle to give an overband performance of ±15ppm over -40deg.C to +80deg.C, a similar SAW device will achieve ±60ppm at best. Bulk wave oscillators have traditionally been given improved stability by the use of temperature compensation techniques. Compensation is achieved by the use of an oscillator whose frequency can be varied by the application of a d.c. voltage combined with a circuit which generates a voltage that varies with temperature. This voltage varies in such a way as to maintain the oscillators frequency at a constant value throughout the operating temperature range.

SAW resonators and delay lines can be used to produce voltage controlled oscillators with linear pulling characteristics. Such devices are ideal for temperature compensation. Various schemes for compensation have been proposed and show that compensation to better than ±10ppm 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

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:

$$T_0(x) = 1$$
$$T_1(x) = x$$
$$T_2(x) = 2x^2 - 1$$
$$T_3(x) = 4x^3 - 3x$$
$$T_4(x) = 8x^4 - 8x^2 + 1$$

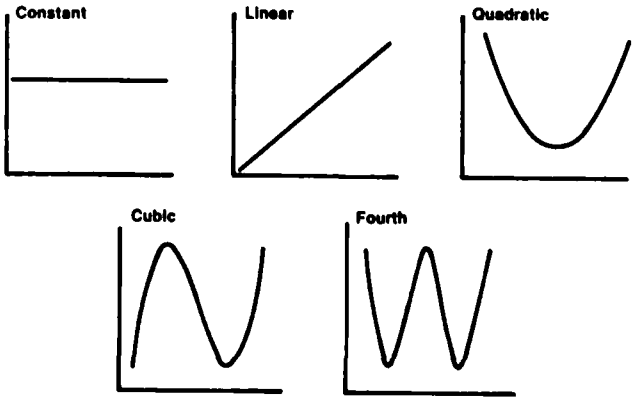



Fig.1. IC VOLTAGES

CH2588-2/88/0000-284 $1.00 C 1988 IEEE

The desired compensation voltage will be of the form:-

$$V(t) = a\ To(x) + a_1 T1(x) + a_2 T2(x) + a_3 T3(x) + a_4 T4(x)$$

For a bulk wave crystal oscillator using AT cut crystals the dominant term will be T3(x). With a SAW device the dominant term will be T2(x).

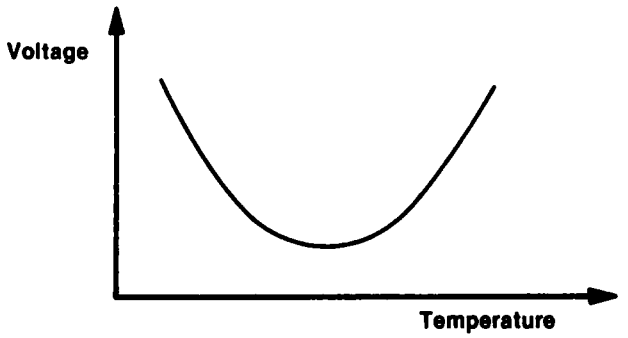



Fig.2. COMPENSATION VOLTAGE

A block diagram of the elements of the IC is shown in figure 3. The constant voltage is generated using a voltage stabiliser and generates a temperature invariant voltage using a bandgap reference. The linearly-varying voltage with temperature is generated using the temperature coefficient of the base-emitter junction voltage of a transistor, operated at constant emitter current. Conventional analogue multipliers comprising four transistors with the collectors cross-coupled are used to generate voltages with quadratic, cubic and fourth order variations with temperature. The summing amplifier is included in the integrated circuit.

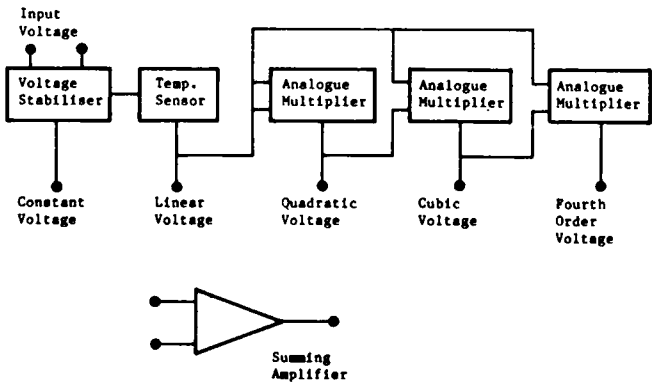



Fig.3. INTEGRATED CIRCUIT SCHEMATIC

## 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 the coefficients To to T4. 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 can be found in.[2]

With the 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.4.

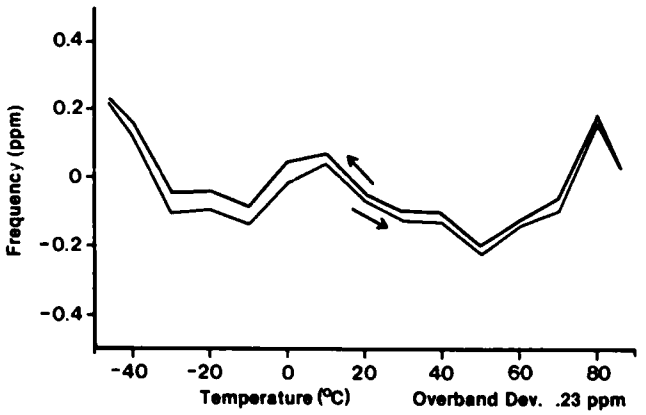



Fig.4. COMPENSATED 5 MHz OSCILLATOR

## SAW VCO

STC has developed a range of SAW VCOs around 600 MHz with a pulling capability of up to ±350ppm for a control voltage range of 0 to 12v. The graph of this frequency variation (fig. 5) shows the good, linear response achieved. The overband performance of the oscillator (fig.6) shows the typical SAW parabolic characteristic which has been slightly degraded by the oscillator circuitry to give a 150ppm maximum variation over -40deg.C to +80deg.C.

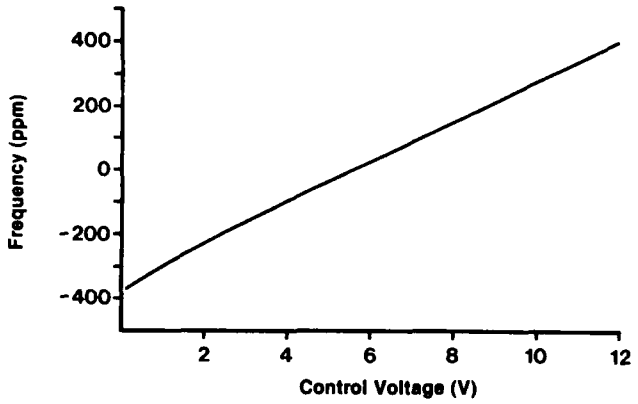



Fig.5. 600 MHz SAW VCO PULLING CHARACTERISTIC

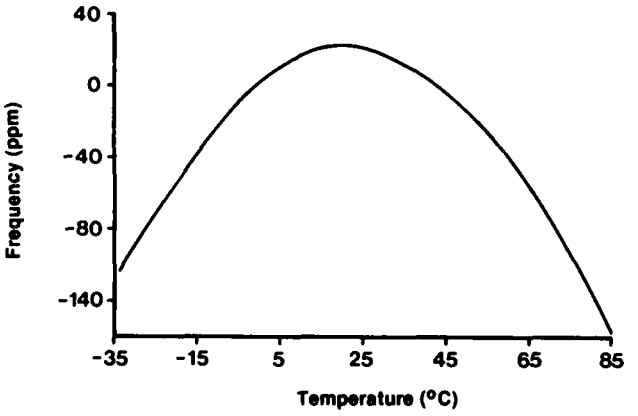



Fig.6. 600 MHz SAW VCO TEMPERATURE PERFORMANCE

This VCO circuit uses a delay line type SAW in a circuit with a phase shifting network consisting of three varactor diodes with associated inductors and capacitors. The use of three varactors was necessary to achieve the very wide pulling and linearity that was required for a particular application. In a VCO designed specifically for compensation the pulling could be considerably reduced and only one varactor should be necessary. This will lead to considerable simplification of the circuit construction and reduce size and cost.

The output from this VCO is a sinewave having an amplitude of 0dBm into a 50 ohm load. The device requires a current of approximately 65 mA from a 12 volt supply.

## SAW Oscillator Compensation

The results obtained from the development 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 the characteristics of the oscillator were combined with those of a typical IC and a theoretical prediction made. This result is shown in fig.7 and indicates that a frequency tolerance of better than ±2ppm may be achieved.

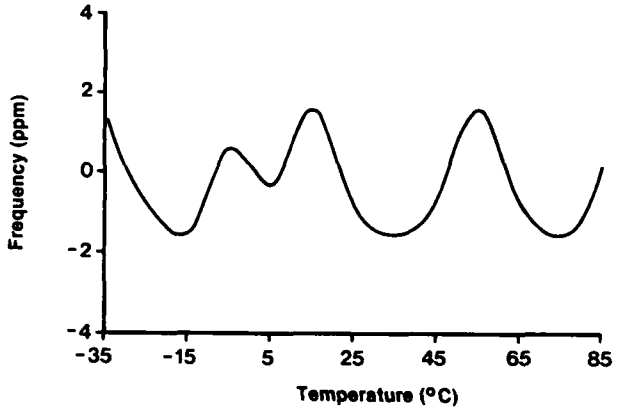



Fig.7. 600 MHz OSCILLATOR THEORETICAL COMPENSATION

One of the SAW oscillators was then built into a circuit with an IC 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 are ideal. The result is shown in fig.8 giving a compensation stability prediction of just over ±2ppm. The resistors were then fitted and a practical performance of ±4ppm as shown in fig.9 was achieved. It is expected to be able to bring this result nearer to the compensation result by slight variation of the resistor values to compensate for resistor tolerances and the loading effects of the resistors when they are fitted into the circuit. This technique is already applied to bulk wave devices.

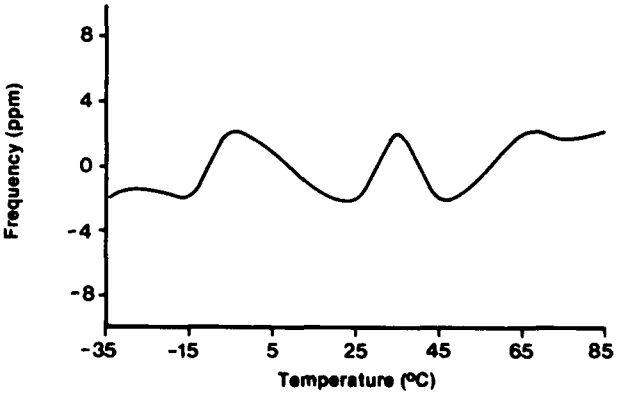



Fig. 8. 600 MHz OSCILLATOR COMPENSATION RESULT

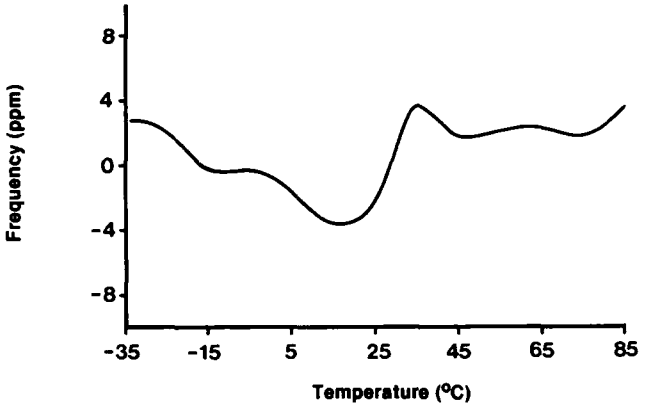



Fig.9. 600 MHz COMPENSATED OSCILLATOR; PRACTICAL RESULT

## CONCLUSIONS

It has been shown that it is possible to apply this compensation technique, developed for bulk wave crystal oscillators, to much higher frequency SAW oscillators.

The oscillator circuit used was not optimum for this application and improvements can be expected from a redesign to reduce the pulling sensitivity and thus increase the inherent stability of the device.

From this initial work it is predicted that this technique will provide highly stable SAW oscillators in a frequency range of 400 MHz to above 1GHz with a frequency variation of less than $\pm$3ppm over a temperature band of 120deg.C; the devices being compact, having low phase noise and requiring minimal additional power consumption (<1mA).

## 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 of 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.Motamedi "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 Untrasonics 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/SSBW Devices" Proceedings of the 1978 Ultrasonic Symposium. p.474-477.

42nd Annual Frequency Control Symposium - 1988

# Frequency Stability of SAW Oscillators at High Temperatures

Bikash K. Sinha*, Joel L. Groves**, Yukio Sudo*, and Shigeru Sato*

* Nippon Schlumberger K.K., 2-2-1, Fuchinobe, Sagamihara-shi, Kanagawa-ken 229, Japan
** Schlumberger-Doll Research, Old Quarry Road, Ridgefield, CT 06877-4108, U.S.A.

## ABSTRACT

This report is a broad overview of various factors affecting the frequency stability of Surface Acoustic Wave (SAW) resonators. The objective of this study has been to unravel the causes of large frequency drifts observed in SAW resonators operating at temperatures of 135°C and beyond. Two major causes of rapid degradation in the long-term frequency stability are found to be the presence of a chromium interface between aluminum and quartz, and moderate to high drive levels in SAW devices with pure aluminum fingers resulting in metal migration into quartz. On the other hand, devices with copper-doped aluminum electrodes maintained excellent long-term stability even at moderately high drive levels.

Experimental data are presented on both the long-term and short-term frequency stabilities of SAW devices at 25°C and 175°C for moderate to high drive levels. Results for the frequency stabilities compare the behavior of SAW devices with pure aluminum and copper-doped aluminum electrodes. In particular, it is shown that the short-term frequency stability of SAW devices with copper-doped aluminum electrodes is measured to be a few parts in $10^{10}$ even at 175°C and for moderately large drive levels. Overall, the best short-term frequency stability is found to be for a gate time of 0.1 sec.

## 1. INTRODUCTION

The medium-term and long-term frequency instabilities in Surface Acoustic Wave (SAW) oscillators are caused by many diverse factors, such as the selection of the crystalline material quality, device design, fabrication, sealing procedures, annealing or pre-aging of the sealed devices, drive level, and operating conditions. Not only are the contributing factors to the SAW oscillator instability numerous, they also lead to a wide variety of effects on the frequency output. For instance, some of the aging mechanisms result in random frequency fluctuations (phase noise), whereas others may lead to systematic, positive or negative frequency drifts. Some of the sources of frequency instabilities are obvious, others are more subtle. Thus, it is not surprising that isolating the various causes of the observed drift in SAW oscillators remains an active area of current effort.

Nevertheless, the long-term stability of SAW resonators has been significantly improved in the recent past. Improvements in the long-term stability from a few ppm to a few-tenths of a ppm have been obtained by subjecting these devices to a good bake-out, avoiding contaminations, and controlling the drive-level etc. [1] - [6]. In order to continue to improve the long-term frequency stability of SAW oscillators, it is essential to understand the role of each mechanism and obtain a quantitative estimate of its contribution to the overall frequency instability.

Latham et al. [1] have provided the first direct evidence of aluminum migration into quartz at moderate drive levels and strikingly large negative frequency drift in SAW oscillators. They have also provided a simple phenomenological relationship between the power dissipation in the SAW resonator and the magnitude of stresses in the IDT fingers.

The first set of measurements attempts to illuminate the role of chromium used as an interface film between aluminum IDT fingers and quartz substrate. This experiment also provides information on the rate of relaxation of the residual or intrinsic stresses in the polished quartz surface as well as in the IDT fingers and the influence of the relaxation rate on the frequency drift.

The second experimental investigation is directed toward understanding the effect of aluminum migration into quartz substrate on the long-term drift of the SAW resonator frequency and the accompanying degradation in the SAW resonator insertion loss and the resonator Q. The objective of this study is to correlate the damage observed in the IDT fingers with the power dissipation in the SAW resonator and the stress magnitudes on the free surface as well as in the IDT film.

The third set of measurements investigates the role of IDT metallization (aluminum on chromium; pure aluminum; and copper-doped aluminum electrodes) on both the long-term and short-term frequency stabilities of SAW oscillators at 25°C and 175°C for moderate to high drive levels.

## 2. THIN FILM STRESSES DUE TO PROPAGATING SURFACE WAVES

The accompanying stresses in the thin film due to the surface wave propagating in a plated substrate are not the same as those in the substrate. Since only the continuity of normal components of the stresses at the film and substrate interface need to be ensured, it is quite common in such plated structures to have a sharp discontinuity in the extensional (or longitudinal) stress components at the boundary between the film and substrate. Thus, it is of interest to estimate extensional stresses in the film for a given power flux carried by the propagating surface wave. The long-term stability of the film, along with any possible aluminum migration into quartz, may very well be dependent on a stress threshold in the film at a given operating temperature.

Furthermore, it is important to understand the role of film material properties and the film thickness on the magnitude of stresses developed by the propagating or standing waves. This information will then enable us to define a stress

CH2588-2/88/0000-288 $1.00 C 1988 IEEE

threshold in the film below which stress induced aluminum migration into quartz will not occur for a given amount of power dissipation in the SAW resonator, or equivalently, in terms of the time average power flux carried by the traveling wave.

2.1 SAW Resonator Equivalent Circuit

In order to estimate the power dissipation in the SAW resonator, it is helpful to determine the equivalent circuit parameters. The equivalent circuit for a two-port SAW resonator is shown in Figure 1. At resonance, the following relationships among the circuit parameters are valid:

$$\text{Insertion loss (I.L.)} = -20 \log\left(1 + \frac{R_1}{R_S + R_1}\right), \quad (1)$$

$$\text{Insertion loss (I.L.)} = -20 \log\left(\frac{1}{1 - Q_l/Q_u}\right), \quad (2)$$

$$Q_l = \frac{f_0}{\Delta f}, \quad Q_u = \frac{\omega L_1}{R_1}, \quad (3)$$

$$\frac{\Delta \phi}{\Delta \omega} = \frac{2 Q_l}{\omega}, \quad (4)$$

where $R_S$ and $R_l$ are, respectively, the source and load impedances while measuring the insertion loss on a network analyzer ($R_S = R_l = 50$ ohm, in our measurement set-up); $R_1$, $L_1$, and $C_1$ are, respectively, motional resistance, inductance, and capacitance, and $\omega$ is the resonant frequency in radians per sec. In Eq. (3), $\Delta f$ is the -3 dB bandwidth, and $f_0$ is the resonant frequency in Hz. $Q_l$ and $Q_u$ are the loaded and unloaded acoustic quality factors of the SAW resonator. In Eq. (4), $\Delta\phi/\Delta\omega$ is the transmission phase slope at the resonant frequency $\omega$.

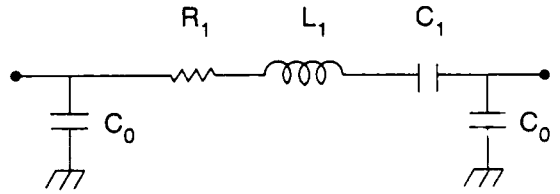



Figure 1. Equivalent circuit of a two-port SAW resonator.

Listed below are the equivalent circuit parameters for the SAW device Q2309 employed in one of the experimental studies.

$f_0 = 187.665705$ MHz, I.L. = -6.76 dB,

$C_0 = 2.5$ pF, $R_1 = 117.8$ ohms,

$L_1 = 5.04$ mH, $C_1 = 1.427 \times 10^{-6}$ F,

$Q_l = 27{,}297$ (in a 50 ohm system), and $Q_u = 50{,}475$.

The SAW resonator input impedance $Z_{11}$, and the output impedance $Z_{22}$ were calculated from $S_{11}$ and $S_{22}$ measurements, respectively, and the results are given below:

$Z_{11} = 133.67 - j\,65.99$ ohms, and

$Z_{22} = 129.47 - j\,62.92$ ohms.

The time average stored power in a SAW resonator can be expressed in terms of the unloaded quality factor $Q_u$ and power dissipation $P_d$ in the SAW device. This relationship is given by

$$P_S = f\,U_S = \frac{Q_u\,P_d}{2\pi}, \quad (5)$$

where $P_S$ is the stored power in the SAW device, f is the frequency, $U_S$ is the stored energy per cycle, and $P_d$ is the time average power dissipation in the device.

A standing wave in the resonant cavity can be considered to be the superposition of two oppositely traveling waves. Therefore, each of these waves will carry half of the total stored power.

2.2 Surface Wave Amplitude and Stresses from the Power Dissipation

2.2.1 Free Substrate

The amplitude of the displacement components at the free surface can then be solved for from the surface wave solution and a given time average, power flux per unit width of the transducer aperture associated with the propagating surface wave.

The results for such amplitudes of the displacement components for the ST-cut ($\Theta = -42.75°$) are given below:

$$|u_1|_{X_2=0} = 2.7899 \times 10^{-6} (P_1/\omega)^{1/2},$$

$$|u_2|_{X_2=0} = 4.2406 \times 10^{-6} (P_1/\omega)^{1/2},$$

$$|u_3|_{X_2=0} = 0.4050 \times 10^{-6} (P_1/\omega)^{1/2}, \quad (6)$$

where $P_1$ is the magnitude of the time average power flux in W/m, and $\omega$ is the surface wave frequency in rad/sec.

Given the amplitude of the surface wave displacement vector, it is quite straight-forward to compute the magnitude of stresses parallel as well as normal to the surface wave propagation direction. For ST-cut quartz, such stress magnitudes are given by

$$|T_{11}|_{X_2=0} = 76.46\,(P_1\omega)^{1/2} \ \text{N/m}^2, \text{ and}$$

$$|T_{33}|_{X_2=0} = 23.59\,(P_1\omega)^{1/2} \ \text{N/m}^2. \quad (7)$$

2.2.2 Plated Substrate

For ST-cut quartz substrate plated with a thin aluminum film ($t_f\xi$ = 0.1, where $t_f$ is the film thickness, and $\xi$ is the wavenumber), we find that the stresses parallel and normal to the propagation direction, $T_{11}$ and $T_{33}$, respectively, are given by

$$|T_{11}|_{X_2=0}(\text{in the substrate}) = 69.64\ (P_1\omega)^{1/2},$$

$$|T_{33}|_{X_2=0}(\text{in the substrate}) = 21.67\ (P_1\omega)^{1/2}. \qquad (8)$$

From the standpoint of stress induced migration of aluminum into quartz substrate, it is more important to estimate the stress magnitudes in the aluminum film (of the IDT fingers) due to the propagating surface waves in the plated substrate. Such accompanying stresses in the film can be determined from the constitutive relations for the film material, provided the relevant strain components in the film are known.

For ST-cut quartz substrate plated with an aluminum film ($t_f\xi$ = 0.1), the extensional stress in the film parallel to the propagation direction can be calculated from Eq. (9).

$$|T_{11}^f|(\text{in the aluminum film}) = 62.9\ (P_1/\omega)^{1/2} \qquad (9)$$

Comparing Eqs. (8) and (9), it becomes obvious that the longitudinal stress $T_{11}$, under a thin film, and in the film are somewhat different from each other for a given magnitude of time average power flux carried by a propagating surface wave at any selected frequency.

## 3. ESTIMATES OF THIN FILM STRESSES FROM MEASURED CIRCUIT PARAMETERS

### 3.1 Power Dissipation versus Drive Voltage

The power dissipation in a SAW resonator placed in an oscillator circuit can be estimated as a function of the (d.c.) drive level of the amplifier. The most reliable approach [2] for the estimation of the power dissipation in a SAW resonator consists of essentially the following two steps:

First, we estimate the current through the SAW device as a function of various drive levels. This is, generally, a critical and not very precise estimate. One possible reason for the uncertainty is that the oscillation frequency may not be at the exact resonant frequency where the equivalent circuit parameters are valid.

Second, we obtain the resistance of the SAW device at resonance from the equivalent circuit parameters. The magnitude of the current together with the resistance readily yields the power dissipation in the device.

We have measured the voltage at the input (A) and output (B) of the SAW resonator (as shown in Figure 2a) with Tektronics FET probe (P6202A) on a spectrum analyzer for different drive levels. The input impedance of the probe is $10^7$+/-2% ohm and the probe capacitance is approximately 2 pF (d.c. to 500 MHz). A difference between these voltages (RMS) provided the voltage drop across the SAW resonator which together with the known value of the device resistance yields the power dissipation in the SAW resonator. Table I contains the input and output voltage measurements converted into RMS values.

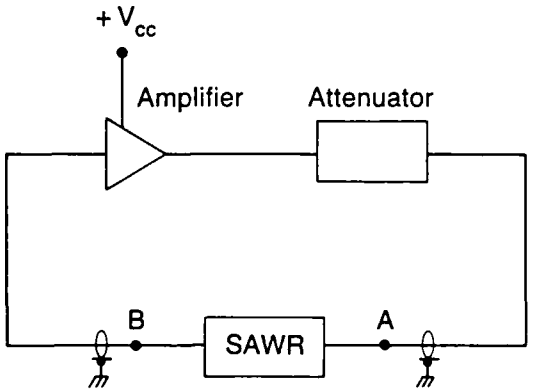



Figure 2a. SAW resonator in an oscillator circuit.

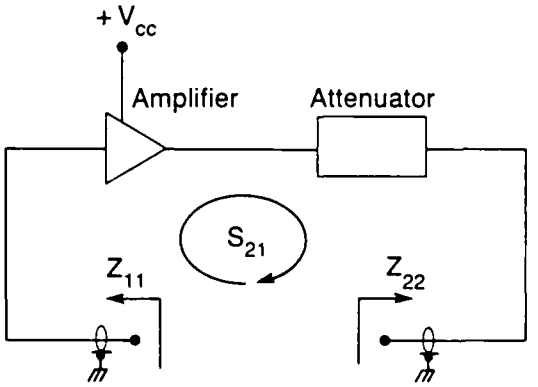



Figure 2b. Input and output impedances of the amplifier, attenuator, and coax.

TABLE I: Input and Output Voltage Measurements at the SAW Resonator (Q2309)

| $V_{cc}$ volts (d.c.) | Input voltage (A) volts (RMS) | Output voltage (B) volts (RMS) |
|---|---|---|
| 5 | 0.309 | 0.062 |
| 6 | 0.447 | 0.088 |
| 8 | 0.646 | 0.135 |
| 10 | 0.813 | 0.174 |
| 12 | 1.059 | 0.204 |
| 15 | 1.303 | 0.26 |

However, the presence of the FET probe in the oscillator loop may affect the RF measurements in an unpredictable way. Therefore, we have also measured the input and output impedances of the loop amplifier, coax cable, and attenuator which are integral parts of the oscillator loop (Figure 2b).

The measurements were performed on a network analyzer with an input power of -13 dBm. The results of these measurements are shown in Table II:

TABLE II: Input and Output Impedances of the amplifier, coax cable, and attenuator

| $V_{cc}$ | $Z_{11}$ | $Z_{22}$ | $S_{21}$ | |
|---|---|---|---|---|
| volts (d.c.) | ohms | ohms | $\|S_{21}\|$ dB | $\angle S_{21}$ deg |
| 6 | 36.96 + j2.1 | 34.97 + j12.1 | 9.64 | 9.61 |
| 8 | 37.36 + j1.4 | 35.41 + j11.65 | 10.17 | 11.19 |
| 10 | 37.7 + j0.96 | 35.8 + j11.37 | 10.47 | 12.29 |
| 12 | 37.73 + j0.66 | 36.08 + j11.0 | 10.67 | 12.97 |
| 15 | 37.74 + j0.52 | 36.51 + j10.6 | 10.85 | 13.49 |

In order to investigate the reliability of the estimate of the power dissipation in the SAW resonator, we have calculated this quantity three different ways: First, we obtained the power dissipation in the device from the voltage drop across the device and its resistance at resonance from the equivalent circuit. Second, we calculated current through the SAW resonator from the equivalent circuit shown in Figure 3, and using the measured value of the input voltage at the SAW device, and the input impedance of the amplifier, coax cable, and attenuator. The third procedure was to use the output voltage measurement at the SAW device in conjunction with the equivalent circuit for the other components in the loop as shown in Figure 4.

A comparison of the results for the power dissipation in the SAW resonator from the three procedures are shown in Table III.

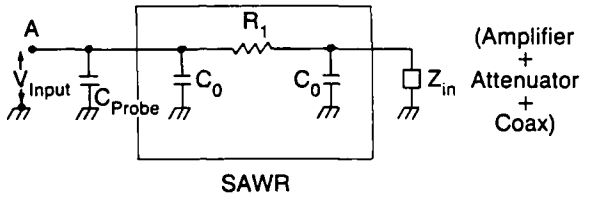



Figure 3. Network representing voltage measurement at the SAW resonator input.

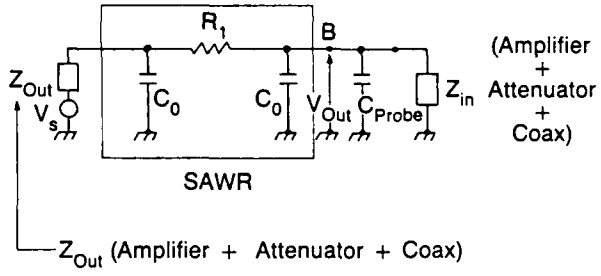



Figure 4. Network representing voltage measurement at the SAW resonator output.

TABLE III: Power dissipation in the SAW Resonator

| $V_{cc}$ volts (dc) | $V_d$ volts (RMS) | $P_d$(I) watts (RMS) | $P_d$(II) watts (RMS) | $P_d$(III) watts (RMS) |
|---|---|---|---|---|
| 5 | 0.247 | $0.516\times10^{-3}$ | --- | --- |
| 6 | 0.359 | $1.092\times10^{-3}$ | $0.976\times10^{-3}$ | $0.816\times10^{-3}$ |
| 8 | 0.511 | $2.215\times10^{-3}$ | $2.043\times10^{-3}$ | $1.614\times10^{-3}$ |
| 10 | 0.639 | $3.466\times10^{-3}$ | $3.228\times10^{-3}$ | $2.507\times10^{-3}$ |
| 15 | 1.043 | $9.235\times10^{-3}$ | $8.30\times10^{-3}$ | $6.799\times10^{-3}$ |

We note that the results for the power dissipation from procedures II and III are, respectively, lower by approximately 10 and 26% than that of the procedure I. Exactly why these differences occur in the estimate of the power dissipation obtained by somewhat different procedures is not yet completely understood. However, it is possible that the assumed equivalent circuit for the SAW resonator is not strictly valid at the oscillating frequency which may very well be somewhat different from the true resonant frequency. The placement of the probe capacitance in the network also alters the estimate of the power dissipation, albeit by a small amount compared with the relatively large difference seen in the estimate from the methods I, II and III. In any event, either of the aforementioned procedures is adequate to estimate the power dissipation with an uncertainty of approximately 10 to 25%. We have employed the first procedure (I) for the estimation of the power dissipation in the SAW resonator.

### 3.2 Surface Wave Displacement Amplitude versus Drive Voltage

Clearly, it is of considerable interest to ascertain the onset of non-linearity in the SAW resonator which will be manifested in terms of amplitude-frequency effect or harmonic generation. Harmonic generation in surface waves lead to a depletion of energy in the fundamental mode and distortion of the waveform which must be avoided in high performance resonators.

The surface wave displacement amplitude can be readily calculated from Eqs. (6) in terms of the time average power flux carried by the surface wave and its frequency. Since we can assume one-half of the stored power in the standing wave to be carried by the oppositely traveling waves, and that the total surface displacement amplitude in the standing wave will be twice of the amplitude due to each of the traveling wave, it is straight-forward to compute the surface displacement amplitude for various drive levels. Subsequent measurements on the SAW device Q2309 were made after re-connecting a lead-wire which was broken during handling. Consequently, some changes may have occurred in the relationship between the supply voltage $V_{cc}$ and power dissipation $P_d$. Table IV summarizes the results for the normal displacement component amplitude for different drive levels.

TABLE IV: Drive Level Dependence of the Surface Wave Amplitude $|u_2|$ (Q2309):

| $V_{cc}$ volts | $P_d/W$ mW/mm | $P_s/W$ mW/mm | $\lvert u_2 \rvert_{X_2=0}$ Å | $\lambda/\lvert u_2 \rvert$ |
|---|---|---|---|---|
| 4 | $7.342 \times 10^{-2}$ | $5.898 \times 10^{2}$ | 42 | 4000 |
| 5 | $1.956 \times 10^{-1}$ | $1.571 \times 10^{3}$ | 69 | 2435 |
| 10 | 2.340 | $1.880 \times 10^{4}$ | 239 | 703 |
| 15 | 5.457 | $4.385 \times 10^{4}$ | 366 | 459 |

### 3.3 Sinusoidal Stresses versus Drive Voltage

Given the unloaded quality factor of the SAW resonator, one can compute the time average power stored in the resonator in terms of the power dissipation from Eq. (5). Since the standing wave in the resonator is equivalent to two oppositely traveling waves, one- half of the stored power can be assumed to be carried by either of the traveling waves. This power flux can be expressed in terms of per unit width of the straight-crested wave, provided the aperture of the SAW resonator is specified. The biaxial stresses $T_{11}$ and $T_{33}$ on the free surface are calculated from Eq. (7). The total stress magnitude due to the standing wave is twice of that due to each of the traveling waves.

The longitudinal stress $T_{11}$ in the plated region of the substrate, under the film, and in the film can be estimated from Eqs. (8) and (9), respectively.

Table V illustrates the stress magnitudes at the free surface and in the IDT film of the SAW resonator. These results have been obtained for the SAW resonator Q2309 whose geometrical aperture width (W) is 1.5 mm. In the calculation for the time average power flux, we have ignored any variation in the width direction because of the apodization of the IDT fingers. Nevertheless, the apodized IDT was employed in these devices in order to suppress higher-order transverse modes.

Computational results for the SAW resonator employed in Ref. [2], are also included at the end of Table V for the sake of reference.

TABLE V: Amplitude of sinusoidal stresses in SAW resonators (25°C)

| $V_{cc}$ volts | $P_d$ mW | $P_d/W$ mW/mm | $P_s/W$ mW/mm | $\lvert T_{11} \rvert$ N/m$^2$ | $\lvert T_{33} \rvert$ N/m$^2$ | $\lvert T_{11}^f \rvert$ N/m$^2$ |
|---|---|---|---|---|---|---|
| (Device Q2309: $Q_u$ = 50,475, $f_0$ = 187.66 MHz, Aperture W = 1.5 mm) | | | | | | |
| 4 | 0.110 | $7.342 \times 10^{-2}$ | $5.898 \times 10^{2}$ | $9.02 \times 10^{7}$ | $2.78 \times 10^{7}$ | $7.42 \times 10^{7}$ |
| 5 | 0.293 | $1.956 \times 10^{-1}$ | $1.571 \times 10^{3}$ | $1.47 \times 10^{8}$ | $4.54 \times 10^{7}$ | $1.21 \times 10^{8}$ |
| 10 | 3.510 | 2.340 | $1.880 \times 10^{4}$ | $5.09 \times 10^{8}$ | $1.57 \times 10^{8}$ | $4.19 \times 10^{8}$ |
| 15 | 8.186 | 5.457 | $4.385 \times 10^{4}$ | $7.78 \times 10^{8}$ | $2.40 \times 10^{8}$ | $6.40 \times 10^{8}$ |
| Shreve et al. [2] device: $Q_u$ = 43,000, $f_0$ = 160 MHz, Aperture W = 1.5 mm | | | | | | |
| | 0.2 (-7 dBm) | 0.133 | 910.207 | $1.03 \times 10^{8}$ | $0.319 \times 10^{8}$ | - |
| | 0.8 (-1 dBm) | 0.533 | 3647.672 | $2.07 \times 10^{8}$ | $0.639 \times 10^{8}$ | - |

## 4. MEASUREMENTS OF THE FREQUENCY SHIFT OF SAW OSCILLATORS AS A FUNCTION OF THE DRIVE VOLTAGE

The drive stability of SAW oscillators is an important consideration for many applications. The objective of this experimental investigation was to study the role of the SAW resonator and the amplifier, coax cable, and attenuator in the oscillator loop on the overall stability of the frequency output as a function of the supply voltage to the amplifier.

The SAW resonator Q2309 was employed together with an amplifier, attenuator, and coax cable in an oscillator circuit. The drive voltage was varied from 4 to 15 volts d.c. The resulting change in the phase shift in the components other than the SAW resonator (i.e., amplifier, attenuator, and coax cable) was measured on a network analyzer (HP 8753A) with an input power of -33 dBm. The transmission gain in the amplifier, and the input and output impedances were also measured.

The contribution of these oscillator components to the oscillator instability can be estimated from Eq. (4). Subtraction of the fractional change in the oscillator frequency from Eq. (4) from the measured fractional change in the oscillator frequency yields the dependence of the SAW resonator frequency on the drive level.

Table VI summarizes the results for the drive level dependence of the amplifier (plus the coax cable and attenuator in the oscillator loop) phase shift and the oscillator frequency output. The contribution of the oscillator components phase shift to the oscillator frequency shift has been calculated from Eq. (4) for a measured value of $Q_l$ = 25,000 at the reference frequency of approximately 187.667 MHz.

TABLE VI: Drive Level Dependence of the SAW Oscillator Frequency (Q2309)

| Vcc | Gain | $\frac{\Delta\phi}{\Delta V_{cc}}$ | $\frac{\Delta\phi}{2\,Q_l \Delta V_{cc}}$ | $\frac{\Delta f}{f\,\Delta V_{cc}}$ | $\frac{\Delta f}{f\,\Delta V_{cc}}$ |
|---|---|---|---|---|---|
| volt | dB | rad/volt | (Amplifier) ppm/volt | (Resonator) ppm/volt | (Oscillator) ppm/volt |
| 4 | 8.836 | $2.079\times10^{-2}$ | 0.416 | -0.192 | 0.224 |
| 5 | 9.394 | $1.639\times10^{-2}$ | 0.328 | -0.077 | 0.250 |
| 10 | 10.494 | $0.606\times10^{-2}$ | 0.121 | 0.220 | 0.341 |
| 12 | 10.679 | $0.304\times10^{-2}$ | 0.061 | 0.300 | 0.361 |

These results raise some very interesting issues. First, looking at the sensitivity of the oscillator frequency to the drive voltage fluctuation, it appears that the oscillator frequency stability will be on the order of 0.01 ppm to 0.04 ppm for a voltage supply regulated to within 1% of its output. Overall, better frequency stability will be obtained at lower supply voltages.

However, the results also indicate that the instability in the oscillator frequency output due to the supply voltage fluctuation is caused by both the SAW resonator and amplifier phase shifts. Another noteworthy result is that while the amplifier phase stability is better at higher supply voltages, the SAW resonator frequency appears to be more stable at moderate voltages, and not necessarily at the lowest supply voltages. The increase in the instability of SAW resonator frequency at high supply voltages may very well be due to the slowly increasing non-linearity in the resonator performance.

## 5. MEASUREMENTS OF THE STABILITY OF SAW DEVICES AT HIGH TEMPERATURES

### 5.1 Devices with Pure Aluminum and Aluminum on Chromium Electrodes

Here we describe our experimental results from a study conducted to pinpoint the causes of large drifts in SAW devices at high temperatures. More specifically, our goal was to understand the role of the IDT metallization parameters, sealing and bonding agent, and the drive-level on the long-term stability of SAW resonators.

To this end, we employed four two-port SAW resonators around 187 MHz fabricated on rotated Y-cut quartz plates. The first two devices were sealed in "An All Quartz" package with devitrifying and vitrifying glass frits, respectively. Both of them had approximately 2800 Å of aluminum IDT fingers. The other two devices were not sealed in any package, but were placed in a baked, stainless steel chamber under high vacuum. The objective in the experiments with the last two devices was to remove any possible effect on the long-term drift at high temperature due to the bonding stress relaxation. The differences in the fabrication and sealing characteristics of these devices are listed in Table VII.

TABLE VII: Fabrication and Sealing Characteristics of SAW Devices

| | Device | Rotation Angle θ | IDT Metallization | Sealing |
|---|---|---|---|---|
| 1. | Q2487 | - 36°20' | 2800 Å aluminum | Devitrifying glass frits |
| 2. | Q2309 | - 36°57' | 2800 Å aluminum | Vitrifying glass frits |
| 3. | A1015H | - 43° | 1700 Å aluminum on 60 Å chromium | Baked, stainless steel, evacuated chamber ($10^{-7}$ Torr) |
| 4. | A1019H | - 43° | 1900 Å aluminum | Baked, stainless steel, evacuated chamber ($10^{-7}$ Torr) |
| 5. | S1001 | - 43° | 2800 Å aluminum on 60 Å chromium | Devitrifying glass frits |
| 6. | 2001 | - 33.5° | 2800 Å aluminum | Devitrifying glass frits-annealed |
| 7. | 2002 | - 34° | 2800 Å aluminum | Devitrifying glass frits-annealed |
| 8. | 3001 | - 34.2° | 2800 Å Cu-doped Al | Devitrifying glass frits-annealed |
| 9. | 3003 | - 34.2° | 2800 Å Cu-doped Al | Devitrifying glass frits-annealed |

The first two devices, sealed in an all quartz package, were subjected to temperatures ranging from 300 to 430°C for varying amount of time for the glazing and sealing processes. Some of the stresses in the IDT fingers as well as on the polished surface may have been annealed to some degree.

The other two SAW devices were placed in stainless steel chambers. These stainless steel chambers had previously been evacuated with an oil-free, high vacuum system and baked at 400°C overnight. After the baking was completed and the system returned to room temperature, the chambers were opened and the SAW devices were mounted inside. The electrical connections were made through vacuum-tight, high temperature feedthroughs. These were then evacuated to $10^{-7}$ Torr and pinch-sealed. These two SAW resonators were not annealed prior to this test.

The four devices were placed in a temperature controlled oven, and the electrical connections were brought out of the oven to a HP 3577A network analyzer for monitoring the electrical characteristics of the devices throughout the tests. The time-base of the network analyzer was controlled by a HP 5061A cesium beam standard. The temperature of the oven was monitored with a calibrated RTD (Resistance Temperature Device).

All of these devices were kept continuously in the oven for about 40 days at temperatures around 135°C. The electrical characteristics of these devices were measured every day using the network analyzer. However, none of these devices was operated in an oscillator circuit during this test.

Figures 5 through 8 display the changes in the resonant frequency of the four devices as a function of the annealing time. The changes in frequency due to temperature fluctuation of the oven have been partially removed. The temperature stability of the oven was on the order of 1°C, and the sensitivity of the RTD was approximately 0.01°C. However, frequency changes due to temperature gradients in the SAW resonator and to differences between the temperature of the SAW device and the RTD were not removed. The fractional change in the resonant frequency has been referenced to the initial frequency at the start of the annealing period.

During the first 1000 hours in the oven, all four devices exhibited a typical annealing behavior. The rate of change of the resonant frequency decreased with increasing annealing time. It is also clear from the results shown in Figures 5, 6 and 8 that these three SAW resonators appear to have stabilized. However, Figure 7 indicates that the third device continues to drift even after approximately 4000 hours. This difference in the behavior from the device in Figure 8 is solely attributable to the presence of chromium in between the quartz substrate and aluminum IDT fingers.

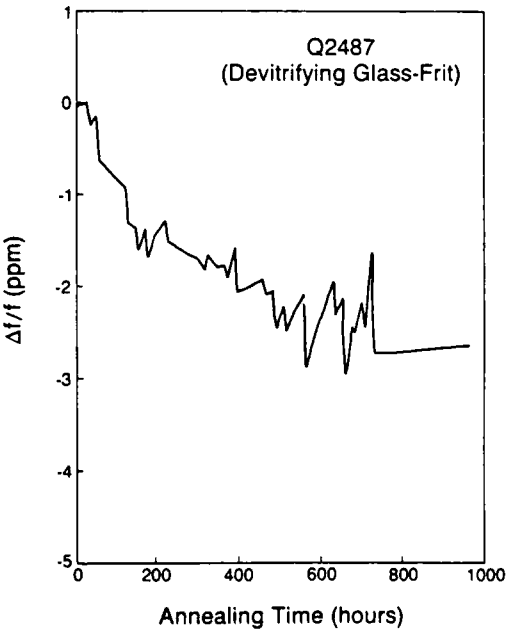



Figure 5. Annealing behavior of a glass-frit sealed SAW resonator Q2487 at approximately 135°C.

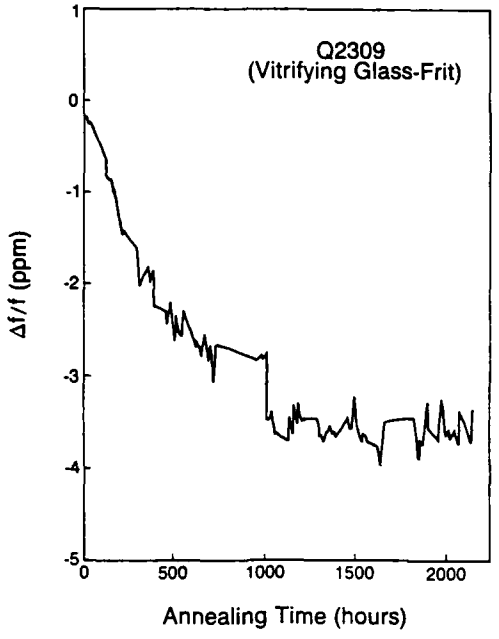



Figure 6. Annealing behavior of a glass-frit sealed SAW resonator Q2309 at approximately 135°C.

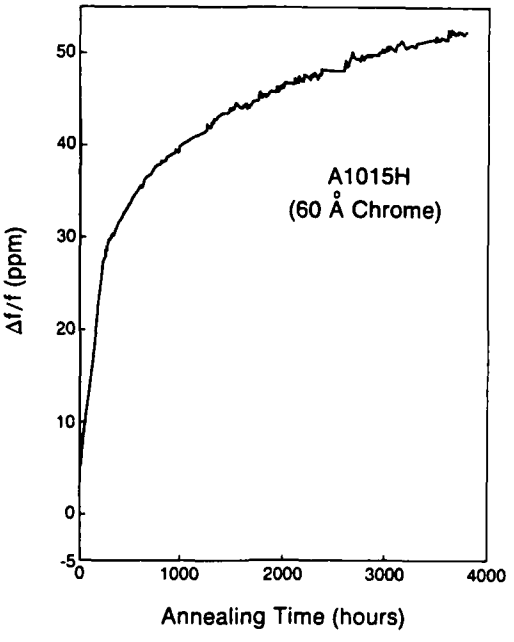



Figure 7. Annealing behavior of an unsealed SAW resonator A1015H placed in a baked, stainless steel, evacuated chamber. IDT fingers were of 1700 Å aluminum on 60 Å chromium. Oven temperature was approximately 135°C.

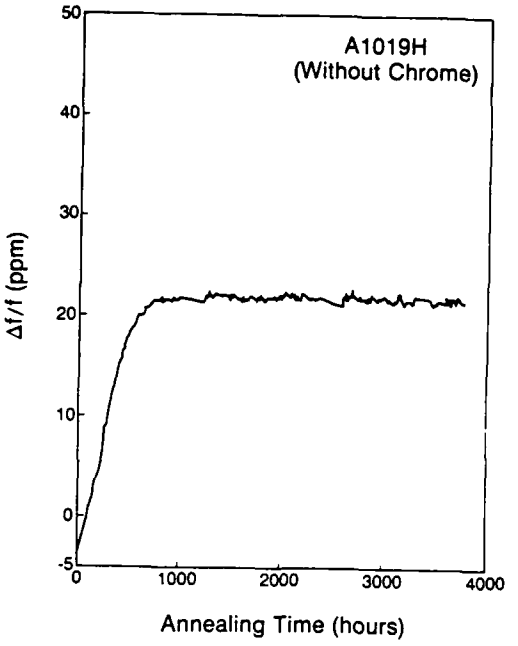



Figure 8. Annealing behavior of an unsealed SAW resonator A1019H placed in a baked, stainless steel, evacuated chamber. IDT fingers were of 1900 Å pure aluminum. Oven temperature was approximately 135°C.

What is most interesting to observe is that, while the first two glass-frit sealed SAW devices exhibited approximately 3 to 4 ppm downward shift in the resonant frequency, the other two unsealed devices showed approximately 25 to 50 ppm upward shift in the resonant frequency. One possible explanation for the relatively smaller shift in the resonant frequency during this annealing period may be the partial annealing of the intrinsic stresses in the SAW resonator that occurred during the high temperature glazing and sealing process. In addition, it appears that the glass-frit seal is essentially subjecting the SAW substrate to a compressive stress along the propagation direction which is parallel to the length of the plate. Since the thermal expansion coefficient of quartz along the length direction is larger than that of the glass-frit, an increase in the temperature of the composite structure will introduce compressive stress in quartz and tensile stress in the frit.

When the sealed device is kept at approximately 135°C in the oven, some relaxation of stress in the glass frit due to plastic deformation or creep resulted in a reduction of the biasing compressive stress along the propagation direction. Such a reduction in the compressive stress will cause a downward drift in the resonant frequency [7], [8] as observed in our experiment. A relatively small downward drift in the frequency of 3 to 4 ppm is indicative of a small magnitude of unstable, compressive stress (on the order of 3 to 4 x $10^5$ $N/m^2$) remaining as a result of glass-frit sealing.

As mentioned earlier, the other two unsealed devices (A1015H and A1019H) were not annealed at all during or after the fabrication process. The upward drift in the resonant frequency during annealing is similar to those observed by other workers [9] during their investigations of aging of SAW devices.

There are at least three potential factors which may have caused upward drift in the frequency during annealing. First, the relaxation of the intrinsic stress in the top damaged layer due to the polishing action may have caused part of the observed frequency drift. Second, Bray et al. [9] have reported that metal films are in a state of compressive stress after evaporation, which results in a state of tensile stress in the adjoining substrate. As a result, any relaxation of the film stress will decrease the magnitude of the biasing tensile stress in the substrate which will, for ST-cut devices, generally, cause an upward frequency drift [7], [8]. Third, any continuation of the oxidation of aluminum IDT fingers during annealing may also have caused an upward frequency drift.

Although much remains puzzling about the slow annealing rate observed for the A 1015H device, it is possible that the relatively thin chromium film (60 Å) is more highly stressed during evaporation and takes much longer to relax to a more stable state compared to thicker aluminum film on quartz.

After the first 1000 hours of annealing at approximately 135°C, the Q2487 SAW device was put into an oscillator circuit. The supply voltage was kept at 10 volts (d.c.). Even though this SAW resonator appeared to have been adequately annealed, a very large frequency drift occurred quite rapidly as soon as the oscillator was turned on and the drive level was kept high at 10 volts (d.c.) supplied to the amplifier. This permanent drift caused by the circulating power in the surface wave was always observed to be downward. Another feature observed in this test was that the frequency drift appeared to halt as soon as the oscillator was turned off. Figure 9 shows an overall view of the frequency changes that occurred in this SAW resonator when used passively as well as actively with a relatively large drive level.

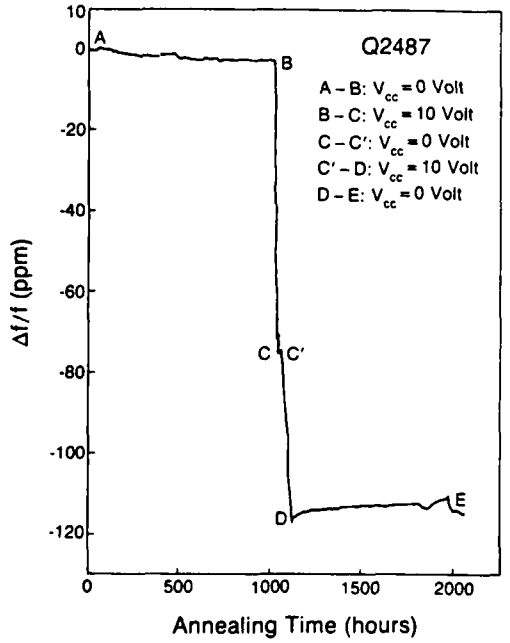



Figure 9. Long-term stability of the SAW resonator Q2487 during annealing at approximately 135°C; and during high drive-level in an oscillator circuit ($V_{CC}$ = 10 volts).

During the periods when this device was operated in an oscillator circuit, the resonator frequency measurements were made with the HP network analyzer by disconnecting the resonator from the oscillator loop. This was done to ensure that any possible phase shift in the amplifier is not responsible for the observed drift.

In order to confirm our suspicion that aluminum migration into quartz may have resulted in the large frequency drift observed for the Q2487 SAW resonator [1], [2], this device was opened for examination of the IDT area. The optical images of the IDT area were obtained and are shown in Figures 10 and 11. These images clearly show the damage that has occurred to the IDTs. This observation is exactly same as that observed by earlier workers [1], [2].

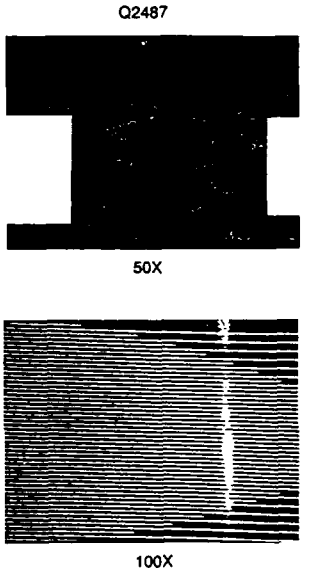



Figure 10. Optical image of the damaged IDT area of the SAW resonator Q2487 due to aluminum migration into quartz. Magnifications = 50 and 100.

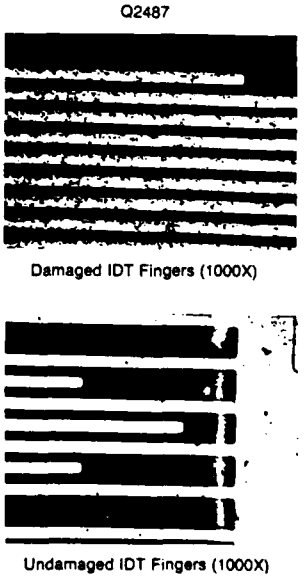



Figure 11. Optical image of the damaged IDT area of the SAW resonator Q2487 due to aluminum migration into quartz. Magnification = 1000.

5.2 Devices with Pure Aluminum or Copper-Doped Aluminum Electrodes

An experiment was performed to study the role of the IDT metallization in a two-port SAW resonator on both the drive voltage induced frequency "shift" and any possible oscillator frequency "drift" as a result of aluminum migration into quartz. The term frequency "shift" refers to a recoverable frequency change which occurs because of the change in the drive level. This phenomenon may be due in large part to the nonlinear, amplitude-frequency effect in SAW resonators. On the other hand the frequency "drift" refers to a non-recoverable frequency change which occurs because of a permanent change in the resonator structure, such as due to the aluminum migration into quartz.

Two SAW oscillators -- one with pure aluminum IDT fingers and the other with copper-doped aluminum fingers, were monitored for their frequency changes over many hours as the supply voltage to the amplifier was varied from 3 to 10 volts. Figure 12 shows the fractional change in the oscillator frequency vs. time at an operating temperature of 175°C. The contrast in the long-term frequency stability of these two oscillators is remarkable. The oscillator with the SAW resonator having pure aluminum IDT electrodes exhibited somewhat larger drive voltage induced frequency shift than that with the copper-doped aluminum IDT electrodes.

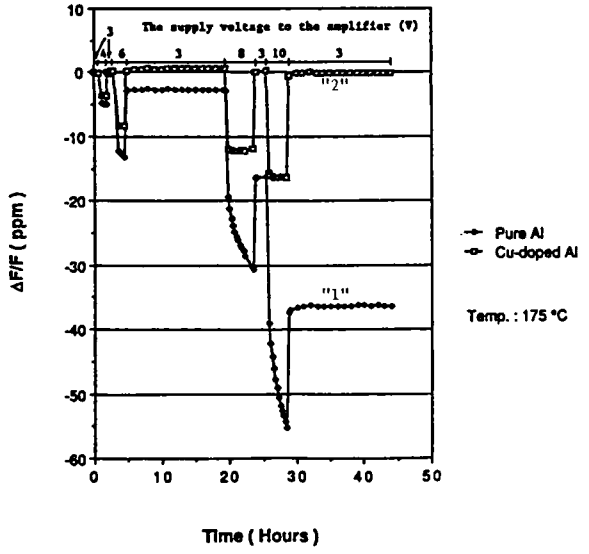



Figure 12. Long-term frequency stability of SAW oscillators at 175°C for various supply voltages. Curves 1 and 2 represent results for devices with pure aluminum (Device: 2002) and copper-doped aluminum (Device: 3003) electrodes, respectively.

The long-term frequency drift in the oscillator with pure aluminum electrodes is even more dramatic at high drive levels.

It is obvious from these results that the long-term stability of SAW oscillator with pure aluminum electrodes is good only when the supply voltage is limited to approximately 4 volts and any increase in the drive voltage leads to almost an instantaneous damage to the IDT area which is associated with a permanent frequency drift. This implies that the thin film stress threshold in the IDT fingers for pure aluminum electrodes is easily attained as the supply voltage is set to approximately 4 volts. On the other hand, the SAW oscillator with copper-doped aluminum IDT electrodes exhibited excellent long-term frequency stability even at high drive levels and operating at 175°C. The long-term stability over many hours was measured to be on the order of a few-tenths of a ppm.

## 6. SHORT-TERM FREQUENCY STABILITY OF SAW OSCILLATORS

The short-term frequency stability of SAW oscillators has been measured for SAW resonators fabricated with either aluminum on chrome, or pure aluminum, or copper-doped aluminum electrodes. All of the SAW resonators were of the same design with a nominal frequency of about 188 MHz as discussed in Sec. 2.1.

Figures 13a and 13b show the short-term frequency stability expressed as the Allan Variance ($\sigma_y(\tau)$), of an oscillator with drive voltage of 5 and 10 volts, respectively. The oscillator stability was controlled by a two-port SAW resonator which had recessed IDT fingers of 2800 Å on 60 Å chromium. Measurements were performed at 26°C in a temperature controlled environment and with various counter gate times. Calculations for the $\sigma_y(\tau)$ were performed with 100 samples. It is evident from these results that for this oscillator, the short-term frequency stability is approaching $10^{-10}$ at 26°C and with a gate time of 0.1 sec. The short-term frequency stability appears to be degraded by a factor of 2 as the drive voltage is increased from 5 to 10 volts.

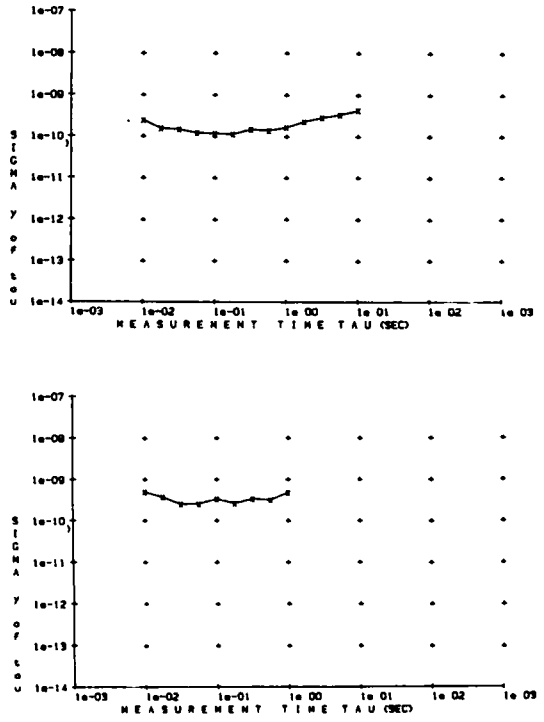



Figure 13. Short-term frequency stability of an oscillator (Device: S1001) at 26°C with a SAW resonator having aluminum on chromium as IDT fingers. (a) supply voltage = 5V, and (b) supply voltage = 10V.

The short-term frequency stability of an oscillator with a SAW resonator having pure aluminum IDT fingers was measured at 25°C and 175°C as a function of drive voltage and for three different counter gate times. The frequency stability expressed as $\sigma_y(\tau)$ has been plotted in Figures 14a and 14b for the oscillator operating at 25°C and 175°C, respectively. What is remarkable in these results is that at 175°C, there is a sharp degradation in the short-term frequency stability as the drive voltage is increased beyond 6 volts. It is possible that crossing the power dissipation threshold for the SAW resonator with pure aluminum electrodes results in an instantaneous onset of metal migration which in turn causes the observed degradation in the frequency stability.

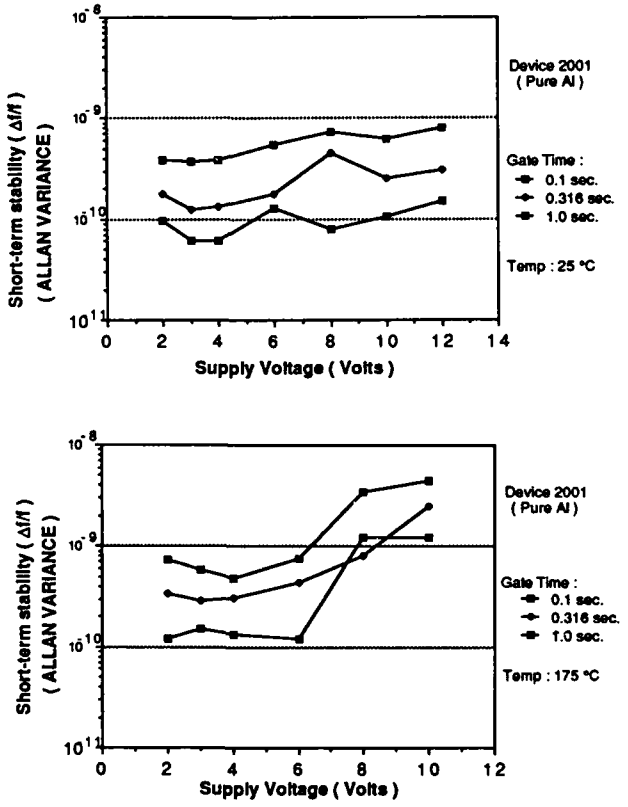



Figure 14. Short-term frequency stability of an oscillator with a SAW resonator having pure aluminum IDT electrodes as a function of supply voltage. (a) Operating temperature = 25°C; and (b) Operating temperature = 175°C.

Figures 15a and 15b illustrate similar results for a SAW oscillator with copper-doped aluminum IDT fingers. Overall, this oscillator showed the best short-term frequency stability of the three different metallizations for the IDT fingers employed in this study. It is also clear from these results that with the copper-doped aluminum electrodes, the short-term frequency stability is maintained even for high drive levels as well as high operating temperatures of up to 175°C.

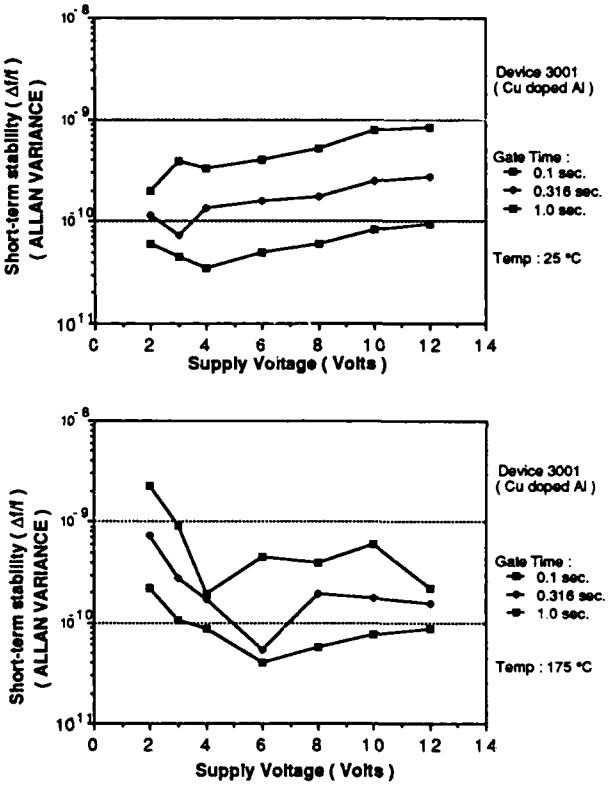



Figure 15. Short-term frequency stability of an oscillator with a SAW resonator having copper-doped aluminum IDT electrodes as a function of supply voltage. (a) Operating temperature = 25°C; (b) Operating temperature = 175°C.

Figure 16 shows the measured short-term frequency stability of two SAW oscillators with pure Al and Cu-doped Al IDT electrodes as a function of the counter gate time. These results were obtained after the two SAW oscillators were driven at different drive levels for varying time intervals at 175°C, as shown in Figure 12. It is evident from these results that because of the aluminum migration into quartz, there is a dramatic degradation in the short-term frequency stability or increase in the phase noise in the SAW resonator with pure Al IDT electrodes.

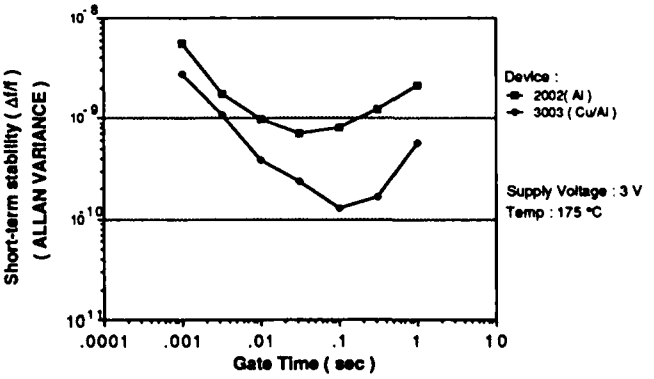



Figure 16. Short-term frequency stability of an oscillator with SAW resonators having pure aluminum and Cu-doped Al IDT electrodes.

## 7. SUMMARY AND CONCLUSIONS

We have studied the role of four distinct factors on the long-term stability of SAW resonators. A summary of results and conclusions follows:

1. Relaxation of intrinsic stresses:

   The two-port SAW resonator (unsealed) with pure aluminum IDT fingers which had not been annealed at all during or after fabrication, showed approximately 25 ppm of upward frequency drift before attaining a stable state. We attribute this drift to be largely due to the relaxation of intrinsic stress in the polished quartz surface and IDT metallic fingers. This observation is similar to that made by other workers [9].

   The two glass-frit sealed SAW resonators relaxed to a more stable state in about 1000 hours at approximately 135°C. During this annealing period, the SAW resonators showed a typical downward frequency drift of about 3 to 4 ppm. We consider this drift to be due to the relaxation of the glass frit sealing induced stresses in the frit as well as the substrate.

2. Presence of chromium interface between aluminum and quartz:

   The presence of chromium interface significantly slows down the complete stress relaxation compared to the device with pure aluminum fingers. In fact, the device under test did not stabilize even after 3000 hours of annealing at approximately 135°C. The stress analysis of thin film plated structure indicates that the relatively thin chromium layer of thickness 60 Å will be more highly stressed compared to aluminum film of thickness 1700 Å. This may have resulted in the prolonged relaxation of the SAW resonator frequency.

3. Dynamic stress induced migration of aluminum into quartz:

   The overall experimental evidence is in complete agreement with the results discussed by Shreve et al. [2]. A yet more significant aspect of this investigation is that it points out the influence of the electrode film thickness, in addition to other important parameters, such as unloaded acoustic quality factor, wave frequency, device aperture, and the power dissipation in the SAW device on the magnitude of the film stress that is developed due to a propagating or standing wave. The film stress threshold for the initiation of the aluminum migration is estimated to be in the same range of 6 to $10 \times 10^7$ N/m$^2$ as reported in Refs. [1] - [4].

   Perhaps, the most important finding is that for devices with pure aluminum electrodes, the tolerance for the safe drive level is quite narrow in order that the film stress does not exceed the aforementioned threshold. Other factors being equal, increasing the film thickness will reduce the stress magnitude in the film for a given power flux in the surface wave. Conversely, a very thin film will be relatively more stressed due to a propagating surface wave.

   This observation leads us to believe that a relatively thin layer of chromium (sometimes used as an adhesion promoting interface) will be more prone to stress induced migration into the substrate as well as into the adjacent aluminum film.

4. Drive level sensitivity of SAW oscillators:

   The drive level sensitivity of the SAW oscillator at the room temperature has been found to be in the range of 0.2 to 0.36 ppm per volt change in the d.c. power supply voltage to the amplifier. This implies that the SAW oscillator frequency stability will be on the order of 0.01 to 0.04 ppm for a voltage supply regulated to within 1% of its output. Particularly illuminating result from this study indicates that the instability in the oscillator frequency output due to the supply voltage fluctuation is caused by both the SAW resonator and amplifier phase shift.

5. Short-term frequency stability of SAW oscillators:

   The short-term frequency stability of SAW oscillators has been studied for three types of IDT electrode metallizations (aluminum on chrome; pure aluminum; and copper-doped aluminum), drive-level, gate time, and the operating temperature (up to 175°C).

   At relatively low temperatures and drive-levels the type of IDT electrode metallization does not significantly influence the observed short-term frequency stability.

   However, for high temperature applications and moderately high drive-levels, copper-doped aluminum electrodes have produced significantly superior short-term frequency stability.

   Overall, the best short-term frequency stability is found to be for a gate time of 0.1 sec.

   In conclusion, we have found that for SAW resonators with pure aluminum IDT fingers, the stress threshold for the onset of aluminum migration is easily crossed even for moderate drive levels. However, if the metal migration can be completely avoided, the excellent long-term stability of SAW oscillators can be ensured even at temperatures much higher than 135°C. Two possible ways of avoiding the onset of metal migration are: (1) to maintain the circulating power level in the surface wave below a relatively low threshold value for devices with pure aluminum IDT fingers; and (2) to employ Cu or Ti doped aluminum IDT electrodes which increases the drive level threshold for the initiation of metal migration.

## ACKNOWLEDGEMENTS

The authors would like to thank H. Hori for the optical inspection of the damaged IDT area, and T. Takahashi for help in the preparation of the manuscript.

## REFERENCES

1. J.I. Latham, W.R. Shreve, N.J. Tolar, and P.B. Ghate, "Improved metallization for surface acoustic wave devices," Thin Solid Films, vol. 64, pp. 9-15,1979.

2. W.R. Shreve, R.C. Bray, S. Elliott and Y.C. Chu, "Power dependence of aging in SAW resonators," in

1981 Ultrasonics Symp. Proc., pp. 94-99, (IEEE Cat. no. 81-CH1689-9).

3. C.A. Adams and J.A. Kusters, "Improved long-term aging in deeply etched SAW resonators," in Proc. 32nd Annu. Symp. Frequency Control, 1978, pp. 74-76.

4. W.R. Shreve, J.A. Kusters, and C.A. Adams, "Fabrication of SAW resonators for improved long-term aging," in 1978 Ultrasonics Symp. Proc., pp. 573-579, (IEEE Cat. no. 78CH1344-1SU).

5. W.J. Tanski, M. Block, and A. Vulcan, "High performance SAW resonator filters for satellite use," in 1980 Ultrasonics Symp. Proc., pp. 148-152, (IEEE Cat. no. 80CH1602-2).

6. T.E. Parker, J. Callerame, and G.K. Montress, "A new all quartz package for SAW devices," in Proc. 39th Annu. Symp. Frequency Control, 1985, pp. 519-525.

7. B.K. Sinha, W.J. Tanski, T. Lukaszek, and A. Ballato, "Influence of biasing stresses on the propagation of surface waves," J. Appl. Phys., vol. 57, pp. 767-776, 1985.

8. B.K. Sinha and S. Locke, "Acceleration and vibration sensitivity of SAW devices," IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control, vol. UFFC-34, pp. 29-38, 1987.

9. R.C. Bray, L.L. Pendergrass, C.A. Johnsen, T.L. Bagwell, and J.L. Henderson, "Annealing behavior and phase noise performance of SAW resonators," in 1985 Ultrasonics Symposium Proc., pp. 247-252, (IEEE Cat. no. 85CH2209-5).

42nd Annual Frequency Control Symposium - 1988

# THEORY AND DESIGN OF LOW PHASE NOISE MICROWAVE OSCILLATORS

Robert G. Rogers
California Microwave Inc.
990 Almanor Avenue
Sunnyvale, California 94086

## Summary

Linear oscillator theory is first examined to show circuit relationships and limitations. Then the conclusions from nonlinear analysis show contributions of second- and third-order nonlinearities to oscillator performance, particularly the phase noise origins.

Making use of the theory, results from several oscillators constructed are presented. Indirect methods of evaluating expected phase noise performance are examined, including second harmonic level, d.c. current shift between oscillating and non-oscillating conditions, and "extinction frequency" which is the shift in oscillation frequency as the d.c. supply voltage is reduced from normal levels to the levels at which oscillation just ceases.

## Theory

The design of any oscillator must begin with the theory. Consider the linear analysis of Robins[1]. His equation for single sideband phase noise in dBc in a 1 Hertz bandwidth is

$$\mathcal{L}(fm) = \frac{FkT}{8CQ^2}\left(\frac{f_o}{f_m}\right)^2 , \qquad (1)$$

where

F = Transistor Noise Figure
kT = Thermal Noise in 1 Hz bandwidth at Temp T = -174 dBm at room temperature
C = Carrier Level
Q = Resonator Q
$f_o$ = Oscillation Frequency

Equation (1) shows the relationships between the factors affecting the noise, but even with a pessimistic noise figure of 10 dB the actual microwave oscillator will be 20 to 25 dB noisier than calculated.

Van der Pol's nonlinear analysis[2] made use of second- and third-order nonlinearities in the active device transfer function. The third order coefficient predicts the steady-state oscillation amplitude, but for analysis he eliminated the second order term from his equations. However, Esdale and Howes[3] used the reflection form of oscillator to show design requirements. With this oscillator, separation into two elements is convenient both in analysis and design.

Figure 1 shows the block diagram, where $\Gamma_T$ is the transistor circuit's reflection coefficient, and $\Gamma_L$ the load reflection coefficient of the resonator with coupled useful output load.

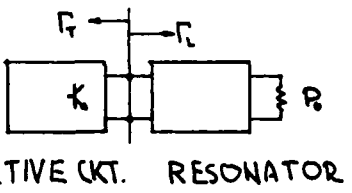



Fig. 1 Block Diagram of Reflection Type Oscillator Circuit

Steady-state oscillation conditions are given by

$$\Gamma_T\Gamma_L = 1 \qquad (2)$$

obtained from summing to zero admittances or impedances around the loop. Equation (2) is actually two equations; the product of reflection coefficient magnitudes must be unity; the sum of their angles must be zero.

Rewriting (2) as

$$\Gamma_L = \Gamma_T^{-1} \qquad (3)$$

allows equating the curves for easier graphical analysis. Since in a practical oscillator the reflection coefficient at startup must be greater than unity, then $\Gamma_T^{-1}$ will plot as in Figure 2 to be smaller than $\Gamma_L$. Initially, Equation (3) will not hold; the two values of $\Gamma_T^{-1}$ and $\Gamma_L$ at $f_o$ will be as shown. But $\Gamma_T^{-1}$ will rise along a trajectory with the instantaneous vector $\frac{d\Gamma_T^{-1}}{dA}$ showing the path with amplitude change. At steady state (3) (or (2)) will be satisfied. Esdale and Howes show that the resonator vector $\frac{d\Gamma_L}{d\omega}$ must make an angle $\alpha$ less than $180^0$ with $\frac{d\Gamma_T^{-1}}{dA}$.

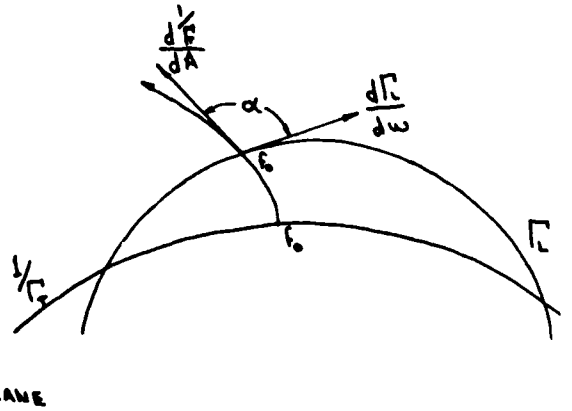



Fig. 2 Plot of $\Gamma_L$ and $\Gamma_T^{-1}$.

CH2588-2/88/0000-301 $1.00 C 1988 IEEE

In summary, nonlinear analysis shows that

1. a signal (usually noise) is required to start oscillations.
2. active device third-order nonlinearity determines steady-state amplitude.
3. second-order nonlinearity determines 1/f noise level.

And for optimum performance

1. $\frac{d\Gamma_T^{-1}}{d\alpha}$ and $\frac{d\Gamma_L}{d\omega}$ intersect orthogonally.
2. the change of $\Gamma_T^{-1}$ magnitude with amplitude is minimized.
3. $\Gamma_L$ change with frequency is maximized (maximum Q).
4. $\alpha = 90^0$.

## Design

Returning to Equation (2), with separation of active and resonator circuits at a practical reference plane, the active circuit can be synthesized to produce $|\Gamma_T| > 1$, and $\Gamma_L$ of reasonable value. An example of a $\Gamma_L$ plot is Figure 3. Marker 3, at 6.5055 GHz is the resonant frequency, and markers 2 and 4 are only 1 MHz apart.

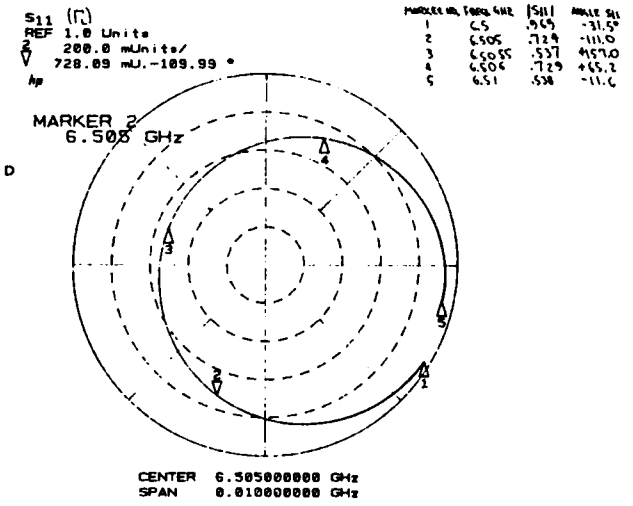



Fig. 3 Resonant Frequency Loop of $\Gamma_L$.

Figure 4 shows a $\Gamma_T^{-1}$ plot. Note here that markers 2 and 4 are 210 mHz apart. Connecting these two circuits at their reference planes, as oscillation amplitude grows the $\Gamma_T^{-1}$ curve will rise along the $\frac{d\Gamma_T^{-1}}{d\alpha}$ trajectory to steady state oscillation, with the two curves tangent at oscillation frequency, as in Figure 2. From Figures 3 and 4, if $\frac{d\Gamma_T^{-1}}{d\alpha}$ rises orthogonally to its initial value, oscillation will be sustained at a frequency about 200 kHz above the 6.5055 GHz cavity resonance frequency. Some additional phase shift to eliminate this frequency shift could be added.

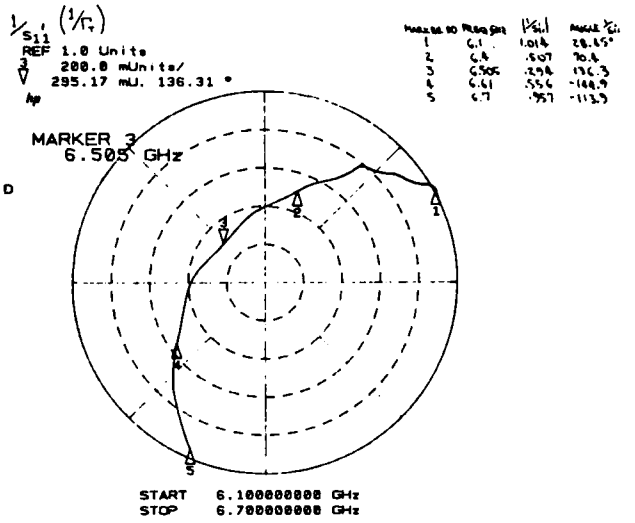



Fig. 4 Plot of $\Gamma_T^{-1}$

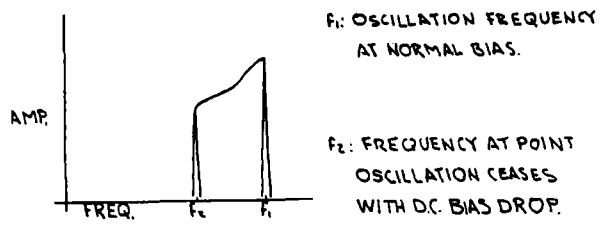
## Evaluation

First, after the oscillator elements have been combined, the angle may be estimated by the discrepancy between oscillation frequency and cavity resonance.

A direct noise measurement is the ultimate evaluation, but particularly in the case of a very low noise oscillator the procedure is tedious. Other indications of oscillation are readily available during initial adjustments. The second order device nonlinearity is the mechanism for placing device 1/f noise around the signal. A Fourier analysis of a distorted waveform will have a frequency invariant, or "dc" term if any even harmonics exist. In an oscillator this "dc" term will appear as a shift in dc collector current between the oscillating and non-oscillating conditions. Quenching oscillation without bias change should show no more than 1% current shift for a good oscillator.

Second harmonic output also shows how close to linear the oscillator is: For a good oscillator, it should be better than 40 dB below the fundamental, and 50 to 60 dB down for a superior one.

A very useful indicator, readily obtainable at the workbench during initial oscillator adjustment is the frequency shift to oscillation extinction with dc bias reduction: the "extinction frequency." It can be defined readily with Figure 5. At design dc bias, oscillation frequency is $f_1$. Slowly dropping power supply voltage, $f_2$ is the frequency at which oscillations just cease. The magnitude of the difference between $f_1$ and $f_2$ is the extinction frequency, which is a function of phase noise since it shows relative insensitivity of the oscillator to bias, or noise fluctuations in the device current.




EXTINCTION FREQUENCY = $f_1 - f_2$

Fig. 5 Extinction Frequency Definition

## Results

Plots of phase noise measured for several frequencies is shown in Figure 6. These are all averages, not worst or best, and are each for a single oscillator, with the exception of the 4.4 GHz curve. Bipolar transistors only, two types, were used. All are free-running and, except for the 6.3 GHz DRO, use empty cylindrical cavities.

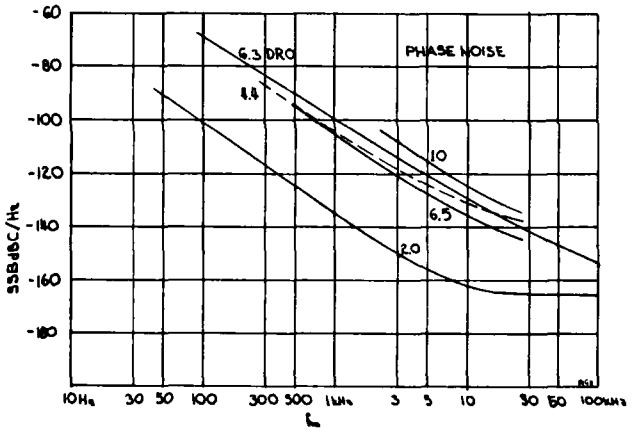



Fig. 6 Single Sideband Phase Noise of Oscillators at Various Frequencies (in GHz)

Figure 7 is a curve that can be useful in noise prediction, making use of the measured extinction frequency as defined above. The single sideband phase noise in a 1 Hertz bandwidth at 25 kHz offset is plotted against extinction frequency. The legend identifies the oscillation frequencies. All are bipolar, two types, and all are empty cavities, except the DRO. One oscillator may be represented by several points during a trimming adjustment. This plot has been successfully used many times to predict performance of a just-completed oscillator.

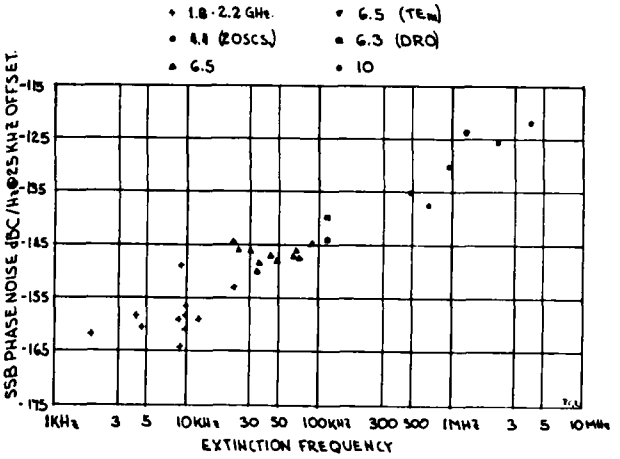



Fig. 7 Phase Noise at 25 kHz Offset vs. Extinction Frequency

Average output power of these oscillators is about +10 dBm, so the phase noise floor in a 1 Hz bandwidth is -184 dBc. A curve averaging these points will approach device noise figure, above this noise floor.

FET device oscillators would generate a curve above this one, presumably parallel to it.

## Conclusions

Adhering to the conditions imposed by theory, performing a painstaking design, and reducing it to reality with care will produce an exceptionally quiet microwave oscillator.

Initial evaluations during adjustment, with easily obtained numbers at the workbench will allow oscillator adjustments to obtain optimum performance.

## References

1. Robins, W. P, "Phase Noise in Signal Sources," Peter Peregrinus Ltd., London, 1982.

2. Van der Pol, B, "The Nonlinear Theory of Electrical Oscillations," Proc. IRE, Vol. 22, pp. 1051-1073, September 1934.

3. Esdale, D. J. and Howes, M. J, "A Reflection Coefficient Approach to the Design of One-Port Negative Impedance Oscillators," IEEE Trans on MTT, Vol. MTT-29, pp. 770-776, August 1981.

42nd Annual Frequency Control Symposium - 1988

# A SIMPLE WAY OF CHARACTERIZING HIGH Q OSCILLATORS

by Jules Goldberg--National Semiconductor, 2129 Hoover Ct. Santa Clara, Ca., 95051
Home: (408)296-8872 Wk: (408)721-7218

## 1.0 ABSTRACT

This paper clarifies the design and analysis of a high Q oscillator by deriving rigorous, practical approximations to all of its poles.

Previously published concepts which separate the resonant part of the crystal(OSCI) from the rest of the circuit(LLATOR) are used.[4,5,6,7,8,9,10,12,13] This paper then uses the high Q response of the crystal to generate approximate linear equations whose solutions are the pole locations plus accurate expressions for the approximation errors.

In the past, this problem was attacked using feedback theory to generate and approximately solve high degree equations specific to the circuit being designed.[4,5,7,8,9,10,12,13] These equations usually contained the Butterworth-Van Dyke crystal model, resulting in an extremely complicated analysis. The complete expression was then simpified assuming the crystal has high Q.

This paper proves that if certain unpublished simple rules are obeyed, the following facts are true:

oThe oscillation frequency equals the crystal resonant frequency plus a small error term nearly proportional to Im(Zo). (Zo is the LLATOR impedance)

oGrowth of oscillation from start up is nearly proportional to the algebraic sum of the crystal's internal resistance and Re(Zo).

oAll non-resonant network poles are close to those of the LLATOR alone.

Some of these facts follow from the 'Negative Resistance Model'.[1,4,9,11,19] However, no rigorous proof or approximation error formula has appeared in the literature to support them. The author believes that the results are analogous to the infinite gain approximation in op-amp feedback design.

This paper applies to any crystal or high Q tuned circuit oscillator regardless of configuration. Since oscillator design is based on the LLATOR, the LLATOR can be specified, designed, and tested by itself.

The results of this paper can be applied directly by recalling that an energized crystal resembles a sinusoidal current source.[1,3,16] SPICE[14] can analyse a distorted voltage response to find the impedance across the crystal at the fundamental frequency.[1,3,16] Output waveforms and oscillation growth patterns can also be observed. Start up can be guaranteed with a SPICE small signal AC analysis, and the LLATOR impedance at spurious crystal harmonics can be analysed to prevent unwanted oscillations. In the past, people have broken feedback loops, plotted loop gain and phase shift, analysed root locus plots[7,13], and tried to simulate actual start up. Since growth can be very slow compared to the oscillation period, simulation of start up has always been impractical.[16]

Design examples for the Pierce and Miller oscillator families, experimental data and sample SPICE programs are given. These examples are different from what has appeared in the literature.[1,4,5,8,9,11,19]

CH2588-2/88/0000-304 $1.00 C 1988 IEEE

## OUTLINE

1.0 Abstract
1. Purpose
2. Previous Methods of Oscillator Design
3. Basic Methodology
4. Expected Results

2.0 Derivation of Oscillator Equations
1. Basic Approach

2.1 Case 1 Design at Crystal Resonance
1. Set up of Master Linear Equation
2. Final Approximate Form of Solution
3. Conditions for Accurate Approximation
4. Application of Formulas to Oscillation Startup and Steady State.

2.2 Case 2 Design Away From Crystal Resonance
1. Set up of Master Linear Equation
2. Final Approximate Form of Solution
3. Conditions for Accurate Approximation

2.3 Conclusions From the Derivation of Basic Crystal Oscillator Equations.

3.0 Two Port Analysis for the Pierce and Miller Oscillators.
1. General two Port Network Based on Y parameters.
2. The Basic Loop Equation.
3. Parasitics in Parallel with the Crystal.

3.1 Pierce Oscillator Design.
1. Basic Design Equation in General Terms.
2. Inclusion of all Parasitics.
3. Stepwise Solution of the Equation Using the Quadratic Formula and no Further Approximations.
4. Minimum Realizable z'/r'

3.2 Miller Oscillator Design.
1. Basic Design Equation in General Terms.
2. Inclusion of all Parasitics.
3. Stepwise Solution of the Equation Using the Quadratic Formula and no Further Approximations.
4. Minimum Realizable z'/r'

3.3 Realizability Limits on z.

3.3.1 Maximizing z'/r' for a Given r.

3.3.2 Limits on z for a Given r.
1. Limit on r Given the Parameters, the Circuit Configuration, and the Type of Oscillator being Designed.

3.4 Concluding Remarks

4.0 Sample Oscillator Designs

4.1 Extracting Circuit Carameters with Spice

4.2 Circuit Configurations to be Analysed

4.3 Design Implementations

4.4 Discussion

4.5 Small Signal AC Simulation
1. Plots of Z over Frequency

4.6 Transient analysis
1. Output Waveform and Amplitude Determination.
2. Growth Simulation

5.0 Experimantal Proof of the Negative Resistance Model.
   1. What Must be Proved.
   2. Experimental Method.
   3. Experimental results.

6.0 Concluding Remarks.

APPENDIX 1 EQUATIONS FOR OSCILLATION GROWTH

BIBLIOGRAPHY

## 2.0 DERIVATION OF BASIC CRYSTAL OSCILLATOR EQUATIONS

The oscillator circuit can be broken into two parts

a) The hi-Q portion of the crystal (OSCI)[1]
b) The rest of the oscillator(LLATOR)
including the crystal parallel case capacitance(Co) and other vibration modes.

Let the LLATOR impedance equal Z(S).

Crystal Resonant Oscillation Mode (OSCI)

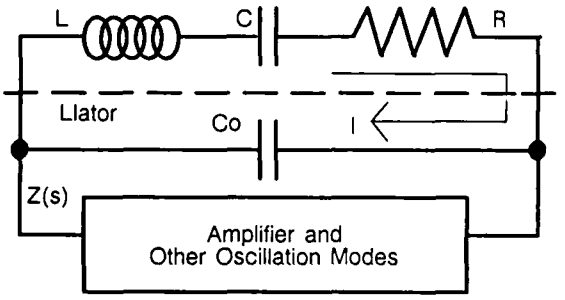



Fig. 1 Crystal Oscillator Split into Two Parts

At a complex pole, the impedance around the current loop of Fig. 1 equals zero because a pole implies current flow without an output voltage. Therefore,

$$Z(s)+sL+1/sC+R = 0. \quad (1)$$

For good design, equation (1) must be solved for all of its roots.

### 2.1 CASE 1 DESIGN AT CRYSTAL RESONANCE

(1) can be rewritten as follows:

$$(Z(s)+R)(s/Wo)/(QR(1+ s/jWo)) + 1-(s/jWo) =0, \quad (2)$$

where Wo =1/SQRT(LC), R is the crystal's motional resistance and Q = 1/(WoCR).

Assuming that the function Z(s) has no poles near crystal resonance and that $|1-s/jWo| << 1$, Z(s) can be written as:

$$Z(s)=Zo + d(s -jWo), \quad (3)$$

where d =dZ(s)/ds at s =jWo.

Using (3), linearization of (2) near Wo produces:

$$(s-jWo)(j2dWo+Zo+R)/(4WoQR) + (Zo+R)jWo/(2WoQR) +1 -s/jWo =0., \quad (4)$$

where the Taylor series[2] remainder, proportional to 1/(QR) is neglected. Solving for s, the complex pole, gives:

$$s - jWo +jWo \frac{j(Zo + R)/(2QR)}{1 -j(j2dWo + Zo + R)/(4QR)} =$$

$$jWo\left(1+ \left(1+ \frac{(j(Zo + R)- 2dWo)/(4QR)}{1 -(j(Zo + R) -2dWo)/(4QR)}\right)j(Zo+R)/(2QR)\right) \quad (5)$$

$$\text{If } |j(Zo+R)/(2QR)| + |dWo/(QR)| << 1, \text{ then} \quad (6)$$

$$\left|\frac{(j(Zo + R)- 2dWo)/(4QR)}{1 -(j(Zo + R) -2dWo)/(4QR)}\right| << 1. \quad (7)$$

(5) is then simplified to

$$s - jWo - Wo(Zo+R)/(2QR) \quad (8)$$

and can be rearranged to get

$$s/jWo-1 - \left(1+ \frac{(j(Zo + R)- 2dWo)/(4QR)}{1 -(j(Zo + R) +2dWo)/(4QR)}\right)j(Zo+R)/(2QR) \quad (9)$$

(8) should be used directly for all high Q oscillator design. Application of (6) guarantees the validity of (8). (7) and (9) can be used to estimate the approximation errors. Physically, (6) says that the reactance of the resonating components within the crystal is much larger than the magnitude of Zo +R and that this reactance changes much faster with frequency than Z does. (Two well known facts.) If the errors are too large, other design proceedures should be used. A less accurate version of the above analysis has been done considering the LLATOR impedance to be relatively constant near crystal resonance[1], I.E. $|dWo/(2QR)| - 0$.

It follows from (8) that a right half plane pole is produced when

$$R+RE(Zo) < 0 \text{ at } Wo. \quad (10)$$

$$RE(s)= -[R+RE(Zo)]Wo/(2QR)= 1/(\text{THE OSC. GROWTH TIME CONST.}). \quad (11)$$

In the steady state, this value = 0. or R=-RE(Zo).

Discrimination between crystal overtones can be achieved by shaping (R+RE(Zo))Wo/(2QR) to be most negative near Wo[3].

It follows from (8) that

$$IM(Zo)/(2QR) - 1-IM(s)/jWo, \quad (12)$$

approximates the fractional amount of frequency deviation from resonance that occurs with a particular crystal[1] and active network. This deviation is very important for designs where the frequency as a function of supply, component and temperature variations must be precisely known. If Q and R are well known, crystal resonant frequency variations with temperature can be compensated for by varying IM(Zo).

As the oscillations grow from small signal to steady state, the value of Z is changing. If (6) is obeyed, then the rate of growth, controlled by the (GROWTH TIME CONSTANT), is much slower than the oscillation frequency and has little short term effect on the waveform. Therefore, the analysis accurately predicts the oscillation frequency and growth rate at any simulated time.[11](5), (6), and (8) are fundamental laws for high Q tuned circuit oscillation and are true regardless of the circuit configuration or technology being used.

It is important to understand from (5) that (8), (10), (11), and (12) can be invalid if |dWo/(QR)| is not << 1 even though |1. - s/jWo| << 1. Therefore, both terms in (6) are critical for high Q oscillator design.

2.2 CASE 2 DESIGN AWAY FROM CRYSTAL RESONANCE

Off resonance design is handled in a manner similar to section 2.1. (1) can be rewritten as:

$$(s/WoQR)Zx(s)/(1 + (s/WoQ) + (s/Wo)^2) + (1-s/So)^n = 0, \quad (13)$$

where So is a pole of Z(s), Zx(s) has all zeros and poles of Z(s) except So, and n is the multiplicity of So.

(13) is then translated to

$$-[-Zy(s)/(QR)]^{1/n} + 1 - s/So = 0, \text{ where} \quad (14)$$

$$Zy(s)/(QR) = (s/Wo)Zx(s)/((QR)(1 + (s/WoQ) + (s/Wo)^2)) =$$

$$Zx(s)/(\text{the resonant circuit impedance at } s) \quad (15)$$

$$\text{If } |1-s/So| << 1, \quad (16)$$

$$\text{then } Zy(s) \sim (s - So)Dy + Zyo, \quad (17)$$

where Dy = dZy(s)/ds at So and Zy(So) = Zyo.

Linearization of (14) and substitution of (17) produces:

$$-(-Zyo/(QR))^{1/n} - (s-So)Dy(-Zyo/(QR))^{1/n}/(nZyo) + 1 - s/So = 0, \quad (18)$$

where the Taylor series[21] remainder, proportional to $1/(QR)^{3/n}$ is neglected. Solving for s, the complex pole, gives:

$$s \sim So\left(1 - \frac{(-Zyo/(QR))^{1/n}}{(SoDy/(nZyo))(-Zyo/(QR))^{1/n} + 1}\right). \quad (19)$$

$$\text{If } (1. + |SoDy/(nZyo)|)|-Zyo/(QR))|^{1/n} << 1., \text{ where} \quad (20)$$

$$|SoDy/(nZyo)| = |So(Zx'/Zx) + (1.-(So/Wo)^2)/(1 + (So/WoQ) + (So/Wo)^2)|/n \quad (20a)$$

then: $s \sim So.$ (21)

It follows from (15) and (20) that for good design, the impedance of the crystal at So should be much larger than Zx(So), all poles and zeros of Z(s) should be far from each other and Wo, and multiple poles are approximated less accurately than simple ones. If all poles of Z(s) are simple, then the restriction on its zeros cancels out.

All of this is under the engineer's control.

(10) and (21) can be used to understand the response to crystal overtones. Applying (10) reveals that no oscillation will occur if R + Re(Z) > 0 at the overtone. The overtone resonant arm of the crystal can be lumped into Z when applying (21) to show that it has a small effect on the main oscillation frequency. Similarly, the main resonant arm of the crystal has a small effect on the location of the overtone pole. This does not mean that large signal output waveforms are not affected. On the contrary, large current spikes can go through the crystal at the overtone frequencies. These spikes will affect the harmonic content of a distorted output.[22]

2.3 CONCLUSIONS FROM THE OSCILLATOR ANALYSIS

Approximate relationships have been found between the oscillator pole locations and those of its componant parts. To estimate the approximation errors, use (7), (9), (19) and (20). It follows from the analysis that calculation accuracy is controlled by the crystal Q and the LLATOR circuit design.

3.0 TWO PORT ANALYSIS FOR PIERCE OSCILLATOR DESIGN

Figure 2 presents a general 2 port network configuration that can be used to analyse many oscillator families.[1] The elements are easily derivable from standard Y parameters[4,5,9] and the translation is given in the figure for reference. Node three is the Y parameter datum node.

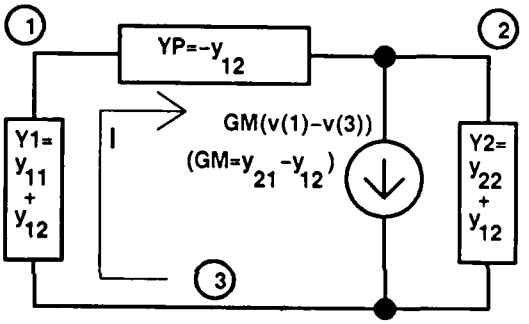



Figure 2. Oscillator Two Port Configuration

The Pierce oscillator family uses a transistor or a tube to produce GM. Y1,Y2 and YP become Resistors in parallel with Capacitors to represent the bias circuit, tube or transistor parasitics etc. A Pierce oscillator is formed by AC grounding node 3. The Colpitts is realized by AC grounding node 2 and the Clapp oscillator is formed by AC grounding node 1. In each case the crystal is placed between nodes 1 and 2.

Other oscillator families exist where the crystal is placed between nodes 1 and 3 or nodes 2 and 3. One member of this family is the Miller oscillator. It is discussed below and in Refs. 2, 4, and 11.

The oscillator equations for fig. 2 will now be derived. According to the theory, we need to find the impedance that the crystal 'sees' when connected between any particular set of nodes.

For any pole, the circuit of fig. 2 will have current flow, I, around a closed voltage loop. The loop equation is as follows:

$$I(1/Y1 + 1/Yp) + (I - GM(V1-V3))/Y2 =$$

$$I(1/Y1 + 1/Yp + 1/Y2) + GM(V3-V1)I/(I*Y2) =$$

$$I(1/Y1 + 1/Yp + 1/Y2 + GM/(Y1*Y2)) =$$

$$1/Y1 + 1/Yp + 1/Y2 + GM/(Y1*Y2) = 0. \quad (22)$$

(22) is a general equation that can be used to find the impedance looking into any pair of nodes in fig. 2. To do this, imagine that Y1, Y2, or Yp is replaced by a current source, I, driving the node pair in question. Figure 3 shows Yp replaced by the source. The impedance the current source 'sees' is Z'=E/I, where E is the voltage across it. Since the source is an active element, it forces I into the network. Therefore, the admittance the source is presenting to the remainder of fig. 3 is I/(-E) =-1/Z'. To solve (22) for the IMPEDANCE 'seen' by Y1, Y2, or Yp, replace it by -1/Z' and solve for Z'. (23) is the well known solution for the Pierce oscillator family[1,4,11]. (Yp is the admittance to be replaced.) It is easy to solve (22) for the other cases.

$$Z' = 1/Y1+1/Y2+GM/(Y1*Y2) \qquad (23)$$

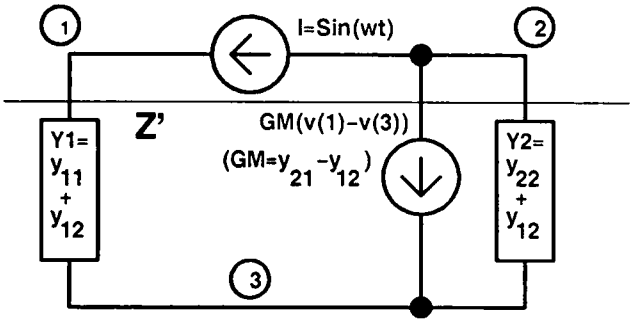



Fig. 3 Yp of fig. 2 Replaced By a Current Source

In a practical circuit Yp is made up of the crystal high Q series resonant circuit in parallel with a parasitic admittance (Yb). The real part of this admittance (g) is used to bias the circuit and the imaginary part (jy) is made up of the crystal case capacitor, Co, plus stray capacitors from the active oscillator circuit. Since Z' calculated in (23) does not contain this stray admittance, Yb, it must be factored in to calculate Z, the impedance the resonant circuit 'sees'. Adding Yb in parallel with Z', we get

$$Z = \frac{Z'}{1+Z'*Yb}. \qquad (24)$$

[7] is a paper that uses the SPICE program and root locus plotting to analyze the effects of Yb from the feedback theory viewpoint.

Note that (23) and (24) are INDEPENDENT of the node that is selected as the reference. Therefore, (23) and (24) apply equally to the Clapp, Colpitts, and Pierce oscillators.

### 3.1 PIERCE OSCILLATOR DESIGN

The small signal startup characteristics of the Pierce oscillator family will now be designed. Assume that Z = r+jz, the llator impedance is specified. Circuit elements Y1, Y2, and GM must be found to produce Z. Since Yb is a parasitic that we have little control over, we can only specify a maximum amount that can be lived with. Anything smaller will make r more negative. The approach is to calculate Z' from Z and Yb. If Yb is too big, Z' will be non-physical. Leaving out the signs of r, z, g, and y,

Z = r+jz, and Yb = g+jy then (solving (24) for Z')

$$Z' = Z/(1.-Z\,Yb) = (Z-|Z|^2\,Yb^*)/|1.-Z\,Yb|^2 =$$

$$r'+jz' = \frac{r-(r^2+z^2)g+j(z+y(r^2+z^2))}{|1.-Z\,Yb|^2} \qquad (25)$$

In most cases, the Pierce family oscillator circuit can be represented as shown in Fig. 4.[1] For high frequency work, GM may have an imaginary part. This is represented by j al in the figure. (al may be positive or negative.) G1 and G2 represent losses such as transistor output conductance, bias resistors, and bipolar transistor base input conductance. al is the result of FT in a bipolar transistor or polysilicon RC gate delay in an MOS transistor. Note that fig. 4 is compatable with the Hybrid-Pi model of a transistor neglecting the series base resistor.[4,5]

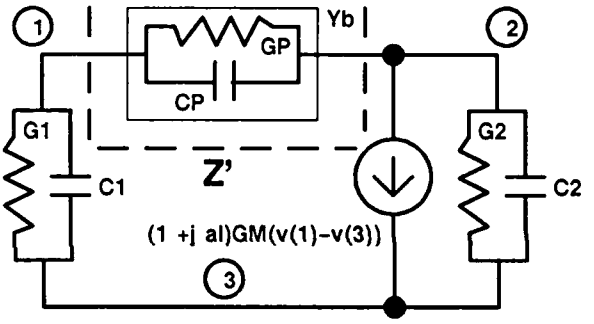



Fig. 4 Full representation of Pierce family oscillator circuit.

From the figure and (23), we get

$$Z' = r' + jz' = 1/Y1+1/Y2+GM/(Y1*Y2) =$$

$$\frac{1.}{jW\,C1+G1} + \frac{1.}{jW\,C2+G2} + \frac{GM(1+j\,al)}{(jW\,C1+G1)(jW\,C2+G2)} \qquad (26)$$

(26) is now solved for Y1 in terms of GM , Y2, and Z'.

$$Y1 = \frac{Y2+GM}{Z'\,Y2-1.} = \frac{(Y2+GM)(Z'^*\,Y2^*-1.)}{|Z'\,Y2-1.|^2} =$$

$$\frac{Z'^*|Y2|^2+GM\,Z'^*\,Y2^*-Y2-GM}{|Z'\,Y2-1.|^2} \qquad (27)$$

where Z'* and Y2* are the complex conjugates of Z' and Y2 respectively. In terms of the full circuit representations:

$$G1 = \frac{(W\,C2)^2\,r'+(W\,C2)GM(al\,r'-z')+GM^2\,G2(r'+al\,z')-GM+r'G2^2-G2}{|Z'\,Y2-1.|^2} \qquad (28a)$$

$$W\,C1 = \frac{-(W\,C2)^2\,z'^2 - (W\,C2)(GM(r'+al\,z')+1.) + GM\,G2(al\,r'-z') - al\,GM - z'^2 G2^2}{|Z'\,Y2 - 1.|^2}, \qquad (28b)$$

(28a) and (28b) are design equations for the Pierce oscillator. By inspection it is seen that there are two degrees of freedom available to the Pierce designer, namely the choice of Y2 and GM. All that is needed to make a physical circuit is that G1 and G2 include all strays. Since a negative C1 represents an inductor, there is no restriction on (28b).

In the real world, G1 and G2 are stray load conductances plus something proportional to GM (Transistor Early conduct[illegible] BETA effects, MOSFET output conductance, for example.) In the interest of workability, we will force the constant parts of G1 and G2 to 0. Therefore,

$$G1 = h1\,GM,\ G2 = h2\,GM \qquad (29)$$

(28a) is expanded into (30), a quadratic expression whose roots represent design limits on W C2. (30) is most positive at mid-range. GM is chosen so that real roots always exist. Since z' and r' are both negative numbers, only the linear coefficient of this expression can ever be positive.

$$(W\,C2)^2 r'^2 + (W\,C2)GM(al\,r'-z') + GM\,G2(r'+al\,z') - GM + r'^2 G2^2 - G2 - h1GM|Z'\,Y2 - 1.|^2$$

$$= (W\,C2)^2 r'^2 + (W\,C2)GM(al\,r'-z') + GM\,h2(r'+al\,z') - GM + r'^2\,h2\,GM^2 - h2\,GM^2$$

$$-h1\,GM\big((W\,C2)^2(z'^2+r'^2) + 2z'W\,C2 + h2\,GM(z'^2+r'^2) + 1 - 2r'^2\,h2\,GM\big) =$$

$$(W\,C2)^2(r'^2 - h1\,GM(z'^2+r'^2)) + (W\,C2)GM(al\,r'^2 - z'(1.+2h1))$$

$$+ GM^2 h2(r'^2(1.+h2+2.h1) + al\,z') - GM^2(1.+h1+h2) - h1\,h2\,GM^3(z'^2+r'^2) > 0. \qquad (30)$$

The discriminant of (30) is a third degree equation in GM. At this point this is impractical to deal with. Instead, maximize (30) minus the term

$$-(W\,C2)^2\,h1\,GM(z'^2+r'^2).$$

Under these conditions (30) is a maximum at

$$W\,C2 = GM\,((z'/r')(1.+2h1) - al)/2. \qquad (31)$$

The entire expression for (30) is now forced to be true when (31) is substituted in. Since (30) has not been truly maximized by (31), it is now true for a small range of W C2 near the value given in (31). After cancelling out GM, (30) becomes the quadratic inequality in -r'GM given by (32).

$$-GM^2((z'/r')(1.+2h1) - al)^2 r'^2/4. - h1\,GM^2 r'^2((z'/r')^2 + 1.)((z'/r')(1.+2h1) - al)^2/4.$$

$$+ GM\,h2\,r'^2(1.+h2+2.h1+al\,z'/r') - (1.+h1+h2) - h1\,h2\,GM\,r'^2((z'/r')^2 + 1.) > 0. \qquad (32)$$

The discriminant of (32) is a 4'th degree polynomial in (z'/r'). Satisfaction of this inequality using the formulas in ref. 20 leads to a design range for GM and a minimum realizable value for z'/r'. The correct root of the discriminant must produce a positive linear term in (32). This root is a function of the circuit parasitics. In section 3.3, z'/r' is related to the maximum realizable value of r.

Once the discriminant of (32) is satisfied:

a) Solve (32) to find the design range of GM.
b) Choose GM within the range.
c) Calculate G2 with (29).
d) Plug G2 and GM into (30) and solve for the permissible range of (W C2).
e) Choose (W C2) within the range. The midpoint gives a large G1. Since all strays are included, G1 > h1 GM.
f) Calculate G1 and (W C1) with (28a) and (28b). A negative C1 represents an inductor.

The Pierce oscillator design is now complete with all conductive parasitics included. The design range for GM and z'/r' may be smaller than necesary because (32) was not optimally obtained. Since the computations are so complex, they have been programmed on a computer.

### 3.2 MILLER OSCILLATOR DESIGN

The Miller oscillator design equation follows from (22) by replacing either Y1 or Y2 with -1/Z' to get (Y1 is arbitrarily chosen because (22) is symmetrical.)

$$-Z' + 1/Yp + 1/Y2 - Z'GM/Y2 = 0. \qquad (33)$$

The well known solution to (33) is

$$Z' = \frac{Yp + Y2}{(Y2 + GM)Yp} \qquad (34)$$

Fig 5. shows fig. 2 with Y1 replaced by a current source. Fig. 6 shows the admittances replaced by their circuit representations. Note the resemblence to figs. 3 and 4.

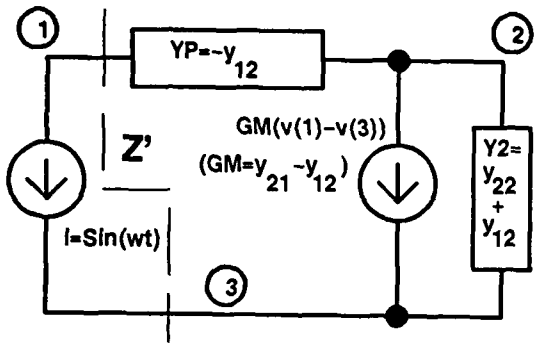



Fig. 5 Basic Miller Oscillator Circuit

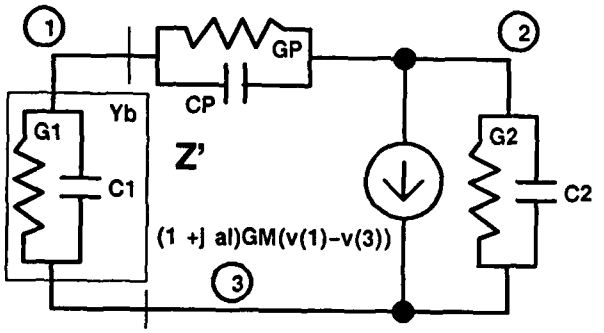



Fig. 6 Miller Oscillator with Admittance Circuit Representations.

As in Section 3.1,

$$Z' = r' + jz' = \frac{jW(Cp + C2) + gp + g2}{(jw\ Cp + gp)(GM + g2 + j(W\ C2 + GM\ al))} \quad (35)$$

Design of the Miller oscillator is very similar to that of the Pierce. Solving (34) for Yp we get

$$Yp = \frac{Y2}{Z'(Y2+GM) - 1} = \frac{Y2(Z'^*(Y2^* + GM^*) - 1.)}{|Z'(Y2+GM) - 1|^2}$$

$$= \frac{|Y2|^2 Z'^* + Y2\, Z'^*\, GM^* - Y2}{|Z'(Y2+GM) - 1.|^2}, \quad (36)$$

where Z'*, Y2*, and GM* are the complex conjugates of Z', Y2, and GM. Addition of full circuit representations gives

$$Gp = \frac{(W\ C2)^2\, r' + (W\ C2)GM(al\ r' + z') + GM\ G2(r' - al\ z') + r'G2^2 - G2}{|Z'(Y2+GM) - 1.|^2} \quad (37a)$$

$$W\ Cp = \frac{-(W\ C2)^2\, z' + (W\ C2)(GM(r' - al\ z') - 1.) - GM\ G2(al\ r' + z') - z'G2^2}{|Z'(Y2+GM) - 1.|^2}. \quad (37b)$$

Gp must be larger than all parasitics for a physical circuit. Inspection of (37a) shows that unless al is large (an undesirable feature at best), (W C2) must be negative to get a positive Gp. Therefore, C2 is physically an inductor. (37a) represents a quadratic inequality to be solved for the permissible range of (W C2) and is maximum at mid-range. To make the mathematics managable, assume (29) plus

$$Gp = hp\ GM. \quad (38)$$

Expansion of (37a) produces

$$(W\ C2)^2\, r' + (W\ C2)GM(al\ r' + z') + GM\ G2(r' - al\ z') + r'G2^2 - G2 - hp\ GM|Z'(Y2+GM) - 1.|^2 =$$

$$(W\ C2)^2\, r' + (W\ C2)GM(al\ r' + z') + GM\ G2(r' - al\ z') + r'G2^2 - G2$$

$$-hp\ GM\{(W\ C2)^2 (r'^2 + z'^2) + 2.(W\ C2)(GM\ al(r'^2 + z'^2) + z') +$$

$$GM\ (al^2 + (1.+h2)^2)(r'^2 + z'^2) + 2.GM(al\ z' - r'(1.+h2)) + 1.\} > 0. \quad (39)$$

The proceedure is similar to the Pierce design. Maximize (39) without the terms

$$-hp\ GM\{(W\ C2)^2 (r'^2 + z'^2) + 2.(W\ C2)GM\ al(r'^2 + z'^2)\}$$ to get

$$W\ C2 = -GM\ (al + (z'/r')(1.-2.hp))/2. \quad (40)$$

Force all of (39) to be true by substituting (40) for (W C2) to produce (41), a quadratic inequality in -r' GM after cancellation of GM.

$$hp\ GM\ r'^2\ (1.+(z'/r')^2)\{al(z'/r')(1.-2hp)-(1+h2)^2 - (al+(z'/r')(1-2hp))^2/4.\} - GM\ r'\{(al+(z'/r')(1-2hp))^2/4$$

$$-(1+h2)(h2+2hp)+al(z'/r')(h2+2hp)\} - hp - h2 > 0. \quad (41)$$

The discriminant of (41) is a 4'th degree polynomial in (z'/r'). Satisfaction of this inequality using the formulas in ref. 20 leads to a design range for GM and a minimum realizable value for z'/r'. The correct root of the discriminant must produce a positive linear term in (41). This root is a function of the circuit parasitics. In section 3.3, z'/r' is related to the maximum realizable value of r.

Once the discriminant of (41) is satisfied:

a) Use (41) to find the the design range for GM.
b) Choose GM within the range.
c) Calculate G2 with (29).
d) Plug G2 and GM into (39) and solve for the permissible range of (W C2).
e) Choose (W C2) within the range. The midpoint gives a large Gp. Since all strays are included, Gp > hp GM.
f) Calculate Gp and (W Cp) with (37a) and (37b). A negative Cp represents an inductor.

The Miller oscillator design is now complete including all conductive parasitics. The design range for GM and z'/r' may be smaller than necesary because (41) was not optimally obtained. Since the computations are so complex, they have been programmed on a computer.

### 3.3 REALIZABILITY LIMITS ON VALUE OF Z FOR PIERCE AND MILLER OSCILLATORS

The realizability limits given in (32) and (41) can be translated into limits on the oscillator specification. For obvious reasons, an engineer wants to know what is possible before he attempts a design. It follows from (25) above that for a given input specification,

$$x = \frac{z'}{r'} = \frac{z + y(r + z^2)^2}{r^2 - g(r + z^2)^2} > 0. \quad (42)$$

z can be chosen to maximize (z'/r') or realizability limits on z can be found so that (z'/r') > a given minimum value.

#### 3.3.1 FIND z TO MAXIMIZE (z'/r')

Since r determines whether the oscillator will start, it is taken as a given constant. z is the independant variable and x is a function of z. Using implicit differentiation, we get from (42)

$$z + y(r + z^2)^2 = x(r^2 - (r + z^2)^2 g) \quad (43)$$

$$1 + 2z\,y = -2x\,z\,g + (dx/dz)(r^2 - (r + z^2)^2 g), \text{ and} \quad (44)$$

$$2y = -2x\,g + (d\,x/dz)(r^2 - (r + z^2)^2 g), \text{ where} \quad (45)$$

(dx/dz) is set to zero to get (45). By inspection of (45), if x > 0, $(d\,x/dz)^2 < 0$ at the extreme value. This guarantees that the derived value of z maximizes x. From (44),

$$1. + 2z\,y + 2z\,g\,x = 0. \quad (46)$$

at the maximum. Plug (42) in to get

$$1. + 2\,z\,y + 2z\,g\,\frac{z + y(r + z^2)^2}{r^2 - g(r + z^2)^2} = 0. \quad (47)$$

(47) reduces to

$$z^2 g + 2\,y\,r\,z + r^2 - r\,g = 0. \quad (48)$$

The negative root of (48) is

$$z = \frac{r^2 - r\,g}{-y\,r + \mathrm{sqrt}((y\,r)^2 - g(r^2 - r\,g))}, \quad (49)$$

a function of r, y and g.

#### 3.3.2 FIND THE REALIZABILITY LIMITS ON z

Let us now satisfy (32) and (41) directly by solving (42) as follows

$$x = \frac{z'}{r'} = \frac{z + y(r + z^2)^2}{r^2 - g(r + z^2)^2} > x, \quad (50)$$

where x is found from (32) or (41). Taking r as a given constant, the quadratic inequality represented by (50) is

$$(y + x\,g)z^2 + z + (y + x\,g)r^2 - xr < 0. \quad (51)$$

The roots are

$$z(\min) = \frac{-1 - \mathrm{SQRT}(1. - 4(x\,g+y)r(r(x\,g+y) - x))}{2\,(x\,g+y)}, \quad (52a)$$

and

$$z(\max) = \frac{2\,r\,(r(x\,g+y) - x)}{-1 - \mathrm{SQRT}(1. - 4(x\,g+y)r(r(x\,g+y) - x))} \quad (52b)$$

z(min) goes to -infinity as g and y approach 0, while z(max) approaches x r. No value of z exists unless

$$1. - 4(x\,g+y)r(r(x\,g+y) - x) > 0. \quad (53)$$

The positive root of (53) and final REALIZABILITY CONDITION ON r is

$$-r < \frac{1.}{2(x\,g + y)(\mathrm{SQRT}(1. + x) + x)}. \quad (54)$$

(54) is the basic REALIZABILITY CONDITION for the Pierce and Miller oscillators.

To specify an oscillator:

a) Use (32) or (41) to find the minimum required value of x for your oscillator.
b) Use x to choose the maximum -r which can satisfy (54).
c) Find z from (52a) and (52b) or calculate z from (49).
d) Use (25) to find Z'
e) Follow the proceedures in sects 3.1 or 3.2 above.

### 3.4 CONCLUDING REMARKS

The reader is aware that larger values of -r will oscillate with crystals having higher motional resistance. Therefore LLATORS having large -r's are more useful than those with smaller -r's. From (54), -r is diminished by g, y, and x. Since x is a function of the parasitics h1, h2, and al {(32), (41)}, x is a figure of merit for the transistor part of a LLATOR. It also follows from (54) that as x ------->0., its effect on -r ---------->0. This is not true for g and y. In fact, as x------->0. the effect of g also approaches 0. (for finite y). This leaves y as the MOST EFFECTIVE PARASITIC THAT CONTROLS PIERCE AND MILLER OSCILLATOR DESIGN. Therefore, y must be minimized at the oscillation frequency. AT HIGH FREQUENCIES, THIS STRAY CAPACITOR EFFECT IS CATASTROPHIC. To counteract it, some designers add a parallel inductor for cancelation.[4] This must be done carefully because the inductor can add a right half plane pole to the LLATOR.

Section 3 has demonstrated exact methods of controlling the startup characteristics of Pierce and Miller oscillators. The proceedures can and have been computerized. The author does not believe that these proceedures have been preveously derived.

## 4.0 SAMPLE OSCILLATOR DESIGNS

Four circuits will be designed using the methods of section 3.

1. A CMOS Inverter Pierce Oscillator
2. A CMOS Miller Oscillator
3. A Bipolar Pierce Oscillator
4. A Bipolar Miller oscillator

The problems associated with the two technologies become apparent as the designs unfold.

### 4 1 EXTRACTING CIRCUIT PARAMETERS WITH SPICE

Recall from Sect. 3.0 that the oscillator two port configuration is based on y parameters. These can be extracted directly with a SPICE[14] small signal analysis as shown in figure 7. The oscillator parameters are then artificialy calculated within the program using controlled current sources.

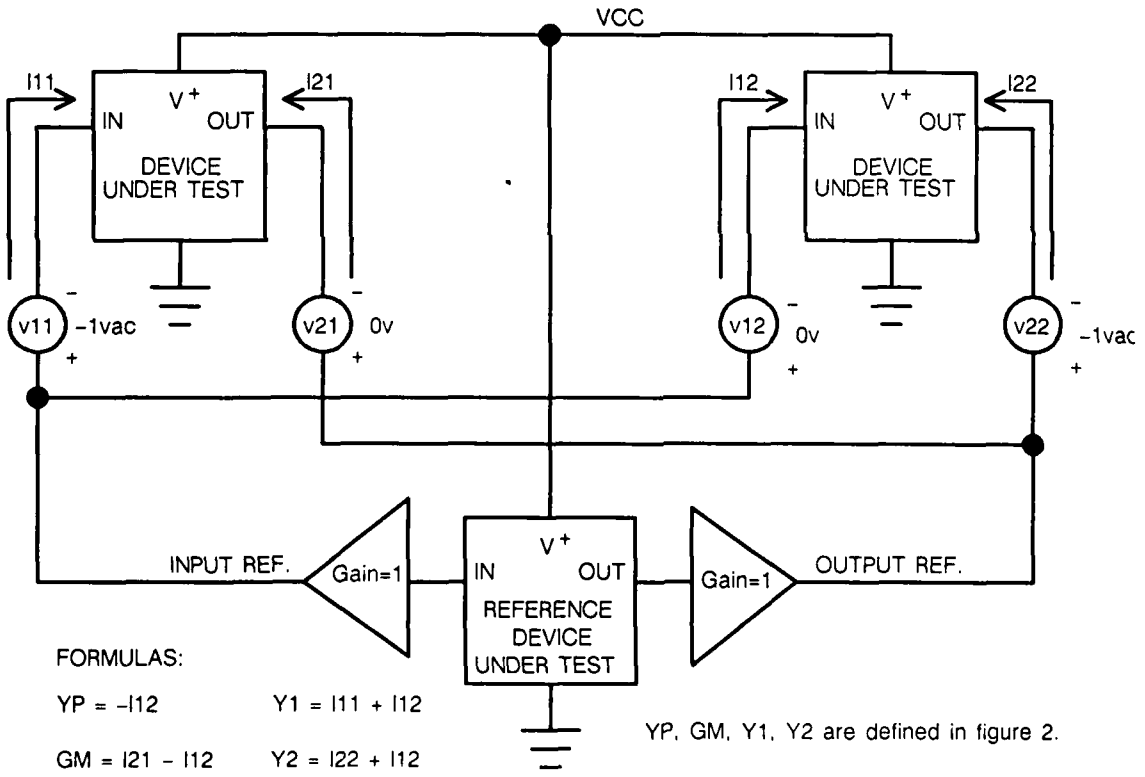



Figure 7. Circuit Diagram for Extracting Oscillator Design Parameters with SPICE.

### 4.2 CIRCUIT CONFIGURATIONS TO BE ANALYSED

The following discussion assumes some knowledge of IC design. Figures 8a and 8b show special CMOS and bipolar inverters adapted for Pierce oscillator design. A prime consideration for these examples is minimization of the H1, H2, hp and GM spread over production. To illustrate the design methods of section 3, the examples are purposely chosen to have inferior characteristics.

The CMOS inverter in figure 8a uses a current mirror in its P-Channel load transistor to control the production spreads of GM and H2. R-REF is to be made from the interconnect layer (p-diff., N-diff., poly. p-well) which has minimum resistivity variation over process and temperature. GM and added Co are caused by Q1 only. H2 is the result of Q1 and Q2 in parallel. All transistors should be drawn at greater than minimum length to avoid GM and H2 variations due to the production spreads in leff. Since all CMOS IC's have an input protection resistor of some form, Rprot is included in the figure for the sake of accuracy. Since it is resposible for generating the parasitics H1 and al, it should be made as small as is practical. The reader should note that since parasitics and process variations are inherent in IC's, the above considerations for figure 8a do not cure the problems in CMOS inverters but only make them somewhat manageable.

Figure 8b shows a bipolar version of the same inverter. Since the process contains only NPN transistors, no PNP current source is used. GM, H1, and H2 are determined by the current through Q1. These parameters are temperature and process sensitive. However, they do not vary as much as those of the CMOS circuit.

The circuits of figures 8a and 8b are used as the device under test' in the Spice program of figure 7. Typical, Fast, and Slow hypothetical process cases were run to illustrate the resultant design task. The results are shown in figures 9 and 10. To save space, full frequency response plots are given only for the slow cases.

The device models are taken from the FACT CMOS and ASPECT bipolar processes and are proprietary to National Semiconductor.

Note that the CMOS y parameters become frequency sensitive above 2mhz. On the other hand, the bipolar inverter

is relatively well behaved at 15mhz.

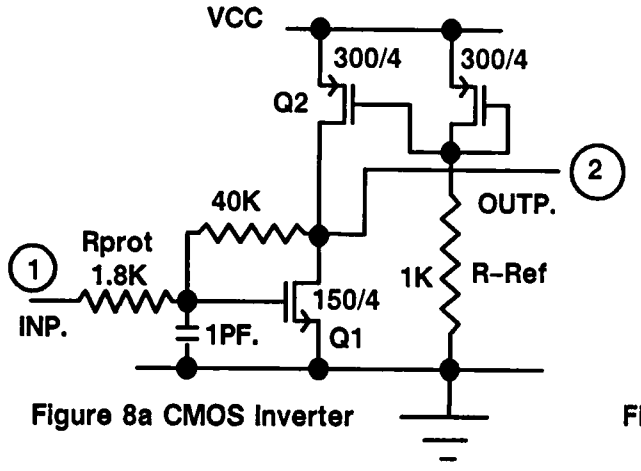



Figure 8a CMOS Inverter

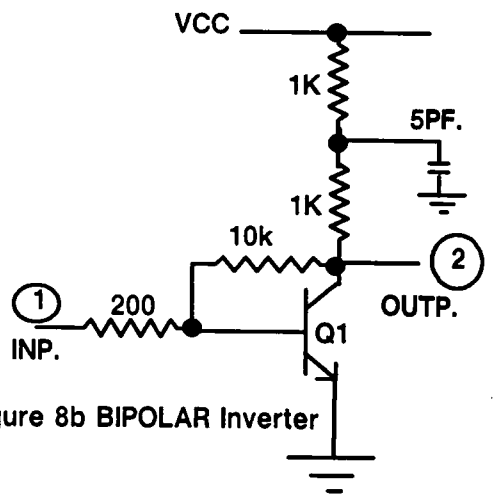



Figure 8b BIPOLAR Inverter

Figure 8 CMOS and BIPOLAR Inverters. The imaginary parts of Y1, Y2, and Yp are to be lumped with the final components.

```
VCC 1 0 5.0 ---- POWER SUPPLY
.SUBCKT DEVICE 1 2 3 5 ----- 'DEVICE PART OF SPICE PROGRAM'
RREF 6 0 1.K
*Typical resistors are shown here. Range on R-REF is 20%.
M1 1 6 6 1 MPS 4U 300U 1050.P 84.U 84.U
M2 1 6 3 1 MPS 4U 300U 1050.P 84.U 84.U
M3 3 5 0 0 MNS 4U 150U  525.P  525.P 42.U 42.U
R2 2 5 1.8K
C2 5 0 1.P
R3 3 5 40.K
.ENDS DEVICE
```

Tabulated Raw Inverter Data for CMOS.

SLOW AC ANALYSIS VCC= 4.5 VOLTS ICC= 3.195MA R-REF +20%

| FREQ | re(Y1) | im(Y1) | re(YP) | im(YP) |
|---|---|---|---|---|
| 1.080E+07 | 1.592E-05 | 9.196E-05 | 2.347E-05 | -2.622E-06 |
| 1.500E+07 | 2.992E-05 | 1.244E-04 | 2.307E-05 | -3.547E-06 |
| 2.084E+07 | 5.497E-05 | 1.645E-04 | 2.236E-05 | -4.689E-06 |
| FREQ | re(GM) | im(GM) | re(Y2) | im(Y2) |
| 1.080E+07 | 1.954E-03 | -3.382E-04 | 1.453E-04 | 2.569E-05 |
| 1.500E+07 | 1.902E-03 | -4.575E-04 | 1.443E-04 | 3.596E-05 |
| 2.084E+07 | 1.810E-03 | -6.050E-04 | 1.425E-04 | 5.061E-05 |

TYPICAL AC ANALYSIS VCC= 5.0 VOLTS ICC= 5.07MA Typical R-REF

| FREQ | re(Y1) | im(Y1) | re(YP) | im(YP) |
|---|---|---|---|---|
| 1.080E+07 | 1.648E-05 | 9.317E-05 | 2.357E-05 | -2.001E-06 |
| 1.500E+07 | 3.094E-05 | 1.259E-04 | 2.326E-05 | -2.704E-06 |
| 2.084E+07 | 5.673E-05 | 1.661E-04 | 2.271E-05 | -3.567E-06 |
| FREQ | re(GM) | im(GM) | re(Y2) | im(Y2) |
| 1.080E+07 | 2.804E-03 | -4.960E-04 | 2.326E-04 | 2.767E-05 |
| 1.500E+07 | 2.727E-03 | -6.702E-04 | 2.313E-04 | 3.878E-05 |
| 2.084E+07 | 2.590E-03 | -8.844E-04 | 2.291E-04 | 5.470E-05 |

FAST AC ANALYSIS VCC=5.5 VOLTS ICC= 7.8625MA R-REF -20%

| FREQ | re(Y1) | im(Y1) | re(YP) | im(YP) |
|---|---|---|---|---|
| 1.080E+07 | 1.715E-05 | 9.460E-05 | 2.369E-05 | -1.278E-06 |
| 1.500E+07 | 3.216E-05 | 1.277E-04 | 2.349E-05 | -1.724E-06 |
| 2.084E+07 | 5.883E-05 | 1.681E-04 | 2.313E-05 | -2.269E-06 |
| FREQ | re(GM) | im(GM) | re(Y2) | im(Y2) |
| 1.080E+07 | 3.916E-03 | -7.100E-04 | 3.709E-04 | 3.273E-05 |
| 1.500E+07 | 3.803E-03 | -9.581E-04 | 3.697E-04 | 4.581E-05 |
| 2.084E+07 | 3.603E-03 | -1.261E-03 | 3.677E-04 | 6.446E-05 |

FIGURE 9a CMOS INVERTER SIMULATION NEAR 15MHZ
(Y parameters are in mhos.)

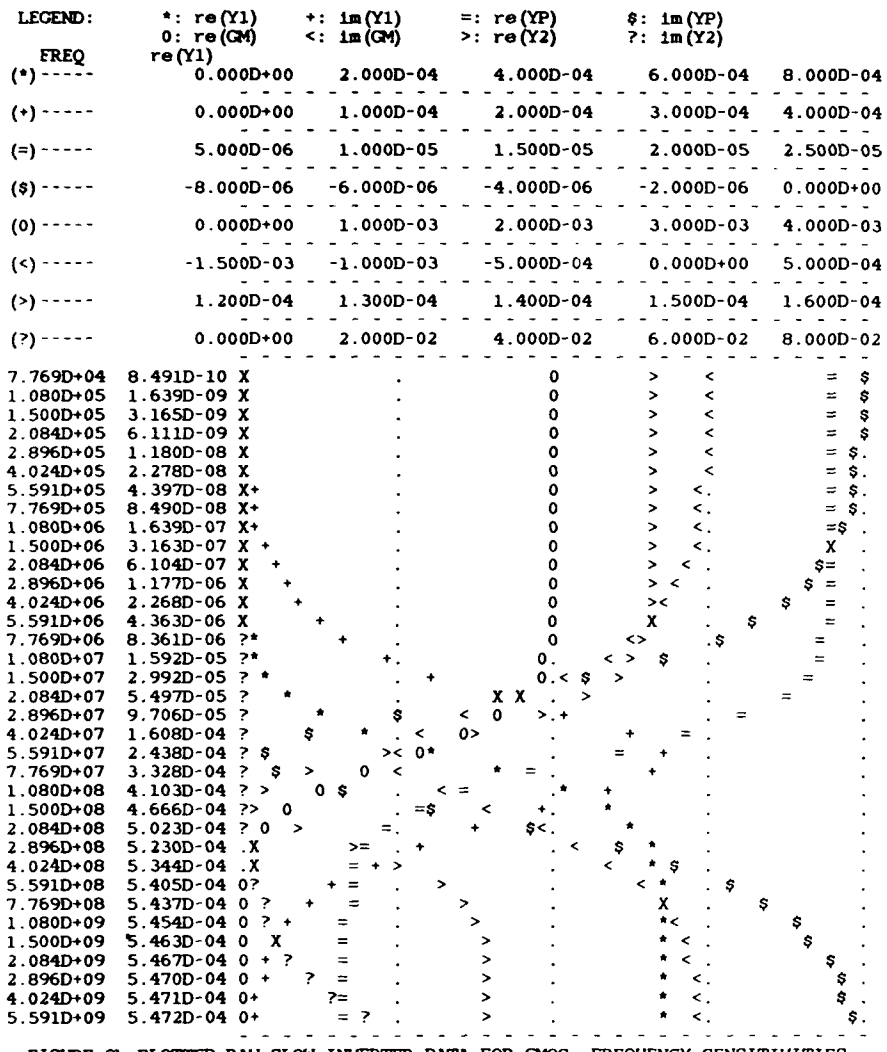



FIGURE 9b PLOTTED RAW SLOW INVERTER DATA FOR CMOS. FREQUENCY SENSITIVITIES BEGIN NEAR 4MHZ (ALL Y PARAMETERS IN MHOS.)

```
VCC 1 0 5.0 ---- POWER SUPPLY
.SUBCKT DEVICE 1 2 3 5
 XQ1 3 4 0 0 M111S2X4 12 (BIPOLAR TRANSISTOR DEVICE SPECIFICATION)
*Typical resistors are shown here. Range on R3 and R1 is 20%.
 R4 2 4 200.
 R2 3 2 10K
 R3 5 3 1.K
 R1 1 5 1.K
 C5 5 0 5P
 ENDS DEVICE
```

Tabulated Raw Inverter Data for Bipolar

SLOW AC ANALYSIS VCC=4.5V R1 and R3 = +20% ICC= 1.51MA

| FREQ | re(Y1) | im(Y1) | re(yp) | im(yp) |
|---|---|---|---|---|
| 1.080E+07 | 2.727E-04 | 1.052E-04 | 1.002E-04 | 4.495E-06 |
| 1.500E+07 | 2.780E-04 | 1.458E-04 | 1.003E-04 | 6.235E-06 |
| 2.084E+07 | 2.882E-04 | 2.015E-04 | 1.006E-04 | 8.636E-06 |

| FREQ | re(GM) | im(GM) | re(Y2) | im(Y2) |
|---|---|---|---|---|
| 1.080E+07 | 4.399E-02 | -2.401E-03 | 4.805E-04 | 1.728E-04 |
| 1.500E+07 | 4.387E-02 | -3.327E-03 | 4.978E-04 | 2.358E-04 |
| 2.084E+07 | 4.364E-02 | -4.599E-03 | 5.283E-04 | 3.173E-04 |

TYPICAL AC ANALYSIS VCC=5.0V TYPICAL R1 and R3 ICC= 2.046MA

| FREQ | re(Y1) | im(Y1) | re(yp) | im(yp) |
|---|---|---|---|---|
| 1.080E+07 | 2.717E-04 | 1.155E-04 | 1.002E-04 | 3.741E-06 |
| 1.500E+07 | 2.779E-04 | 1.601E-04 | 1.003E-04 | 5.188E-06 |
| 2.084E+07 | 2.898E-04 | 2.211E-04 | 1.005E-04 | 7.184E-06 |

| FREQ | re(GM) | im(GM) | re(Y2) | im(Y2) |
|---|---|---|---|---|
| 1.080E+07 | 5.509E-02 | -3.201E-03 | 5.742E-04 | 1.678E-04 |
| 1.500E+07 | 5.492E-02 | -4.435E-03 | 5.897E-04 | 2.300E-04 |
| 2.084E+07 | 5.459E-02 | -6.126E-03 | 6.178E-04 | 3.116E-04 |

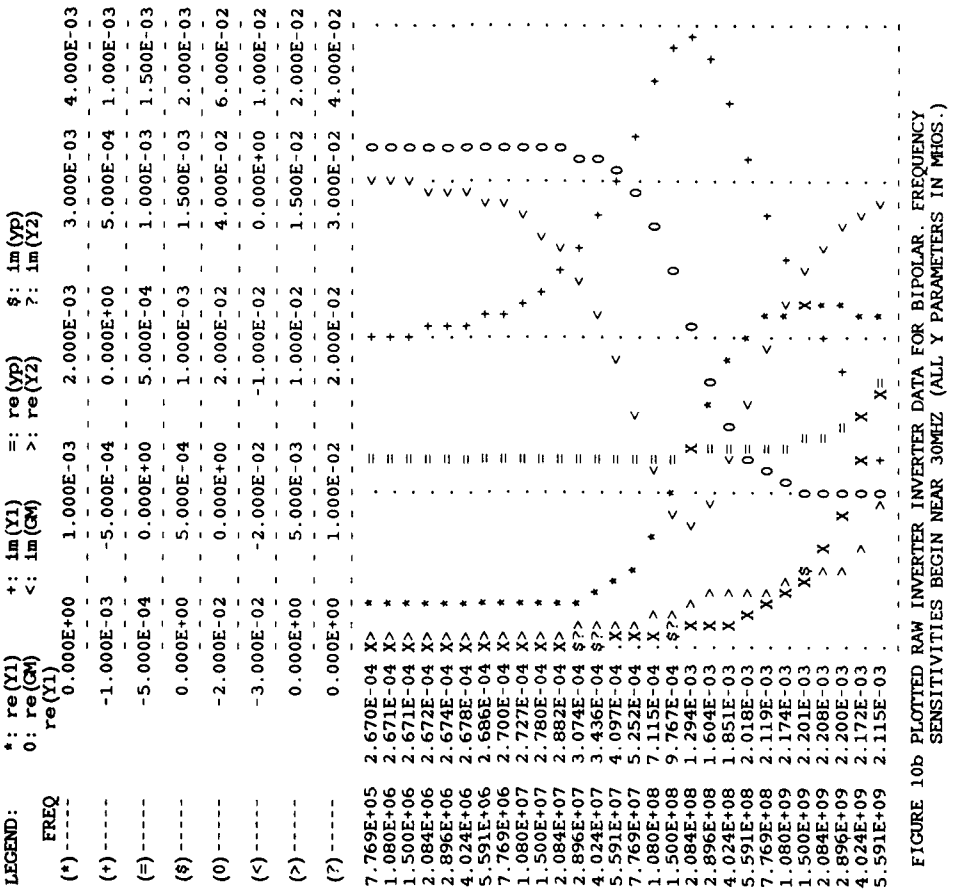
FAST AC ANALYSIS VCC=5.5V R1 and R3 = -20%. ICC= 2.84MA

| FREQ | re(Y1) | im(Y1) | re(yp) | im(yp) |
|---|---|---|---|---|
| 1.080E+07 | 2.739E-04 | 1.347E-04 | 1.002E-04 | 2.973E-06 |
| 1.500E+07 | 2.817E-04 | 1.865E-04 | 1.003E-04 | 4.121E-06 |
| 2.084E+07 | 2.965E-04 | 2.573E-04 | 1.005E-04 | 5.703E-06 |

| FREQ | re(GM) | im(GM) | re(Y2) | im(Y2) |
|---|---|---|---|---|
| 1.080E+07 | 6.604E-02 | -4.143E-03 | 7.115E-04 | 1.568E-04 |
| 1.500E+07 | 6.580E-02 | -5.736E-03 | 7.247E-04 | 2.158E-04 |
| 2.084E+07 | 6.534E-02 | -7.915E-03 | 7.492E-04 | 2.942E-04 |

10a BIPOLAR INVERTER SIMULATION NEAR 15MHZ

(Y parameters are in mhos.)




### 4.3 Design Implementations

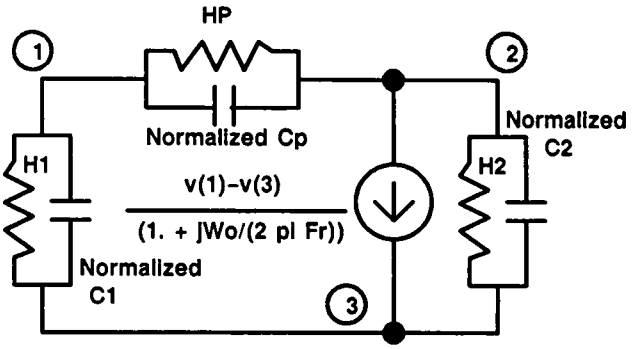
Computerized CMOS and BIPOLAR design results are given in figure 13. The figure is based on an arbitrary 1 pole roll off model for GM. The data in figs. 9 and 10 are normalized by the equivalent DC value of GM to get h1, h2, hp, al, and the normalized parasitic capacitances. A schematic with normalization formulas is shown in shown in figure 11. Fr is the equivalent roll off frequency of GM and GMN is the effective GM at DC. (Fr is infinity if GM has no imaginary part.) A one pole model for the inverter allows design estimates at frequencies other than where the data is taken. To find the scaled circuit, multiply all elements in figure 11 by GM and scale the components in figure 8 by (GM/GMN).

Our illustrative designs are done at 15MHZ. The value of re(GM), chosen from the ranges calculated in (32) and (41), is used to find GM at DC plus parasitics and device sizes. If one cannot change the inverter, its GM is the only allowed value. If this does not fall within the design range, another circuit should be used.

Figure 12 is the normalized version of figs. 9 and 10. Note that the CMOS normalized parasitics are about 5x as large as those of the bipolar circuit.

Since all designs in figure 13 were done by computer, only the results are shown. Equation numbers used to calculate them are noted along side. The computer also prints out the performance of each design over production. When no equation number is given, the values are based on the spread calculations.

CRAT is the fractional variation assumed for the external parts of C1, C2 and Cp. 'MIN.', 'TYP.' and 'MAX.' responses are calculated with this capacitor variation. 'INT. C1', 'INT. C2', 'INT. CP', 'INT. G1', 'INT. G2', and 'INT. GP' are the inverter's parasitic capacitors and conductances. The computer includes them when evaluating the production spreads.




AL= im(GM)/re(GM)

Fr(mhz) =-15./AL (Data taken at 15MHZ.)

2

GMN=(1.+AL )re(GM) (GMN is the equivalent GM at DC)

Normalized Cp= im(yp)/(2. pi 15.E6 GMN)

Normalized C1= im(y1)/(2. pi 15.E6 GMN)

Normalized C2= im(y2)/(2. pi 15.E6 GMN)

HP= re(yp)/GMN

H1= re(y1)/GMN

H2= re(y2)/GMN

FIGURE 11 NORMALIZED INVERTER WITH 1 POLE GM ROLL OFF

| min. (L) | typ. (T) | max. (H) |
|---|---|---|
| FR(L)\\\\=0.6236E+02 | FR(T)\\\\=0.6103E+02 | FR(H)\\\\=0.5954E+02 |
| CRAT(1,L)=0.1150E+01 | CRAT(1,T)=0.1000E+01 | CRAT(1,H)=0.8500E+00 |
| CRAT(2,L)=0.1150E+01 | CRAT(2,T)=0.1000E+01 | CRAT(2,H)=0.8500E+00 |
| CRAT(P,L)=0.1150E+01 | CRAT(P,T)=0.1000E+01 | CRAT(P,H)=0.8500E+00 |
| H1(L)\\\\=0.1487E-01 | H1(T)\\\\=0.1070E-01 | H1(H)\\\\=0.7952E-02 |
| H2(L)\\\\=0.7172E-01 | H2(T)\\\\=0.7999E-01 | H2(H)\\\\=0.9141E-01 |
| HP(L)\\\\=0.1147E-01 | HP(T)\\\\=0.8044E-02 | HP(H)\\\\=0.5808E-02 |
| CP(L)\\\\=-.1870E+02 | CP(T)\\\\=-.9922E+01 | CP(H)\\\\=-.4523E+01 |
| GMN(L)\\\\=0.2012E-02 | GMN(T)\\\\=0.2892E-02 | GMN(H)\\\\=0.4044E-02 |
| C1(L)\\\\=0.6560E+03 | C1(T)\\\\=0.4620E+03 | C1(H)\\\\=0.3350E+03 |
| C2(L)\\\\=0.1896E+03 | C2(T)\\\\=0.1423E+03 | C2(H)\\\\=0.1202E+03 |
| I/GMN(L)\=0.1588E+01 | I/GMN(T)\=0.1753E+01 | I/GMN(H)\=0.1944E+01 |

FIGURE 12a NORMALIZED CMOS INVERTER DATA CALCULATED FROM FIG. 9a USING FORMULAS FROM FIG. 11 (FR in MHZ, Cap. in Pfd., I in amps, GMN in MHOS)

| min(L) | typ. (T) | max. (H) |
|---|---|---|
| FR(L)\\\\=0.1978E+03 | FR(T)\\\\=0.1857E+03 | FR(H)\\\\=0.1721E+03 |
| CRAT(1,L)=0.1150E+01 | CRAT(1,T)=0.1000E+01 | CRAT(1,H)=0.8500E+00 |
| CRAT(2,L)=0.1150E+01 | CRAT(2,T)=0.1000E+01 | CRAT(2,H)=0.8500E+00 |
| CRAT(P,L)=0.1150E+01 | CRAT(P,T)=0.1000E+01 | CRAT(P,H)=0.8500E+00 |
| H1(L)\\\\=0.6301E-02 | H1(T)\\\\=0.5027E-02 | H1(H)\\\\=0.4249E-02 |
| H2(L)\\\\=0.1128E-01 | H2(T)\\\\=0.1067E-01 | H2(H)\\\\=0.1093E-01 |
| HP(L)\\\\=0.2273E-02 | HP(T)\\\\=0.1814E-02 | HP(H)\\\\=0.1513E-02 |
| CP(L)\\\\=0.1499E+01 | CP(T)\\\\=0.9958E+00 | CP(H)\\\\=0.6595E+00 |
| GMN(L)\\\\=0.4412E-01 | GMN(T)\\\\=0.5528E-01 | GMN(H)\\\\=0.6630E-01 |
| C1(L)\\\\=0.3506E+02 | C1(T)\\\\=0.3073E+02 | C1(H)\\\\=0.2985E+02 |
| C2(L)\\\\=0.5670E+02 | C2(T)\\\\=0.4415E+02 | C2(H)\\\\=0.3454E+02 |
| I/GMN(L)\=0.3422E-01 | I/GMN(T)\=0.3702E-01 | I/GMN(H)\=0.4284E-01 |

FIGURE 12b NORMALIZED BIPOLAR INVERTER DATA CALCULATED FROM FIG.10a USING FORMULAS FROM FIG. 11. (FR in MHZ, Cap. in Pfd., I in amps, GMN in MHOS)

Pierce Osc. Design Results

| | CMOS | | | BIPOLAR | | | |
|---|---|---|---|---|---|---|---|
| | min | typ(or chosen) | max | min | typ(chosen) | max | Eqn. # |
| most neg. r(ohm) | | -189.1 | | | -281.7 | | (32), (54) |
| value of r | -99.07 | -140. | -111.8 | -178.2 | -220. | -192.1 | |
| value of z | -216.2 | -315.7 | -449.5 | -228.1 | -360.3 | -530.2 | (49) |
| GM des. lim.(mho) | .0151 | .05985 | .1046 | .0315 | .1521 | .2727 | (32) |
| GM var. | .0416 | .05985 | .08369 | .1214 | .1521 | .1824 | |
| C2 des. lim.(pfd) | 83.84 | 253.8 | 423.8 | 76.43 | 333.6 | 590.7 | (30) |
| C2 variation | 290. | 253.8 | 218.6 | 382.8 | 333.6 | 284.2 | |
| Int. C2 | 7.895 | 8.517 | 10.06 | 6.88 | 6.72 | 6.3 | |
| C1 Variation | 65.73 | 61.06 | 56.43 | 114.6 | 100.7 | 87.04 | (28a) |
| Int. C1 | 27.32 | 27.65 | 28.036 | 4.256 | 4.674 | 5.44 | |
| Cp(C0) Variation | 11.82 | 12. | 12.22 | 12.03 | 12. | 11.97 | |
| Int. Cp | -.779 | -.5938 | -.3785 | .182 | .151 | .121 | |
| G2 Var.(mho) | 2.99E-3 | 4.79E-3 | 7.65E-3 | 1.37E-3 | 1.62E-3 | 1.99E-3 | |
| int. G2(mho) | 2.99E-3 | 4.79E-3 | 7.65E-3 | 1.37E-3 | 1.62E-3 | 1.99E-3 | |
| G1 Variation | 1.71E-3 | 1.73E-3 | 1.76E-3 | 2.32E-3 | 2.32E-3 | 2.33E-3 | (28b) |
| Int. G1 | .619E-3 | .64E-3 | .666E-3 | 7.65E-4 | 7.65E-4 | 7.65E-4 | |
| Gp(G0) Var. | .496E-3 | .5E-3 | .505E-3 | .5E-3 | .5E-3 | .5E-3 | |
| Int. Gp | .478E-3 | .481E-3 | .486E-3 | 2.8E-4 | 2.8E-4 | 2.8E-4 | |
| Supply Curr(MA) | 66.1 | 104.9 | 162. | 4.15 | 5.631 | 7.814 | |

Miller Osc. Design Results

| | CMOS | | | BIPOLAR | | | |
|---|---|---|---|---|---|---|---|
| | min | typ(or chosen) | max | min | typ(chosen) | max | Eqn. # |
| most neg. r(ohm) | | -95.59 | | | -253.3 | | (41), (54) |
| value of r | -37.82 | -60. | -50.64 | -176.9 | -200. | -197.9 | |
| value of z | -272. | -241.1 | -201.4 | -413.6 | -350.5 | -297. | (49) |
| GM des. lim.(mho) | .00141 | .0196 | .0378 | 6.07e-4 | .0207 | .0408 | (41) |
| GM var. | .01364 | .0196 | .02741 | .01653 | .0207 | .2484 | |
| L2 des. lim.(uh) | .4995 | .8177 | 2.254 | 1.296 | 2.089 | 5.374 | (39) |
| L2 variation | .7083 | .8177 | .9697 | 1.813 | 2.089 | 2.462 | |
| Int. C2(pfd.) | 2.59 | 2.79 | 3.29 | .937 | .914 | .858 | |
| CP variation | 16.54 | 14.41 | 12.29 | 5.162 | 4.488 | 3.813 | (37b) |
| Int. CP | -.26 | -.194 | -.124 | .025 | .0206 | .0164 | |
| C1(C0) Var. | 11.89 | 12. | 12.13 | 11.94 | 12. | 12.11 | |
| Int. C1 | 8.95 | 9.06 | 9.18 | .58 | .636 | .74 | |
| G2(mho) var. | 9.78e-4 | 1.57e-3 | 2.5e-3 | 1.86e-4 | 2.21e-4 | 2.72e-4 | (37a) |
| Int. G2(mho) | 9.78e-4 | 1.57e-3 | 2.5e-3 | 1.86e-4 | 2.21e-4 | 2.72e-4 | |
| G1(G0) var. | 4.93e-4 | 5.e-4 | 5.08e-4 | 5.e-4 | 5.e-4 | 5.e-4 | |
| Int. G1 | 2.03e-4 | 2.1e-4 | 2.18e-4 | 1.042e-4 | 1.041e-4 | 1.055e-4 | |
| Gp variation | 2.84e-4 | 2.85e-4 | 2.87e-4 | 6.76e-5 | 6.76e-5 | 6.76e-5 | |
| Int. Gp | 1.56e-4 | 1.58e-4 | 1.59e-4 | 3.76e-5 | 3.76e-5 | 3.76e-5 | |
| Supply Curr.(ma) | 21.7 | 34.4 | 53.3 | .566 | .767 | 1.064 | |

FIGURE 13 PIERCE AND MILLER OSC. DESIGN RESULTS. THE SYMBOLS ARE FROM SECT 3

All designs are based on:

a)Typical device parameters
b)Oscillation frequency of 15mhz.
c)Min. parallel biasing resistor of 2k (G0=5.E-4mho.)
d)12pfd. capacitance in parallel with the resonant circuit.(Co=12pfd.)
e)The optimum value of z according to (49)
f)The design methods of section 3.

For the rest of section 4, GM denotes its effective value at DC as calculated in fig. 11.

### 4.4 Discussion

The best llator design has:

a) The most negative value of r over production.
b) The lowest internal normalized parasitic capacitors and conductances.
c) The lowest normalized power.

From figure 13, the bipolar design clearly wins on all counts. Note that the negative value of r in the CMOS Miller design is practically unacceptable, and that the limiting value of r (95.59ohms) calculated from (41) and (54) is the lowest among the 4 candidates. INT Co -9.06pfd., is a large part of the assumed 12pfd. limit, leaving only 2.94pfd. for the crystal and package. The CMOS Pierce design, although inferior to both bipolar designs is usable if supply current of -100ma is not an issue.

From the data in figure 12, one can see that the normalized BIPOLAR stray capacitors are about 1/5 those for CMOS and that I/GMN, the supply current per mho of GM is about 50x as high in CMOS as it is in bipolar. This can be reduced somewhat by not using a current mirror, or by using large CMOS transistors as was done in [1] to design a watch oscillator running at 32khz. The design biased the transistors in the sub-threshold region where the I/V characteristics are exponential. This approach may not be practical at higher frequencies. The conclusion one comes to is that a Bipolar llator biased in the active region is more robust than a similar CMOS llator and that CMOS can be made to work if necessary.

In most cases, external G1 and Gp are required to make the design meet the exact specification on Z. (See Fig. 13) However, r will be more negative if they are left out.

### 4.5 SMALL SIGNAL AC SIMULATIONS

Spice can be used to plot the value of Z over frequency. The Spice network is shown in figure 14. Inverter scaling is simulated with artificial controlled current sources at the I/O nodes. N is the required scale factor. Since external inductors and conductances may throw off the inverter's dc biasing, parallel RC and LC structures are converted to series equivalents.

If re(Z) is not selective enough, C1, C2, or CP can become frequency sensitive by means of a tank circuit. Since this technique is well known[4], it will not be pursued further. Figure 15 is a composite small signal AC plot of z and r as a function of frequency for the typical case in all four designs. The plots show that all four designs can oscillate only near 15mhz. This result is not true in all cases.

### 4.6 TRANSIENT ANALYSIS

Output waveforms can be plotted by driving the llator with a sinusoidal current source. These waveforms can be analysed using the Fourier transform capabilities of Spice to see if the oscillations are still growing when the crystal is operating at a particular current level[16].

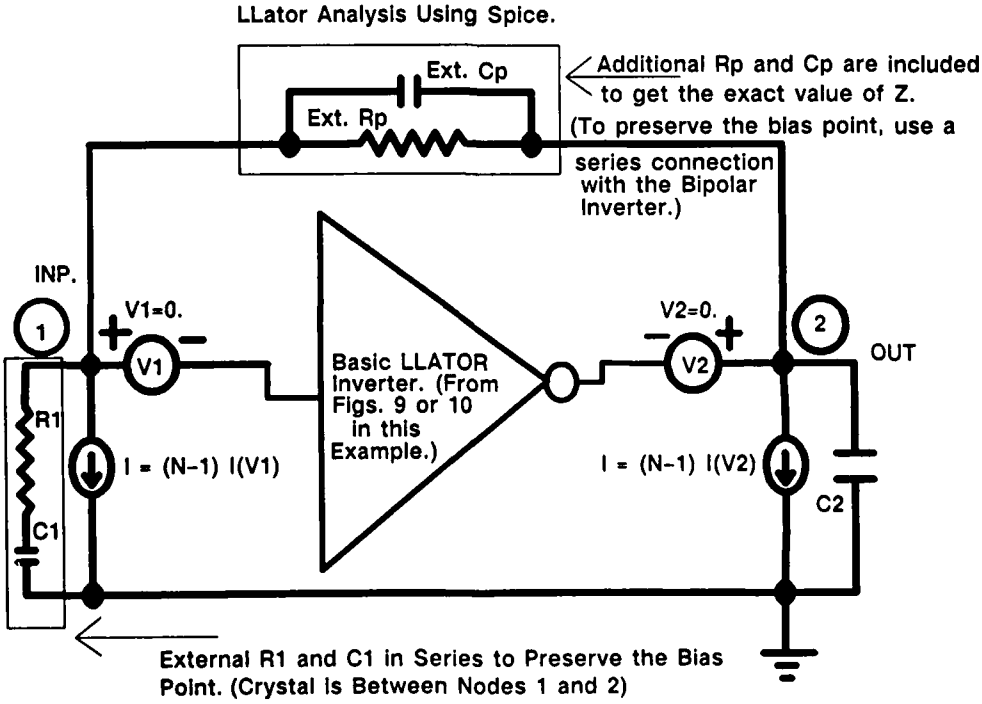



Figure 14 Spice Network to used to Plot Frequency Response. (Crystal is Between Nodes 1 and 2 for Pierce and Between Gnd. and one of the Nodes For the Miller Oscillator.)

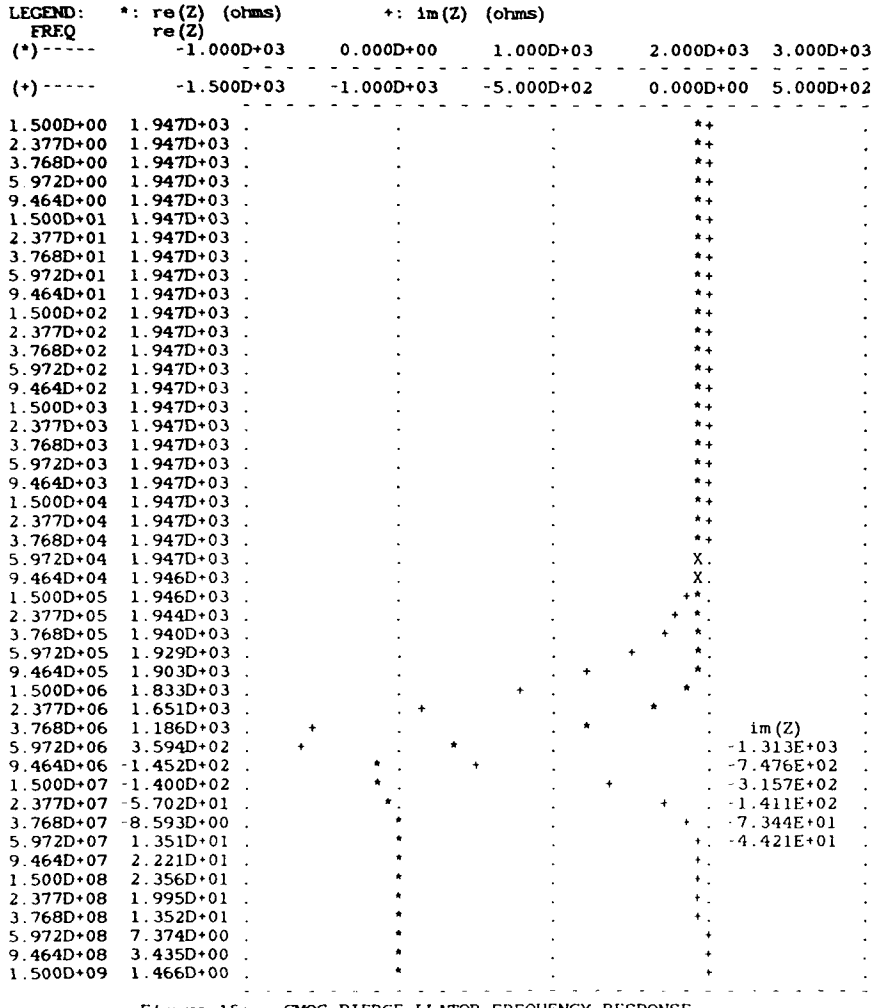



Figure 15a CMOS PIERCE LLATOR FREQUENCY RESPONSE

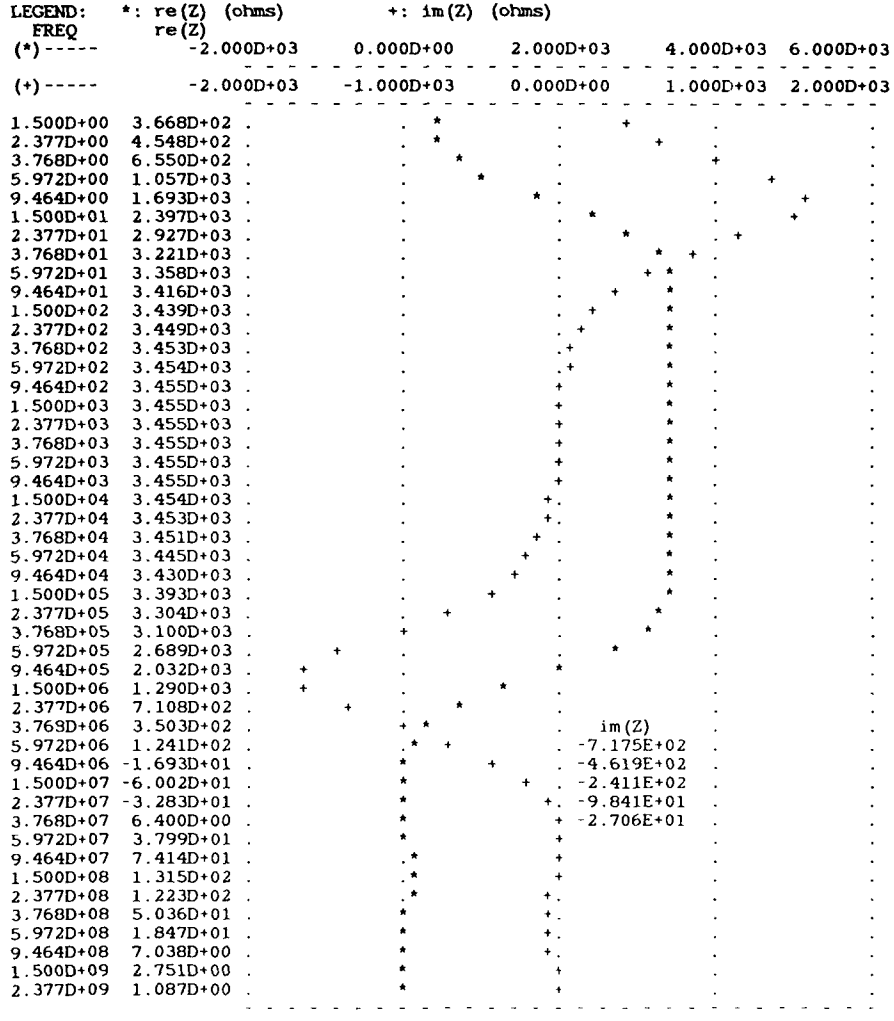



Figure 15b CMOS MILLER LLATOR FREQUENCY RESPONSE

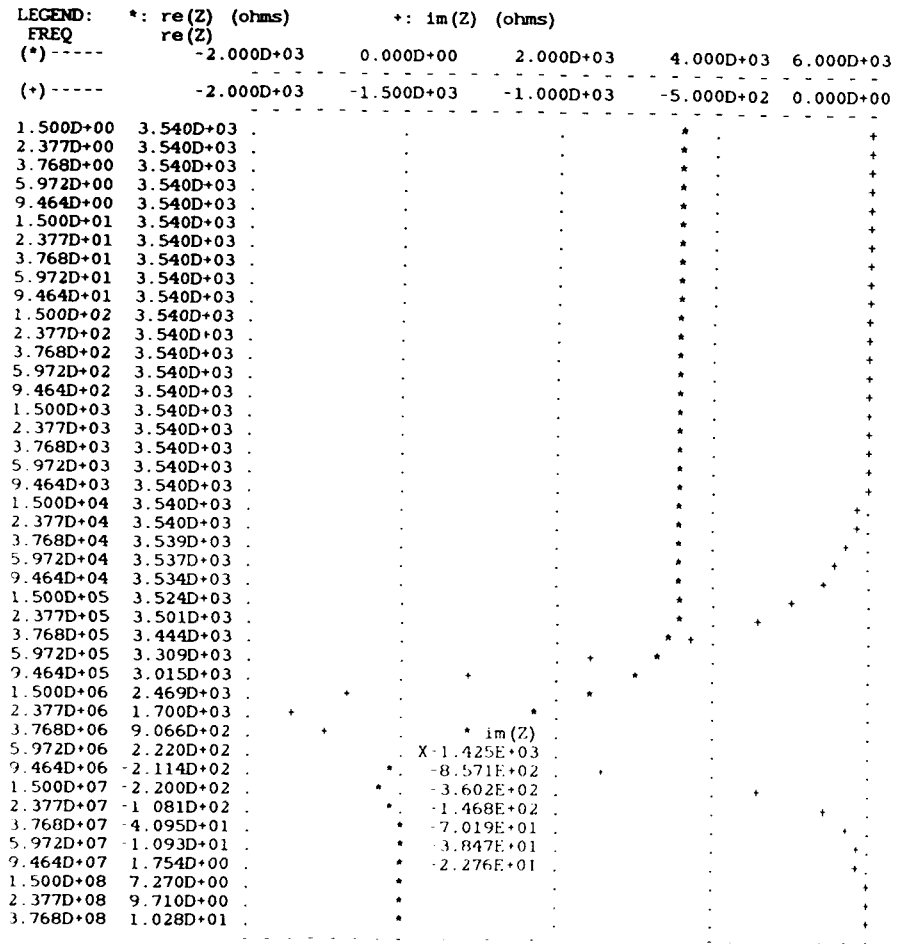



Figure 15c BIPOLAR PIERCE LLATOR FREQUENCY RESPONSE

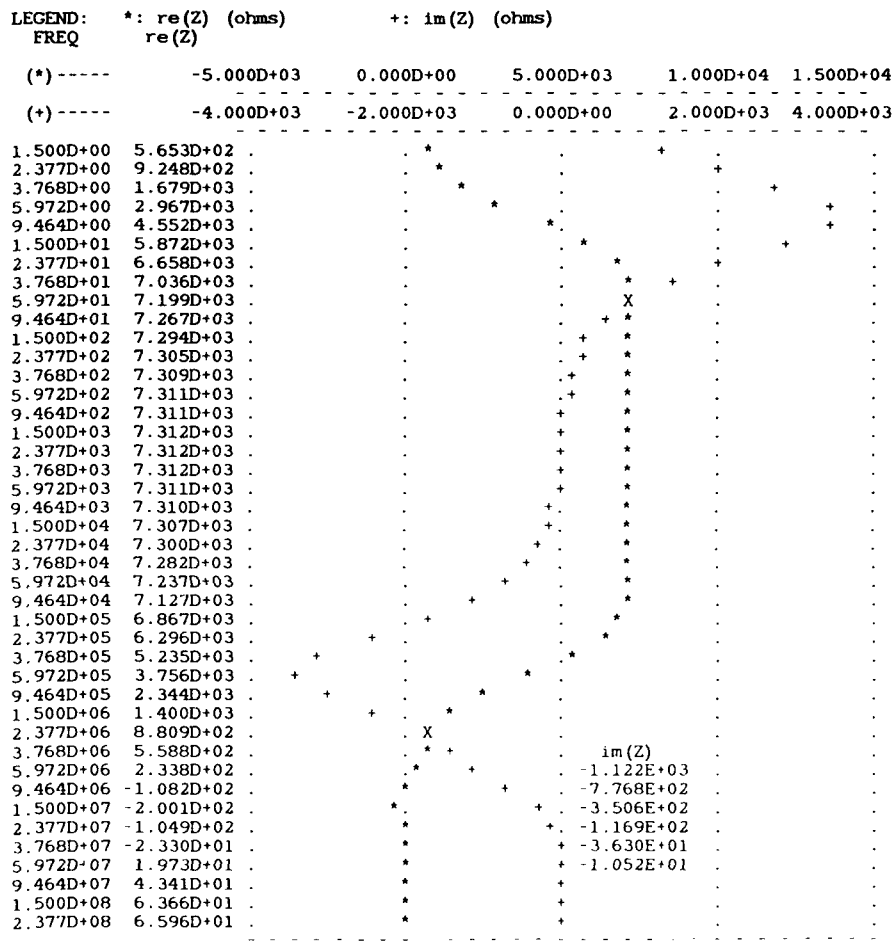



Figure 15d BIPOLAR MILLER LLATOR FREQUENCY RESPONSE

Figure 16 is a table of the oscillation properties vs. current level. The table was generated by doing consecutive SPICE runs of the typical CMOS Pierce LLATOR circuit from sect. 4.3 starting with .01ma and multiplying by $10^{.1}$ each time. In figure 16, 'INP' represents node 1 of fig. 8a, 'OUT' represents node 2 and 'DIFF' denotes v(2)-v(1). Z is the impedance across the resonant circuit. P-P means peak to peak voltage and 'offset' is the shift of the average value from the bias point. For this circuit, re(Z) becomes positive and the P-P voltages aproach the 5v rails at about 20ma. The offsets are noticable at 2ma. Such offsets are exacerbated if non-linear devices are used as resistors.

Since the initial turn on of any current source causes a transient disturbance to the LLATOR bias points, fourier analysis has to be performed after equilibrium has been re-established. A quick simulation shows that in this case a disturbance decays from about 2v to 3.e-5v in 500ns. (-8 cycles at 15mhz.) Therefore, all Fourier analysis was done after an 8 cycle delay.

Due to non-linearities, the bias point during oscillation can be different from its quiescent value. This affects Z. The rate of change is controlled by LLATOR time constants. Faster growing oscillations which barely satisfy (6) and (20) will therefore exhibit a warm up curve which is not time scalable to the high Q case. Similarly, all non resonant LLATOR poles including spurious crystal modes will affect the warm up characteristic in one way or another. Sustained bias point movements can be particularly dangerous because a changing Z modulates the frequency and amplitude. This can cause squegging and phase noise.[4] However, if the oscillation growth time constant is much longer than all others and noises are small, steady state waveform analysis applies.

Figure 17 contains plots of the waveform properties as a function of normalized time assuming slow growth and a 40ohm crystal resistance. (To find actual time, multiply the scale by 2QR/Wo = 2L.) Linear variation is assumed between consecutive values in figure 16. A formula is derived in appendix 1 which describes the time for the oscillations to grow from one current level to the next. Assuming that the initial current is .02ma, the formula is used to calculate the values in figure 17. Figures 17a-d are then plotted from the calculations using 'SPICE' as a paintbrush.

Fig. 17a shows current growth and im(Z) variation vs normalized time. This means that oscillation frequency varies during growth and settles to a final value. Figures 17b-17d show oscillation envlopes as seen from various points in the circuit. Envelopes were plotted in this form so that they could be related qualitatively to published photos. Such photos were published by A. Rusnyak[19] and are included as figures 17e-f for comparison. His results look like figs. 17c-d once oscillation starts.

The computer runs were done on a CRAY. Each steady state current level run took about 9.1 sec. I could not get the direct simulation to work in less than two hours even though Q was taken to be 212, a 5% error according to (7).

All steady state Fourier SPICE data is a property of the LLATOR only. R was added to the growth simulation to represent a 40ohm crystal. Other growth simulations can be run with different crystal resistors without changing anyting else.

The engineer can now use Spice to analyse his oscillator in great detail before building a breadboard or commiting the design to silicon.

| RUN # | CURR. LEV. (ma) | re(Z) (K) | im(Z) (K) | P-P INP (V) | P-P OUT (V) |
|---|---|---|---|---|---|
| 0.000E+00 | -3.891E-09 | -1.400E-01 | -3.157E-01 | -1.309E-09 | -7.417E-10 |
| 1.000E+00 | 1.000E-02 | -1.402E-01 | -3.157E-01 | 2.595E-03 | 6.251E-03 |
| 2.000E+00 | 1.259E-02 | -1.402E-01 | -3.157E-01 | 3.267E-03 | 7.869E-03 |
| 3.000E+00 | 1.585E-02 | -1.402E-01 | -3.157E-01 | 4.113E-03 | 9.907E-03 |
| 4.000E+00 | 1.995E-02 | -1.402E-01 | -3.157E-01 | 5.178E-03 | 1.247E-02 |
| 5.000E+00 | 2.512E-02 | -1.402E-01 | -3.157E-01 | 6.520E-03 | 1.570E-02 |
| 6.000E+00 | 3.163E-02 | -1.401E-01 | -3.156E-01 | 8.208E-03 | 1.977E-02 |
| 7.000E+00 | 3.982E-02 | -1.402E-01 | -3.157E-01 | 1.033E-02 | 2.490E-02 |
| 8.000E+00 | 5.013E-02 | -1.402E-01 | -3.157E-01 | 1.301E-02 | 3.134E-02 |
| 9.000E+00 | 6.311E-02 | -1.402E-01 | -3.158E-01 | 1.638E-02 | 3.945E-02 |
| 1.000E+01 | 7.944E-02 | -1.402E-01 | -3.157E-01 | 2.062E-02 | 4.967E-02 |
| 1.100E+01 | 1.000E-01 | -1.402E-01 | -3.157E-01 | 2.595E-02 | 6.252E-02 |
| 1.200E+01 | 1.259E-01 | -1.402E-01 | -3.157E-01 | 3.267E-02 | 7.869E-02 |
| 1.300E+01 | 1.585E-01 | -1.402E-01 | -3.157E-01 | 4.114E-02 | 9.908E-02 |
| 1.400E+01 | 1.996E-01 | -1.402E-01 | -3.157E-01 | 5.179E-02 | 1.247E-01 |
| 1.500E+01 | 2.512E-01 | -1.402E-01 | -3.157E-01 | 6.520E-02 | 1.570E-01 |
| 1.600E+01 | 3.163E-01 | -1.401E-01 | -3.156E-01 | 8.208E-02 | 1.977E-01 |
| 1.700E+01 | 3.982E-01 | -1.402E-01 | -3.157E-01 | 1.033E-01 | 2.489E-01 |
| 1.800E+01 | 5.013E-01 | -1.402E-01 | -3.157E-01 | 1.301E-01 | 3.134E-01 |
| 1.900E+01 | 6.311E-01 | -1.401E-01 | -3.156E-01 | 1.638E-01 | 3.944E-01 |
| 2.000E+01 | 7.945E-01 | -1.401E-01 | -3.156E-01 | 2.061E-01 | 4.964E-01 |
| 2.100E+01 | 1.000E+00 | -1.401E-01 | -3.155E-01 | 2.597E-01 | 6.248E-01 |
| 2.200E+01 | 1.259E+00 | -1.400E-01 | -3.153E-01 | 3.269E-01 | 7.861E-01 |
| 2.300E+01 | 1.585E+00 | -1.399E-01 | -3.151E-01 | 4.116E-01 | 9.889E-01 |
| 2.400E+01 | 1.997E+00 | -1.398E-01 | -3.148E-01 | 5.184E-01 | 1.244E+00 |
| 2.500E+01 | 2.515E+00 | -1.394E-01 | -3.142E-01 | 6.531E-01 | 1.563E+00 |
| 2.600E+01 | 3.165E+00 | -1.388E-01 | -3.131E-01 | 8.225E-01 | 1.960E+00 |
| 2.700E+01 | 3.984E+00 | -1.361E-01 | -3.111E-01 | 1.033E+00 | 2.428E+00 |
| 2.800E+01 | 5.015E+00 | -1.232E-01 | -3.066E-01 | 1.283E+00 | 2.884E+00 |
| 2.900E+01 | 6.313E+00 | -1.019E-01 | -2.988E-01 | 1.578E+00 | 3.301E+00 |
| 3.000E+01 | 7.947E+00 | -7.704E-02 | -2.872E-01 | 1.932E+00 | 3.666E+00 |
| 3.100E+01 | 1.000E+01 | -5.325E-02 | -2.727E-01 | 2.374E+00 | 3.982E+00 |
| 3.200E+01 | 1.259E+01 | -3.235E-02 | -2.558E-01 | 2.992E+00 | 4.242E+00 |
| 3.300E+01 | 1.585E+01 | -1.490E-02 | -2.380E-01 | 3.814E+00 | 4.422E+00 |
| 3.400E+01 | 1.997E+01 | -7.458E-04 | -2.208E-01 | 4.877E+00 | 4.528E+00 |
| 3.500E+01 | 2.515E+01 | 1.052E-02 | -2.051E-01 | 6.197E+00 | 4.583E+00 |
| 3.600E+01 | 3.166E+01 | 1.914E-02 | -1.920E-01 | 7.869E+00 | 4.642E+00 |
| 3.700E+01 | 3.984E+01 | 2.628E-02 | -1.803E-01 | 9.986E+00 | 4.718E+00 |
| 3.800E+01 | 5.018E+01 | 3.130E-02 | -1.708E-01 | 1.267E+01 | 4.821E+00 |
| 3.900E+01 | 6.319E+01 | 3.571E-02 | -1.623E-01 | 1.607E+01 | 4.956E+00 |
| 4.000E+01 | 7.953E+01 | 3.964E-02 | -1.547E-01 | 2.034E+01 | 5.125E+00 |
| 4.100E+01 | 1.001E+02 | 4.156E-02 | -1.459E-01 | 2.558E+01 | 5.292E+00 |

FIGURE 16a OSCILLATION PROPERTIES VS. CRYSTAL CURRENT LEVEL

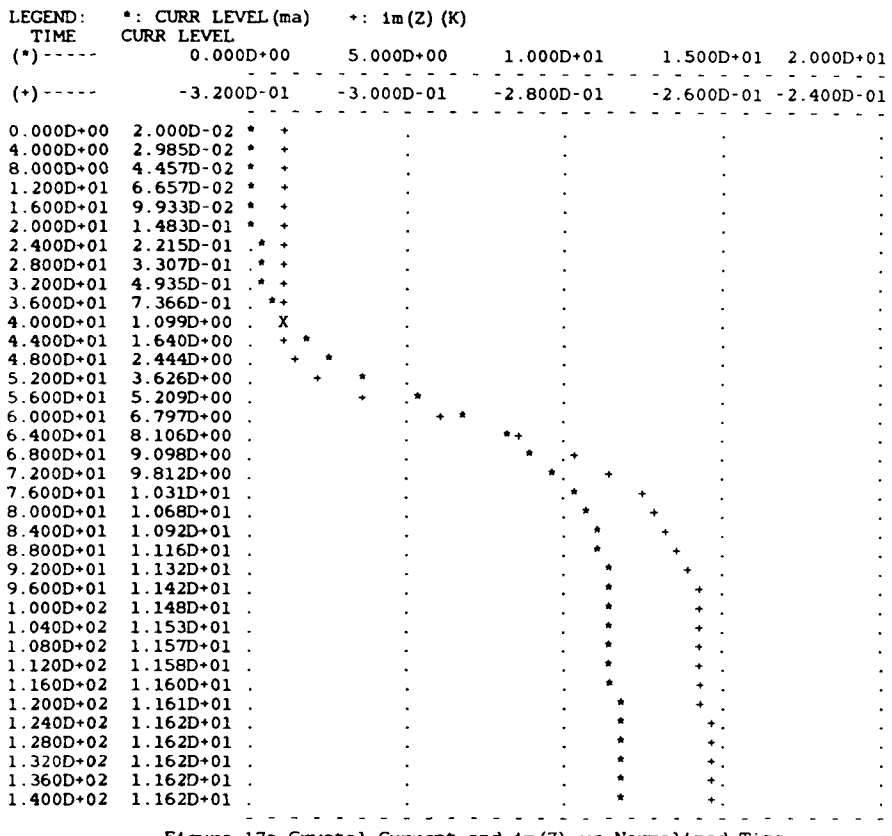



Figure 17a Crystal Current and im(Z) vs Normalized Time
(To find actual time, multiply the scale by 2QR/Wo=2L)

| RUN # | CURR. LEV. (ma) | P-P DIFF (v) | INP. OFFST (v) | OUTP. OFFST (v) | DIFF. OFFST (v) |
|---|---|---|---|---|---|
| 0.000E+00 | -3.891E-09 | 9.490E-10 | 2.486E-09 | 2.091E-09 | -6.980E-10 |
| 1.000E+00 | 1.000E-02 | 3.454E-03 | -1.528E-07 | -1.635E-07 | 4.491E-08 |
| 2.000E+00 | 1.259E-02 | 4.348E-03 | -2.384E-07 | -2.956E-07 | -2.420E-08 |
| 3.000E+00 | 1.585E-02 | 5.474E-03 | -4.136E-07 | -4.244E-07 | -2.045E-08 |
| 4.000E+00 | 1.995E-02 | 6.892E-03 | -6.445E-07 | -6.489E-07 | -1.150E-08 |
| 5.000E+00 | 2.512E-02 | 8.677E-03 | -1.043E-06 | -1.040E-06 | 2.515E-08 |
| 6.000E+00 | 3.163E-02 | 1.092E-02 | -1.651E-06 | -1.660E-06 | -1.302E-07 |
| 7.000E+00 | 3.982E-02 | 1.375E-02 | -2.646E-06 | -2.505E-06 | 2.692E-07 |
| 8.000E+00 | 5.013E-02 | 1.731E-02 | -4.170E-06 | -4.365E-06 | 2.763E-08 |
| 9.000E+00 | 6.311E-02 | 2.179E-02 | -6.614E-06 | -6.740E-06 | 7.914E-08 |
| 1.000E+01 | 7.944E-02 | 2.743E-02 | -1.042E-05 | -1.036E-05 | 1.512E-07 |
| 1.100E+01 | 1.000E-01 | 3.453E-02 | -1.655E-05 | -1.641E-05 | -1.015E-08 |
| 1.200E+01 | 1.259E-01 | 4.346E-02 | -2.644E-05 | -2.632E-05 | -1.226E-07 |
| 1.300E+01 | 1.585E-01 | 5.472E-02 | -4.170E-05 | -4.200E-05 | 2.799E-07 |
| 1.400E+01 | 1.996E-01 | 6.887E-02 | -6.642E-05 | -6.593E-05 | -3.469E-08 |
| 1.500E+01 | 2.512E-01 | 8.670E-02 | -1.051E-04 | -1.056E-04 | 1.324E-07 |
| 1.600E+01 | 3.163E-01 | 1.091E-01 | -1.669E-04 | -1.666E-04 | -3.810E-07 |
| 1.700E+01 | 3.982E-01 | 1.373E-01 | -2.649E-04 | -2.647E-04 | -2.077E-07 |
| 1.800E+01 | 5.013E-01 | 1.727E-01 | -4.194E-04 | -4.188E-04 | 4.201E-07 |
| 1.900E+01 | 6.311E-01 | 2.173E-01 | -6.642E-04 | -6.631E-04 | 1.477E-06 |
| 2.000E+01 | 7.945E-01 | 2.734E-01 | -1.055E-03 | -1.051E-03 | -3.964E-07 |
| 2.100E+01 | 1.000E+00 | 3.439E-01 | -1.669E-03 | -1.673E-03 | -4.375E-06 |
| 2.200E+01 | 1.259E+00 | 4.322E-01 | -2.652E-03 | -2.653E-03 | -3.339E-06 |
| 2.300E+01 | 1.585E+00 | 5.431E-01 | -4.206E-03 | -4.209E-03 | -6.837E-06 |
| 2.400E+01 | 1.997E+00 | 6.817E-01 | -6.669E-03 | -7.212E-03 | -5.390E-04 |
| 2.500E+01 | 2.515E+00 | 8.549E-01 | -1.062E-02 | -1.103E-02 | -4.120E-04 |
| 2.600E+01 | 3.165E+00 | 1.070E+00 | -1.692E-02 | -1.694E-02 | -1.981E-05 |
| 2.700E+01 | 3.984E+00 | 1.327E+00 | -2.457E-02 | -2.917E-02 | -4.589E-03 |
| 2.800E+01 | 5.015E+00 | 1.616E+00 | -2.373E-02 | -3.772E-02 | -1.402E-02 |
| 2.900E+01 | 6.313E+00 | 1.955E+00 | -1.472E-02 | -3.572E-02 | -2.103E-02 |
| 3.000E+01 | 7.947E+00 | 2.339E+00 | -1.073E-03 | -2.908E-02 | -2.799E-02 |
| 3.100E+01 | 1.000E+01 | 2.771E+00 | 1.487E-02 | -1.879E-02 | -3.370E-02 |
| 3.200E+01 | 1.259E+01 | 3.241E+00 | 3.159E-02 | -1.018E-02 | -4.180E-02 |
| 3.300E+01 | 1.585E+01 | 3.738E+00 | 4.417E-02 | -9.631E-03 | -5.388E-02 |
| 3.400E+01 | 1.997E+01 | 4.355E+00 | 5.114E-02 | -1.443E-02 | -6.564E-02 |
| 3.500E+01 | 2.515E+01 | 5.089E+00 | 6.458E-02 | -2.080E-02 | -8.537E-02 |
| 3.600E+01 | 3.166E+01 | 6.013E+00 | 8.727E-02 | -2.623E-02 | -1.135E-01 |
| 3.700E+01 | 3.984E+01 | 7.179E+00 | 1.149E-01 | -2.726E-02 | -1.424E-01 |
| 3.800E+01 | 5.018E+01 | 8.635E+00 | 1.567E-01 | -2.641E-02 | -1.830E-01 |
| 3.900E+01 | 6.319E+01 | 1.045E+01 | 2.091E-01 | -2.280E-02 | -2.320E-01 |
| 4.000E+01 | 7.953E+01 | 1.269E+01 | 2.715E-01 | -1.843E-02 | -2.907E-01 |
| 4.100E+01 | 1.001E+02 | 1.541E+01 | 3.442E-01 | -1.811E-02 | -3.620E-01 |

FIGURE 16b OSCILLATION PROPERTIES VS. CRYSTAL CURRENT LEVEL

```
LEGEND:   *: OUTPUT OFFSET(v)     +: OUTPUT PEAK(v)     =: OUTPUT VALLEY(v)
 TIME     OUTPUT OFFSET
(*+=)---        -2.500D+00    -1.250D+00     0.000D+00     1.250D+00  2.500D+00
0.000D+00 -6.543D-07 .                .              X              .          .
4.000D+00 -1.493D-06 .                .              X              .          .
8.000D+00 -3.364D-06 .                .              X              .          .
1.200D+01 -7.509D-06 .                .              X              .          .
1.600D+01 -1.620D-05 .                .              X              .          .
2.000D+01 -3.711D-05 .                .             =*+             .          .
2.400D+01 -8.274D-05 .                .             =*+             .          .
2.800D+01 -1.838D-04 .                .             =*+             .          .
3.200D+01 -4.071D-04 .                .            = * +            .          .
3.600D+01 -9.136D-04 .                .           =  *  +           .          .
4.000D+01 -2.047D-03 .                .          =   *   +          .          .
4.400D+01 -4.608D-03 .                .        =     *     +        .          .
4.800D+01 -1.050D-02 .                .     =        *        +     .          .
5.200D+01 -2.383D-02 .                .  =           *           +  .          .
5.600D+01 -3.743D-02 .             =  .              *              +          .
6.000D+01 -3.375D-02 .          =     .              *              +          .
6.400D+01 -2.829D-02 .       =        .              *              +          .
6.800D+01 -2.332D-02 .    =           .              *              +          .
7.200D+01 -1.975D-02 .    =           .              *              +          .
7.600D+01 -1.778D-02 .  =             .              *              +          .
8.000D+01 -1.656D-02 .  =             .              *              +          .
8.400D+01 -1.575D-02 .  =             .              *              +          .
8.800D+01 -1.495D-02 .  =             .              *              +          .
9.200D+01 -1.442D-02 .  =             .              *              +          .
9.600D+01 -1.409D-02 .  =             .              *              +          .
1.000D+02 -1.387D-02 .  =             .              *              +          .
1.040D+02 -1.371D-02 .  =             .              *              +          .
1.080D+02 -1.361D-02 .  =             .              *              +          .
1.120D+02 -1.353D-02 .  =             .              *              +          .
1.160D+02 -1.348D-02 .  =             .              *              +          .
1.200D+02 -1.345D-02 .  =             .              *              +          .
1.240D+02 -1.342D-02 .  =             .              *              +          .
1.280D+02 -1.341D-02 .  =             .              *              +          .
1.320D+02 -1.341D-02 .  =             .              *              +          .
1.360D+02 -1.341D-02 .  =             .              *              +          .
1.400D+02 -1.341D-02 .  =             .              *              +          .
```

Figure 17c Output Voltage Growth Envelope and Offset vs Normalized Time
(To find actual time, multiply the scale by 2QR/Wo=2L)

```
LEGEND:   *: INPUT OFFSET(v)     +: INPUT PEAK(v)     =: INPUT VALLEY(v)
 TIME     INP. OFFSET
(*+=)---        -2.000D+00    -1.000D+00     0.000D+00     1.000D+00  2.000D+00
0.000D+00 -6.506D-07 .                .              X              .          .
4.000D+00 -1.487D-06 .                .              X              .          .
8.000D+00 -3.352D-06 .                .              X              .          .
1.200D+01 -7.425D-06 .                .              X              .          .
1.600D+01 -1.634D-05 .                .              X              .          .
2.000D+01 -3.693D-05 .                .              X              .          .
2.400D+01 -8.280D-05 .                .              X              .          .
2.800D+01 -1.841D-04 .                .             =*+             .          .
3.200D+01 -4.075D-04 .                .             =*+             .          .
3.600D+01 -9.165D-04 .                .             =*+             .          .
4.000D+01 -2.045D-03 .                .          =   *   +          .          .
4.400D+01 -4.531D-03 .                .        =     *     +        .          .
4.800D+01 -1.007D-02 .                .     =        *        +     .          .
5.200D+01 -2.123D-02 .                .  =           *           +  .          .
5.600D+01 -2.239D-02 .             =  .              *              +          .
6.000D+01 -1.067D-02 .          =     .              *              +          .
6.400D+01  1.711D-04 .       =        .              *              +          .
6.800D+01  7.862D-03 .    =           .              *              +          .
7.200D+01  1.340D-02 .    =           .              *              +          .
7.600D+01  1.685D-02 .  =             .              *              +          .
8.000D+01  1.924D-02 .  =             .              *              +          .
8.400D+01  2.080D-02 .  =             .              *              +          .
8.800D+01  2.236D-02 .  =             .              *              +          .
9.200D+01  2.340D-02 .  =             .              *              +          .
9.600D+01  2.402D-02 .  =             .              *              +          .
1.000D+02  2.446D-02 .  =             .              *              +          .
1.040D+02  2.476D-02 .  =             .              *              +          .
1.080D+02  2.497D-02 .  =             .              *              +          .
1.120D+02  2.512D-02 .  =             .              *              +          .
1.160D+02  2.522D-02 .  =             .              *              +          .
1.200D+02  2.528D-02 .  =             .              *              +          .
1.240D+02  2.533D-02 .  =             .              *              +          .
1.280D+02  2.535D-02 .  =             .              *              +          .
1.320D+02  2.535D-02 .  =             .              *              +          .
1.360D+02  2.535D-02 .  =             .              *              +          .
1.400D+02  2.535D-02 .  =             .              *              +          .
```

Figure 17b Input Voltage Growth Envelope vs. Normalized Time
(To find actual time, multiply the scale by 2QR/Wo=2L)

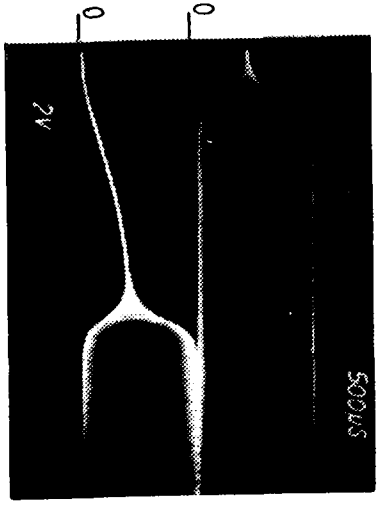



Figure 17e Output Node Growth Envelope from A. Rusniak's Paper.

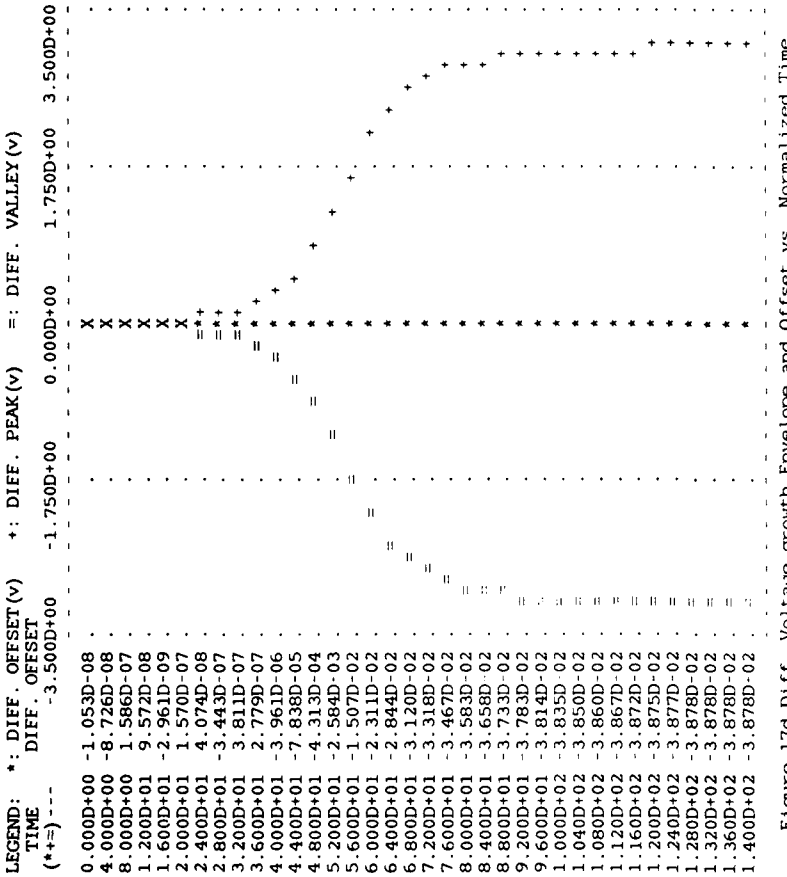



Figure 17d Diff. Voltage growth Envelope and Offset vs. Normalized Time (To find actual time, multiply the scale by 2QR/Wo=2L)

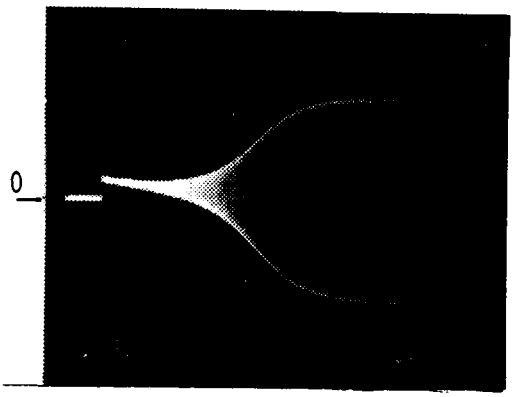



Figure 17f V(out)-V(in) Growth Envelope from A. Rusniak's Paper.

## 5.0 EXPERIMENTAL VERIFICATION OF THE THEORY

To verify the theory presented in this paper, it must be shown:

a) How to measure the LLATOR impedance the crystal 'sees' during oscillation. (To verify that s- jWo -Wo(Zo+R)/(2QR)

b) How to measure the crystal parameters without using them in an oscillator. (To verify that s- jWo -Wo(Zo+R)/(2QR))

c)That the oscillator will not start unless re(Zo)+R < 0.

d)That the oscillator frequency - jWo -Wo im(Zo)/(2QR).

e) That the crystal acts like a capacitor across the oscillator terminals at non-resonant frequencies. (To verify that s - So at all non-resonant poles.)

### 5.1 MEASUREMENT OF THE ACTIVE CIRCUIT IMPEDANCE ACROSS THE CRYSTAL

Fig. 18 is a schematic of the actual test circuit.

In the figure, a monolithic Pierce oscillator active circuit is shown with its external components. The test crystal is a 66.6667mhz Saronix SRX 2967 3'rd harmonic 'Nymph'. In order to observe the effects of the crystal on the active circuit, the impedance across nodes 1 and 2 is calculated for the oscillator minus the test crystal. Results are shown in fig. 20a and fig. 20b. The value of Z(66.667mhz) is calculated by interpolation.

To calculate the impedance between nodes 1 and 2, the $Z^{4,5}$ parameters vs frequency are measured. This can be done with an HP 8503 S parameter meter using HP software to convert the S parameters to Z parameters. In order to do its measurement, the 8503 drives the network and measures it's response. The drive signal should be 'small' (less than -50mv) to simulate oscillation startup. Formulas to convert from S parameters to Z paramaters are given in [15]. Once the Z parameters are obtained, the active network impedance across the crystal can be calculated.

If the crystal has one side connected to ground, then z11 or z22 is the desired impedance. However, if the crystal isn't connected to ground as in the Pierce oscillator, then the following calculation applies.

In fig. 19, a network is driven with a current source between nodes 1 and 2.

In the figure, I2 = -I1 producing the following Z parameter equations:

$$v1 = z11\ I1 + z12\ (-I1) \quad (55)$$

$$v2 = z21\ I1 + z22\ (-I1) \quad (56)$$

The impedance across the current source = (solving (55)-(56))

$$(v1-v2)/I1 = z11+z22-z12-z21 = Z \quad (57)$$

### 5.2 CORRELATION OF CRYSTAL IMPEDANCE WITH OSCILLATOR PERFORMANCE.

To verify item e) in sect. 5.0, plot the crystal conductance and "capacitance" (susceptance/radian frequency) vs. frequency from 10mhz to 100mhz. The thing to notice here is that the crystal is almost a perfect capacitor

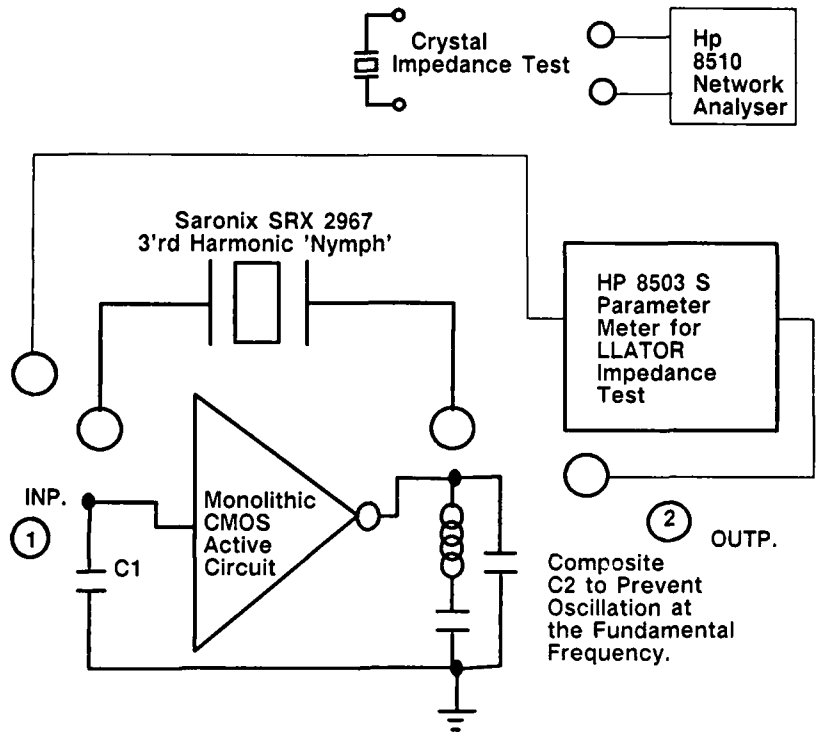



Fig. 18 Actual test circuit on which S parameters were taken

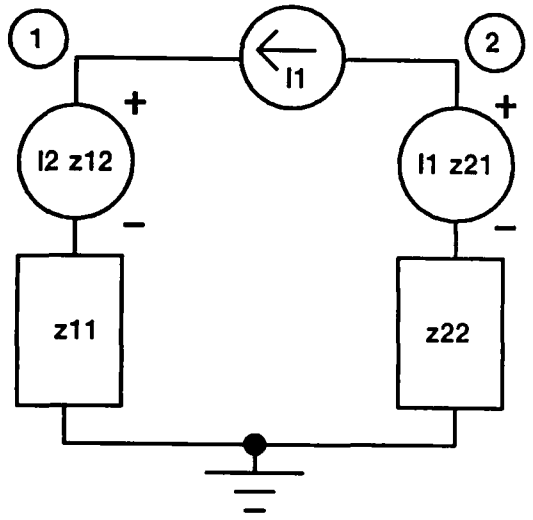



Figure 19. Generalized network for Z parameter calculations driven by a current source.

except at frequencies near resonance (-22.3mhz and -66.7mhz). Co can therefore be measured at any frequency far from resonance. The resulting value is 5.837pfd.

To measure the crystal motional resistance, while the crystal is controlling an oscillator, add a variable capacitor in series with the test crystal in fig. 18. Decrease the capacitor until the oscillator will not start, noting the capacitor value. Use a Tektronix 6022 current probe in the crystal leads to observe the oscillation current. Increase the capacitance slightly until the oscillator just starts, recording the capacitor value. Since the reactance of the added capacitor is high, THE RESULTING OSCILLATION FREQUENCY WILL BE UNSTABLE AND SHOULD NOT BE USED FOR CALCULATION. The two capacitor values are given in Figure 20f. To get the negative resistances presented to the crystal, thier impedances are added to that of the llator at 66.667mhz and the value of Co is factored in using (24) from sect. 3.0. The results are listed in Figure 20g.

If the variable capacitor is replaced by a small resistor (24ohms in our case) in series with a relatively large capacitor (-18pfd), the oscillation frequency (Wa in fig. 20c) will be stable and of low amplitude.

Replace the RC network by a resistor whose value will just allow oscillation to start(-56ohm). DO NOT USE THE VALUE OF THIS RESISTOR FOR CALCULATION BECAUSE SMALL MEASUREMENT ERRORS WILL BE EXAGERATED. The new oscillation frequency is Wb in fig. 20c.

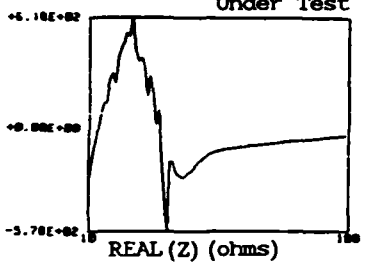
Since the oscillations are weak, Z(66.667mhz) from fig. 20a can now be used in conjunction with series impedance data to verify item d) in sect. 5.0. Using an HP 8510 high resolution network analyser, or some other means, measure the complex impedance(Za) of the crystal in series with 24ohms and 18pfd. and the current probe at Wa. Then measure the complex impedance(Zb) of the crystal in series with 56ohms and the current probe at Wb.

Since Wa - Wb - 66.667mhz, the amplifier impedance(Z(66.667mhz)) is relatively constant for the above test. To obey (1), the crystal impedance must adjust itself to compensate for the varying impedances placed in series with it while oscillating at Wa and Wb. If (1) is obeyed, then Kirchoff's voltage law for fig. 1 predicts that the values of Za(at Wa), Zb(at Wb) and -Z(66.667mhz) should be close to each other within acceptable tolerances. The results given in fig. 20c verify this.

MEASURED RESULTS

| Frequency (mhz) | Re(Z) (ohms) | | IM(Z) (ohms) |
|---|---|---|---|
| 64. | -103.6 | | -115.3 |
| 65. | -102.1 | | -111.1 |
| 66. | - 99.25 | | -106.8 |
| 66.667(interpolated) | - 98.48 | (This is Z') | -104.133 |
| 67. | - 98.09 | | -102.8 |
| 68. | - 94.97 | | - 98.5 |

Figure 20a Measured impedance Between nodes 1 and 2 of Pierce Oscillator Under Test (No Crystal)



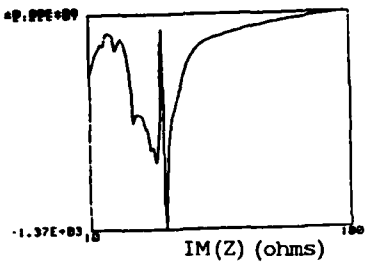
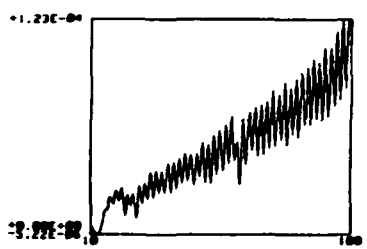




Figure 20b Plots of the Real and Imaginary Parts of Z vs Frequency(mhz) (No Crystal)

| | FREQ(mhz.) | REAL(ohms) | IMAGINARY(ohms) |
|---|---|---|---|
| Za(xtal+24ohms and 18pfd.) | 66.66752(Wa) | 101.44 | 94.5 |
| Zb(xtal+56ohm) | 66.6652(Wb) | 104.7 | 94.5 |
| -Z(66.667mhz) of amplifier circuit. | 66.667 | 98.48 | 104.133 |

fig.20c Measured values of Crystal impedance in Series with 6022 current probe 56ohm resistor, and series RC network



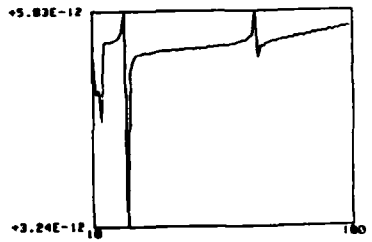




Figure 20d Plot of Crystal Admittance vs Frequency(mhz) (Not the test crystal)

| | FREQ(mhz.) | REAL(ohms) | IMAGINARY(ohms) |
|---|---|---|---|
| at Wa | 66.66752 | 75.9 | 205.6 |
| at Wb | 66.6652 | 43.1 | 94.5 |
| at Wc (zero reactance) | 66.662475 | 26.2 | 0. |

Figure 20e Measured Crystal Impedances Close to Resonance

Co, the parallel Stray Capacitance = 5.837pfd.
Series capacitance to just start oscillation = 11.04pfd
Series Capacitance to just stop oscillation = 9.757pfd

Figure 20f Important Capacitance Values Measured at 8mhz

| | FREQ(mhz.) | REAL(ohms) | IMAGINARY(ohms) |
|---|---|---|---|
| Wa | 66.66752 | 33.1 | 141.17 |
| Wb | 66.6652 | 29.61 | 68.2 |
| Wc | 66.662475 | 26.1 | 1.66 |
| Negative resistance presented to crystal | (osc. just starts) | -30.4 | -------- |
| | (osc. just stops) | -28.28 | -------- |

Figure 20g Resonant Portion of Crystal Impedance Calculated from (58)

| DF(mhz) | DZr(ohms) | Calculated L |
|---|---|---|
| 66.66752-66.662475 | 141.17-1.66 | 2.2mh |
| 66.6652-66.662475 | 68.2-1.66 | 1.943mh |
| 66.66752-66.6652 | 141.7-68.2 | 2.5211mh |

Figure 20h Calculated values of L using (60)

$|D/L| = 2.244e\text{-}4 << 1.$
$|(R+Zo)/(2WoL)| = .62e\text{-}4 << 1.$

Figure 20i Max. Calculated values of |D/L| and |(R+Zo)/WoL| to Verify condition (6) (L SUBSTITUTED FOR QR/Wo)

## 5.3 MEASUREMENT OF CRYSTAL PARAMETERS WHILE THE CRYSTAL IS NOT CONTROLLING OSCILLATION

To calculate the crystal resonant circuit parameters without having them control oscillation, the complex impedance of the physical crystal (including Co) must be measured at two frequencies close to series resonance (Wa and Wb from section 5.2). A third measurement is taken at the frequency (Wc) where the crystal reactance is zero. Results are given in figure 20e above.

Resonant circuit impedances at Wa, Wb, Wc and startup can now be calculated using Co, Z', figures 20e, 20f and (58) below. (58) follows from (25) of sect. 3.0. The crystal resonant circuit impedance is analogous to Z' and Wo Co is analogous to Yb.

$$R + jZr = \frac{Re(Zc) + j[Im(Zc) + Wo\ Co(Im(Zc)^2 + Re(Zc)^2)]}{(1 + Wo\ Co\ Im(Zc))^2 + (Wo\ Co\ Re(Zc))^2}, \quad (58)$$

where R = the crystal internal resistance (should be a constant),

Zr = the reactance of the series resonant part of the crystal at the test frequency,

Re(Zc) = the real part of the crystal complex impedance measured at the test frequency,

Im(Zc) = the imaginary part of the crystal complex impedance measured at the test frequency,

Wo = the crystal series resonant frequency.

Since R is a constant, all calculated values of R should agree reasonably well. (See figure 20g) They should also be within the limits set by the oscillator start up tests (sect. 5.2). Due to experimental error, this is not the case. However, the values are close enough so that item c) in sect 5.0 is believable.

To find L, the effective inductance of the crystal, first differentiate Zr with respect to Wo.

$$d(Zr/d(Wo) = d(Wo\ L - 1/Wo\ C)/d(Wo) = L + 1/(CWo^2) - 2L \quad (59)$$

Since the experiment is run at three slightly different frequencies (Wa, Wb, Wc), three different values of Zx (shown in fig. 20g) are obtained. Use them to calculate three values of L. If frequency variation is proportional to im(Z), the obtained values agree reasonably well with each other as shown in figure 20h.

$$L = .5dZr/DWo - .5(Za-Zb)/(Wa-Wb) \quad (60)$$

To verify that condition (6) applies, evaluate each term using the values of L, Co, and the data of fig. 20a. To do this, use (24) of sect 3.0 to add Co in parallel with the active circuit impedance data at 66mhz and 67mhz to evaluate:

$$D=[Z(67MHZ)-Z(66MHZ)]/(2.x3.1415927e6) \quad (61)$$

The results, given in figure 20i show that condition (6) is obeyed very well.

## 6.0 CONCLUSIONS

This paper presents a simple and accurate theory for the design of high Q tuned circuit oscillators. The crystal is separated from the active oscillator and all high Q approximations are handled before the design begins. Preveously published concepts (OSCI-LLATOR, Negative Resistance))[1-4] are rigorously used as a basis for the final formulas. Technology independent equations for oscillation build up time constant and accurate frequency determination are given along with validity checks on the calculations. The final result is that high Q oscillator design is reduced from a two port to a 1 port network problem. All remaining mathematical complexity comes from the design of the active impedance and is not a specific part of crystal oscillator theory. The derived equations were experimentally verified to 'reasonable' measurement accuracy.

Pierce and Miller Oscillator design methods based on the theory are demonstrated and 'Spice' runs are made to demonstrate the theory's application.

## APPENDIX 1 EQUATIONS FOR OSCILLATION GROWTH

(11) in sect 2.2 says that at any simulated time,

$$-[R+re(Zo)]Wo/(2QR) - 1/(\text{THE OSCILLATION GROWTH TIME CONSTANT}). \quad (11)$$

This means that for any small time window,

$$I(t+dt)/I(t) - \exp(-dt[R+re(Zo)]Wo/(2QR)), \quad (a1)$$

where I(t) is the crystal current amplitude as a function of time, t is time and dt is a relatively small time period.

Taking logarithms gives

$$Ln(I(t+dt)/I(t)) - dI/I(t) - -dt[R+re(Zo)]Wo/(2QR), \text{ where} \quad (a2)$$

dI is the small change in crystal current over the period dt, L=QR/Wo, and R+re(Zo) is a function of I(t). Simulation data for the typical CMOS Pierce llator is given in figure 16. Solving (a2) for dt/dI gives

$$- 2L/(I(t)(R+re(Zo)) - dt/dI. \quad (a3)$$

If the points in figure 16a are connected by straight lines, then

$$R+re(Zo) - R(n) + Dn\ (I(t)-I(n)) \quad (a4)$$

between the n'th and n+1'th SPICE runs. (Called the n'th region. n is the run # tabulated in figure 16.)

$$Dn = \frac{R(n+1)-R(n)}{I(n+1)-I(n)}. \quad (a5)$$

R(n) = R +re(Zo) at I(n), the n'th current level. In figure 16a, $I(n) = .01\ 10^{(n-1)}$ and I(0)=0. Substitution of (a4) into (a3) produces

$$-2L/(I(t)(R(n) + Dn(I(t)-I(n)) = dt/dI \quad (a6)$$

This equation is valid within the n'th region. The solution for the time spent in the n'th region is

$$t - to = 2L\left(Ln\left(\frac{I(t)\ \{R(n)+Dn(Io-I(n))\}}{Io\ \{R(n)+Dn(I(t)-I(n))\}}\right)\right)/(Dn\ I(n)-R(n)). \quad (a7)$$

to is the time that the oscillations first enter the n'th region, and Io is the crystal current at that time. During growth, the oscillations enter the n'th region at Io = I(n). I is assumed to be Io at t=0.

Warm up is the sum of the times spent in all regions.

BIBLIOGRAPHY


[1]Eric A. Vittoz, "Quartz Oscillators for Watches" 1979 Proceedings of the International Congress on Chronomometry.

[2]Robert J. Matthys, "Crystal Oscillator Circuits" Wiley-Interscience 1983 Section 6.1.

[3]Marvin E. Frerking, "Spurious Oscillation in Crystal Oscillators" Proc. 20'th Annual Symposium on Frequency Control, April,1966, pp.501-516.

[4]Bejamin Parzen with Arthur Ballato "Design of Crystal and Other Harmonic Oscillators" Wiley-Interscience 1983 pp.13-15, 402-410.

[5]Marvin E.Frerking "Crystal Oscillator design and Temperature Compensation" Van Nostrand Reinhold 1978.

[6]R. J. Holbeche, and G. Allen "Influence of Series Reactance on Quartz Crystal Resonators" IEE Proceedings Vol. 130, Pt. G, #4, August 1983.

[7]M. A. Unkrich and R. G. Meyer "Conditions for Startup in Crystal Oscillators", IEEE Journal of Solid State Circuits, Vol. SC-17, #1, Feb., 1982.

[8]R. G. Meyer and David Soo "MOS Crystal Oscillator Design" IEEE Journal of Solid State Circuits Vol. SC-15, #2, April,1980.

[9]T. W. Pickerell "Single Mosfet Crystal Oscillator Design for Integrated Circuits", Monochip Application Note APN-29. Interdesign, 1983.

[10]L. Odess "Analysis of the Frequency and Sensitivity of Crystal Controlled Oscillators", The Radio and Electronics Engineer Vol 46, #3 pp109-116 March, 1976.

[11]J. L. Creighton, H. B. Law and R. J. Turner "Crystal Oscillators and their Application to Radio Transmitter Control", Journal of Instn. Elect Engineers, 94, Pt 3A, No. 12, 1947.

[12]P. J. Baxandall "Transistor Crystal Oscillators and the Design of a 1-MC/s Oscillator Circuit Capable of Good Frequency Stability", The Radio and Electronics Engineer, April, 1965.

[13]J. T. Santos and R. G. Meyer "A One Pin Oscillator for VLSI Circuits" IEEE Journal of Solid State Circuits" Vol SC-19, #2, April,1984.

[14]A. Vladimeriscu, A.R. Newton, D. O. Pederson "Spice Version 2F User Guide", Dept of Electrical Engineering and Computer Sciences, University of California, Berkeley.

[15]D. Anderson "S Parameter Techniques for Faster, More Accurate Network Design" Hewlett Packard Journal Vol. 18, No. 6, Feb. 1967. Also available as H.P. Application Note 95-1.

[16]T. Adachi, M. Hirose, Y. Tsuzuki "Computer Analysis of Colpitts Crystal Oscillator" Proceedings of the 39'th Annual Frequency Control Symposium-1985, pp.176-182.

[17]M. Toki, Y. Tzusuki "Measuring Method of Equivalent Series Capacitance and Negative Resistance of Crystal Oscillator Circuits", Electronic Letters Vol. 19, No. 23, pp 959-960, Nov. 1983.

[18]E. Hafner "Theory of Oscillator Design" Proceedings of the 11'th Annual Frequency Control Control Conference, May, 1963, pp508-536.

[19]A. Rusnyak "Start up Time of CMOS Oscillators" IEEE Transactions on Circuits and Systems Vol. Cas-34, No. 3, March 1987.

[20]G. Korn, T. Korn "Mathematical Handbook for Scientists and Engineers" McGraw Hill Book Company

[21]F. B. Hildebrand "Advanced Calculus for Applications", Prentice Hall, 1963

[22]S. Nonaka "Abnormal Crystal Oscillator Frequency Change with Load and its Elimination" Proceedings of 28'th Frequency Control Symposium - 1974 pp203-210

42nd Annual Frequency Control Symposium - 1988

# A NEW DIGITALLY TEMPERATURE COMPENSATED CRYSTAL OSCILLATOR FOR A MOBILE TELEPHONE SYSTEM

T. MIYAYAMA, Y. IKEDA and S. OKANO
Toyo Communication Equipment Co., Ltd.
Samukawamachi, Kozagun, Kanagawa pref., JAPAN

## Summary

This paper relates to a digitally temperature compensated crystal oscillator suitable for mobile telephone system applications. The newly developed DTCXO has the following characteristic features:

1. High frequency stability (+/- 0.5ppm max. / -30 to +80°C)
2. Low power consumption (3mA max. / +5V operation )
3. Compactness (small size of only 4cm³)

To develop the DTCXO, it was necessary to develop new devices. Thus, for example, the requirements for greater compactness (size reduction) and lower power consumption have led the authors to use ASICs (application specific integrated circuits) with C-MOS technology. Another example is the X-cut crystal unit. It has been newly developed to function as the temperature sensor, being accommodated in an extremely small receptacle. For software development, the authors developed a temperature compensation method which enables us to calculate successfully temperature compensation data at the non-measurement points by the interpolation formula.

This paper presents the satisfactory results obtained with the DTCXO in application experiments to demonstrate its suitability for mobile telephone system.

## Introduction

Mobile telephone systems are currently being operated worldwide in the 800MHz and/or 900MHz frequency band, and as the numbers of subscribers increase , the access channels become more saturated and radio frequency channel spacings are narrowed down from 25kHz to 12.5kHz. This gives rise to frequency stability problems, with much tighter specifications being demanded to ensure greater stability. The mobile telephone systems with the narrower channel spacing of only 12.5kHz is expected to require a very stringent frequency stability of a +/- 0.5ppm in the specified temperature range[1,2]. Although a crystal oscillator which meets this stringent frequency stability requirement (+/- 0.5ppm), can be achieved with the conventional temperature compensation method using a thermistor and resistor network, the problem is that in mass-production terms this would limit the temperature range which is over -10 to +70°C for the severe frequency stability of +/- 0.5ppm.

The conventional temperature compensation techniques cause enormous problems, when extending this temperature range to -30 to +80°C for mass-production of oscillators maintaining a frequency stability within +/- 0.5ppm. We have therefore resorted to digital technology to achieve this high frequency stability. This has led to our development of a digitally temperature compensated crystal oscillator.

Historically, the development of the DTCXO dates back to a fairly remote past when research started, mainly in the United States, over 15 years ago.

In the course of development, DTCXO has been proposed with various temperature compensation techniques. As shown Figure 1 a and b, broadly these temperature compensation methods can be classified into two categories. Figure 1a shows a PROM used as a temperature compensation data mapping. The temperature compensation data pre-written into the PROM are read out on the bassis of the temperature data that have been converted to digital codes through an analog-to-digital (A/D) converter. These temperature compensation data are converted to the corresponding analog voltage signals through a D/A converter , and then applied to the VCXO frequnecy control terminal to ensure smothly that the frequency error of the VCXO is flattened out to zero. The above system operates on the basis of an extremely simple sequence and has been extensively used by a large number of researchers in the field[3,4,5,7].

Figure 1b, however, shows a system in which the PROM of Figure 1a has been replaced by a microprocessor. In this case, the microprocessor is not used simply as a temperature data map, but also functions as an intelligent control device. Thus, the microprocessor has a calculating capability that allows itself to reduce the scale of the temperature map by interpolation and also permits to use timing signals for controlling the DTCXO system[6]. Of the two systems given above, we have chosen the PROM version shown in Figure 1a, as it offers a greater potential for compact design, low power consumption, and economical benefits.

## Basic Construction of the DTCXO

A simplified block schematic of the DTCXO developed by the authors is shown in Figure 2. In this figure, except for part of the VCXO and the PROM, we have achieved a greater degree of integration with a C-MOS gate array, one of the available ASIC methods. This has made a significant contribution to improving the compact design, power consumption and economic performance of the DTCXO. The temperature compensation data are stored in the PROM which takes the form of a small outline (s.o.) packaged C-MOS one-time PROM that makes surface mounting a practical possibility in terms of compactness and low power consumption. The temperature range in which operation is assured was previously checke at the time of selecting the PROM. This was essential, because the ambient conditions applicable to mobile telephone systems utilizing the DTCXO, cover a wide temperature range from -30 to +80°C.

Let us now outline in more practical terms the design principles for the DTCXO.

## Design Concept of the DTCXO

The design of the circuit blocks composing the DTCXO was a painstaking development process based on the following essential ideas or requirements:

CH2588-2/88/0000-327 $1.00 C 1988 IEEE

Figure 1a. Basic DTCXO by Using PROM

Figure 1b. Basic DTCXO by Using Microprocessor

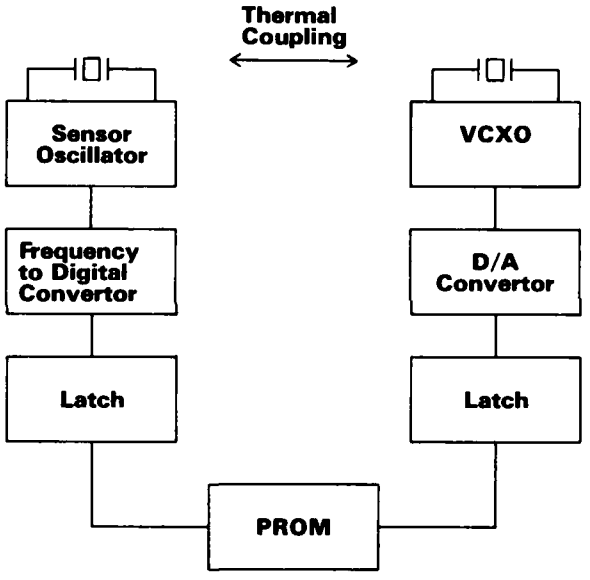



Figure 2. Simplified Block Diagram of the DTCXO

1. The system must ensure an excellent frequency stability.
2. The system must offer the benefits of a compact design (small size and light weight).
3. The system must provide for low power consumption.
4. The system must have a high reliability.

## Design of the VCXO

A crystal oscillator with a C-MOS inverter is generally claimed to have a low power consumption. The authors have tried to go further by trying to include the VCXO crystal oscillator, composed of a C-MOS inverter, into a gate array circuit. The problem, however, is that the power consumption of a C-MOS inverter varies proportionally to the frequency, as can be found from the equation below[9].

$$P_D = C_L * V_{DD}^2 * F \qquad (1)$$

Where $P_D$ = Power consumption
$C_L$ = Load capacity including the stray capacity
$V_{DD}$ = Supplied voltage
$F$ = Operating frequency

It therefore follows that within the low-frequency range, the power consumption $P_D$ can be made to have small values. In the high or higher frequency range above 10MHz, however, it assumes rather high values.

The authors have therefore constructed a VCXO circuit with bipolar transistors, without using a C-MOS inverter as the VCXO crystal oscillator.

These bipolar transistors were connected in a cascade arrangement to successfuly reduced the power consumption of the VCXO to 7mW or less. One of these two transistors was used in a high-loaded Q Colpitts type crystal oscillator circuit while the other transistor served as a base-mounted type amplifier. The literature refers to DTCXOs using a switched capacitor array with C-MOS technology as a frequency variable device[10].
The authors, however, adopted a proven varactor diode in the conventional TCXO as a variable reactance device.

## Memory Size of the PROM

The following considerations were made to allow for the memory size of the PROM required to meet the +/- 0.5ppm frequency stability tolerance within the wide temperature range from -40 to +90°C (a range which has a 10°C margin at either end over and above the actual temperature range).

The frequency - temperature characteristics of an AT-cut quartz crystal unit with a zero temperature coefficient over a relatively extensive range can be expressed by the following polynomial equation.

$$dF/F = A(T-To) + B(T-To)^2 + C(T-To)^3 \qquad (2)$$

Where To is the reference temperature. The factors A, B, and C, with To = 25°C, are given as follows, on the basis of the experimental values determined by M. Aruga:

$$A = (3.314\ E\text{-}07) - (5.035\ E\text{-}06) * THETA \qquad (3)$$

$$B = (-5.475\ E\text{-}10) - (4.871\ E\text{-}09) * THETA \qquad (4)$$

$$C = (1.175\ E\text{-}10) - (1.272\ E\text{-}11) * THETA \qquad (5)$$

In Figure 3 shows the frequency - temperature curve of a AT-cut crystal unit obtained by substituting the above equation (3), (4), and (5) in equation (2) and taking the offset angle THETA from the cut angle of the quartz crystal with a zero temperature coefficient as a parameter.

With respect to the curves presented in Figure 3, Figure 4 gives the curves obtained from the first-order differential against the temperature, to show the amount of variation in frequency associated with one degree centigrade. Within the temperature range (-40 to +90°C), the curves are those of ⑧, with the maximum frequency variation for a change in temperature by 1°C being +1.05ppm/°C at -40°C.

To achieve a temperature stability of +/- 0.5ppm, it is essential to allow for the various error factors so as to achieve a maximum frequency variation of 0.6ppm (+/-0.3ppm) after temperature compensation. In other words, if instead of sampling every 1°C, the sampling rate for curve ⑧ in Figure 3 is changed to every 0.6°C, it is possible to achieve the desired frequency stability.

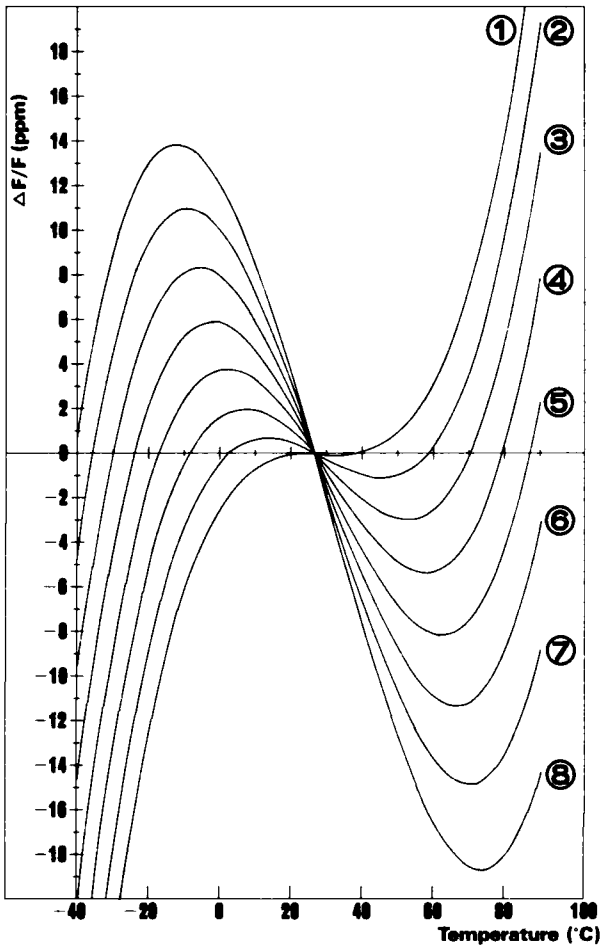



Figure 3. Frequency Temperature Curve of the AT-cut Crystal Units

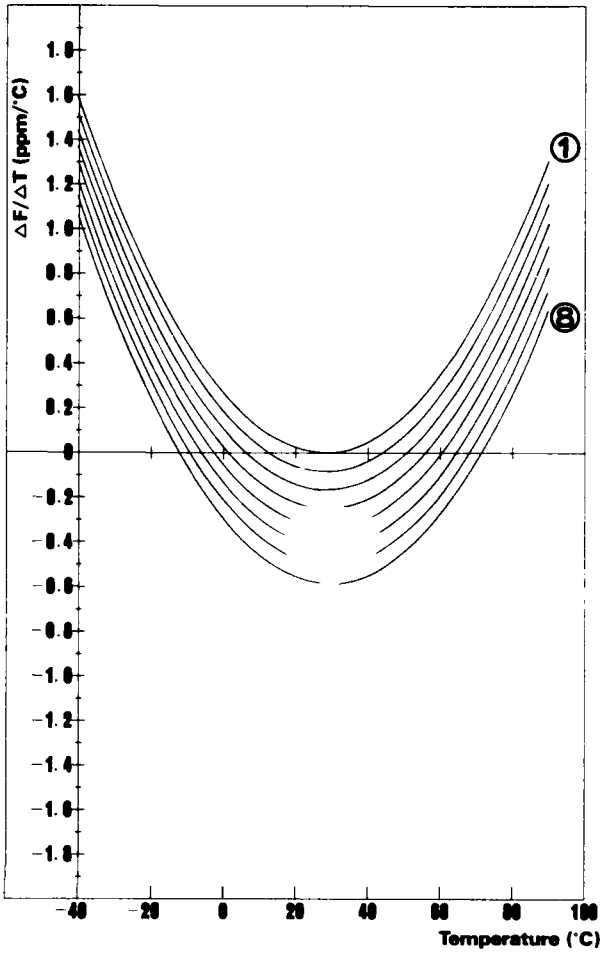



Figure 4. First-order Differential for the Frequency Temperature Curve of the AT-cut Crystal Units

This means that if temperature compensation in the 130°c range from -40 to +90°C is carried out in 217 (=130/0.6) steps, it is possible to attain a frequency stability of +/- 0.5ppm or better. All that is required, therefore, is that should be possible to have a memory table with an 8 bit ( 256 steps ) structure in the PROM memory.

## Temperature Detection

Figure 5 shows the DTCXO's temperature detection circuit design. For temperature sensing, the C-MOS inverter oscillates a 40kHz crystal sensor whose frequency sensitively responds to the ambient temperature, and this temperature-sensitive frequency is divided into $1/2^M$ by a frequency divider, thereby acting as a temperature-variable gate for frequency counting.

The standard pulse signals to be counted, however, provide the input for the gate signal generator after the VCXO output frequency has been divided into $1/2^N$ by a frequency divider. The standard pulse signals that have been fed into the temperature-variable gate are counted by a frequency counter, with the result of this count being output in the form of an 8 bit digital code.

This digital code is entered as an address input into the PROM, and the temperature compensation data on the previously entered PROM map are read out.

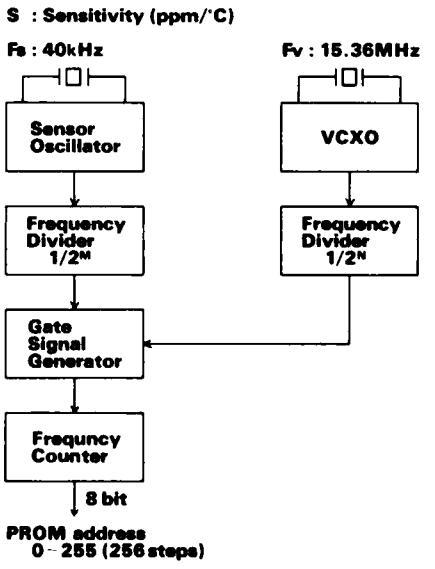



Figure 5. Temperature Detection of the DTCXO

The next equation gives the relationship between the individual configuration elements shown in Figure 5. In this equation let us assume that the amount of change of the pulse counted by the frequency counter within the compensation temperature range, is delta NP. This delta NP value will correspond to the extent of change of the address value entered into PROM.

$$\text{delta NP} = 2^{(M-N-1)} * Tw * Fr * S/Fs \qquad (6)$$

where M : Dividing number for the sensor frequency
N : Dividing number for the VCXO frequency
Tw: Compensation temperature range ( °C )
Fv: VCXO oscillation frequency ( Hz )
Fs: Sensor oscillation frequency ( Hz )
S : Sensor sensitivity ( ppm/°C )

The sensitivity required for the crystal temperature sensor can be obtained from the next equation by rearranging equation (6).

$$S = (\text{delta NP} * Fs)/2^{(M-N-1)} / Tw / Fr \qquad (7)$$

Where equations (6) and (7) are valid provided that Sv << S and that the differentiation coefficient for the VCXO's frequency - temperature curve is Sv (ppm/°C).

Let us now actually determine the sensitivity S of the temperature sensor by using equation (7) in much the same manner as (8).

The VCXO's frequency is here taken as having a value of 15.36MHz, a frequency often used as the reference for the frequency synthesizer of the US cellular mobile telephone systems. The other factors are as follows.
M = 11, N = 2, Tw = 130°C (-40 to +90°C), Fv = 15.36E6 Hz, Fs = 40E3 Hz, and delta NP = 256 (maximum step number due to the 8 bit address):

$$S = \frac{256 \times 40 \times 10^{3}}{2^{(11-2-1)} \times 130 \times 15.36 \times 10^{6}} \qquad (8)$$

$$= 20 \times 10^{-6}$$

This means that a temperature sensor with a temperature sensitivity of 20ppm/°C is required as the temperature sensing device. It is for this reason that the small crystal temperature sensor was developed. This crystal unit is of the X-cut tuning fork type, located in a cylindrically shaped metal housing (RV-26) of 2mm diameter and 6mm length. The actual frequency - temperature characteristic is shown in Figure 6. The X-cut crystal unit has an essentially parabolic frequency - temperature curve. Therefore, to achieve a sensitivity of 20ppm/°C over a wide temperature range from -40°C to +90°C, the temperature at the apex of the parabola has been set to an extremely low temperature.

The advantages of using a 40kHz crystal temperature sensor can be summarized as follows:

1. Compact - small, lightweight
2. Superior mass-production capability and high reliability
3. Frequency adjustment free
4. Superior aging characteristics
5. Since the frequency is low at only 40kHz, the current consumed by the oscillator circuit of the C-MOS is extremely low, amounting to only 100uA.

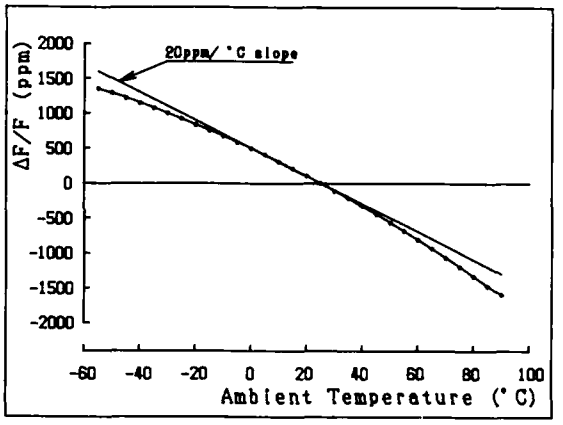



Figure 6. Frequency - Temperature Characteristic of the X-cut Quartz Crystal Sensor

## Compensation Procedure

Figure 7 shows a fully automated temperature compensation system. For simplicity, the system is shown with only one DTCXO compensated. The data bus lines for the DTCXO output signals and temperature compensation can be switched under scanner control so that this temperature compensation system is suitable for mass-production and capable of expansion.

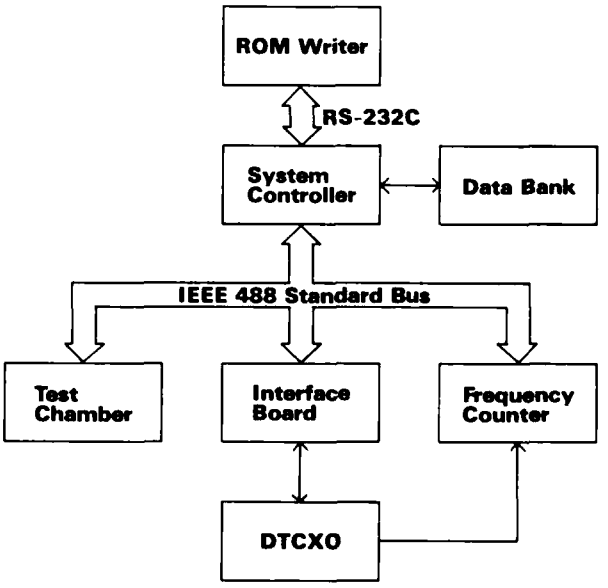



Figure 7. Fully Automated Temperature Compensation System of the DTCXO

Figure 8 shows the temperature compensation procedure. First, the measuring conditions for the temperature compensation system are entered into the controller.

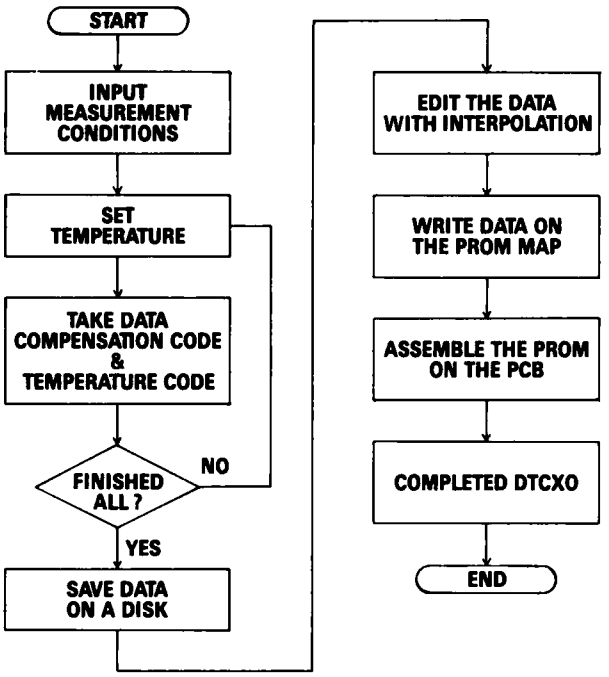



Figure 8. Compensation Procedure of the DTCXO

These measurement conditions include the frequency to be set for the DTCXO, the temperature range in which compensation should take place, and the temperature setting interval for the temperature testing chamber. Moreover, the digitalized 8 bit temperature code and the 8 bit compensation code for compensating the frequency error of the DTCXO are measured after the temperature in the test chamber has sufficiently stabilized at the corresponding temperature setting points. The measurement data are then stored in the controller memory. After the measurements have been completed at all of the set temperature points, the data are stored on a disk. To enhance the accuracy of compensation, the measurements may also be taken at intervals of 1°C each, throughout the compensation temperature range. This does have the drawback, however, that the measurements take a very long time, an aspect which would be a serious impediment to the DTCXO's mass-production. The authors have tried to overcom this problem by using an interpolation method to determine the compensation codes for all temperature at which no measurements are performed. This interpolation approach has made it possible to reduce the time required for completing the measurements since it allows the number of measurement points to be significantly reduced and the temperature measuring intervals to be correspondingly spaced out. The data edited by polynomial interpolation are finally entered into the DTCXO's one-time PROM by means of a ROM writer to complete the compensation procedure.

The following section sums up the test results obtained with the DTCXO that has been temperature compensated by the above procedure.

## Experimental Results

Figure 9 shows the frequency - temperature curve prior to temperature compensation of the VCXO which is to be digitally temperature-compensated. From the curve it is clear that the sensitivity of VCXO control has been set to 15ppm/V. Let us assume that the 8 bit digital-to-analog (D/A) converter has a full scale range (FSR) of 2.56V. The control voltage corresponding to one step is therefore 10mV. By consequence, the sensitivity of control for the VCXO per digital code step is therefore 0.15ppm/step.

This value is adequate as the sensitivity for controlling a +/-0.5ppm grade DTCXO (at a width of 1ppm).

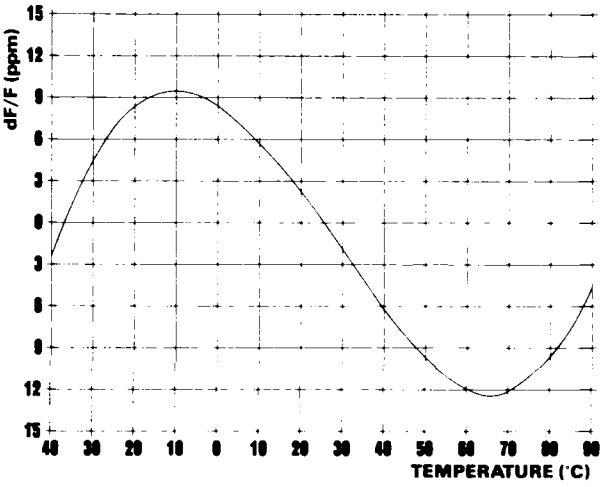



Figure 9. Frequency - Temperature Curve before Compensation

Figure 10 and 11 show the frequency - temperature curve after temperature compensation. The data are acquired for temperature compensation at intervals of 10°C each. Thus, the curve shown in Figure 10 is arrived at by linear approximation of the temperature compensation codes at each 10°C intervals while the curve given in Figure 11 is based on the results obtained by polynomial interpolation.

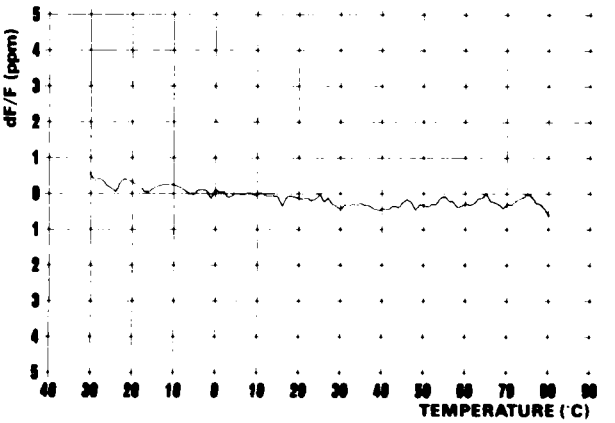



Figure 10. Result after Compensation with Linear Approximation

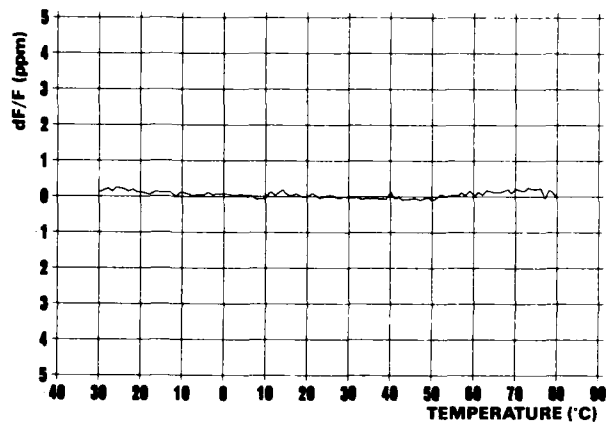



Figure 11. Result after Compensation with Polynomial Interpolation

The results are very favorable, remaining within +/- 0.3ppm in the wide temperature range from -30°C to +80°C.

Figure 12 shows the current consumption of the newly developed DTCXO. The total current consumption is 2.9mA, a value similar to that of the conventional analog TCXO used in mobile telephone systems.

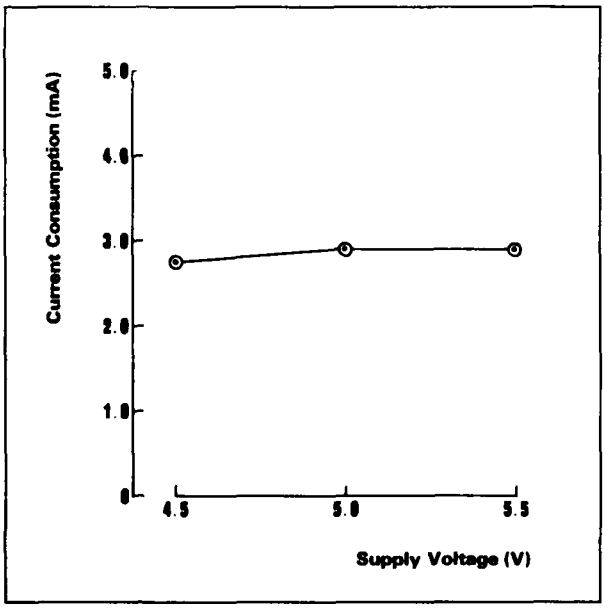



Figure 12. Current Consumption of Newly Developed DTCXO

## Application for Mobile Telephone Systems

The DTCXO descrived above may be used as a standard oscillator for the frequency synthesizers of the mobile telephone system shown in Figure 13. Through its phase-locked loop configuration, this standard oscillator stabilizes the voltage controlled oscillator (VCO) with a poor frequency stability. Figure 14 shows the detailed specifications for the DTCXO designed for mobile telephone systems called the TCO-103 series.

The outline and internal structures of the DTCXO are shown in Figure 15. This DTCXO is composed with two printed circuit boards. The newly developed C-MOS gate array is mounted on both side of the PCB shown as the right side of the picture.

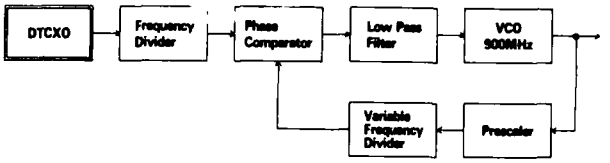



Figure 13. Simplified Diagram of A Frequency Synthesizer

1. Output Frequency : 12.8MHz or 15.36MHz
2. Supply Voltage : +5V ± 0.5V
3. Current Consumption : 3mA
4. Output Level : 1Vp-p, Clipped Sinewave in 20K // 5pF Load
5. Frequency Stability
   - vs Temperature : ± 0.5ppm -30 ~ +80°C
   - vs Supply Voltage : ± 0.2ppm +5V ± 0.5V
   - Aging : ± 0.8ppm year
6. Frequency Adjustment : ± 3ppm min. by internal trim. ± 3ppm min. by external voltage +2.5V ± 1V
7. Dimensions and Pin Configurations

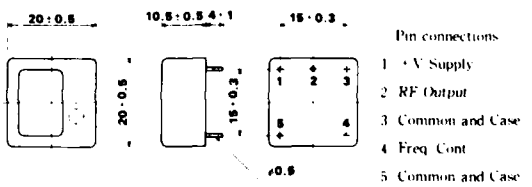



Figure 14. Specifications of the DTCXO (TCO-103 Series) for A Mobile Telephone System

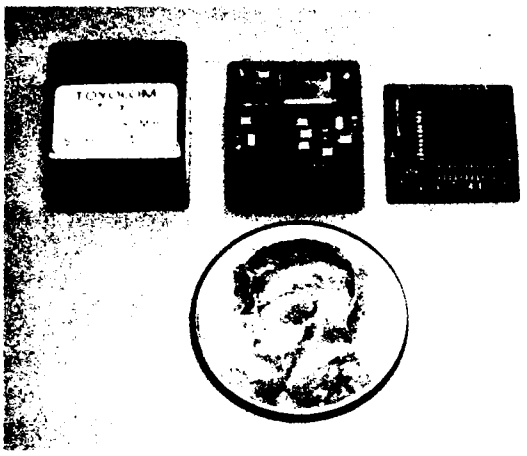


Figure 15. Outline and Internal Structure of the Newly Developed DTCXO

## Conclusions

This paper gives a detailed description of the DTCXO designed for mobile telephone system applications for which it is essential to have a narrower channel spacing.

The authors have used a polynomial interpolation technique to achieve a DTCXO with a high frequency - temperature stability within +/- 0.3ppm.

The compact (4cm') low-power consumption (15mW) DTCXO has been developed by using a flat-packaged C-MOS gate array in the circuit of the digital unit.

In the near future, ASIC may be available in a mixed form including bipolar and C-MOS. By that time, it may be possible to achieve a complete one-chip design for the DTCXO and to accomodate the VCXO also on the same chip. This may bring a further reduction to 1/2 the present size within the bounds of possibility.

## Acknowledgment

The authors wish to express their gratitude for the guidance and encouragement received from Dr. Y. Tsuzuki of Yokohama National University. They would like to thank S. Matsumura for assistance of designing the C-MOS gate array.

## References

1. Nordic Mobile Telephone Group, Automatic Cellular Mobile Telephone System NMT-900 , Technical specification for the mobile station., NMT DOC.900-3, p.18, 1985

2. Uno, T. "Electromechanical Devices for Mobile Communication Systems," Proc. 15th EM Symposium, 1986, pp.45-50.

3. Buroker, G. E., and Frerking, M. E. "A DIGITALLY COMPENSATED TCXO," Proc. 27th Annual Frequency Control Symposium, 1973, pp.191-198.

4. Mroch, A. B., and Hykes, G. R. "A MINIATURE HIGH STABILITY TCXO USING DIGITAL COMPENSATION," Proc. 30th Annual Frequency Control Symposium, 1976, pp.292-300.

5. Scott, P. J. "DESIGN CONSIDERATIONS FOR A DIGITALLY TEMPERATURE COMPENSATED CRYSTAL OSCILLATOR," Proc. 31st Annual Frequency Control Symposium, 1977, pp.407-411.

6. Frerking, M. E. "THE APPLICATION OF MICRO-PROCESSORS TO COMMUNICATIONS EQUIPMENT DESIGN," Proc. 33rd Annual Frequency Control Symposium, 1979, pp.431-435.

7. Pollard, G. B. "A DIGITALLY COMPENSATED HYBRID CRYSTAL OSCILLATOR," Proc. 36th Annual Frequency Control Symposium, 1982, pp.474-479.

8. Onoe, M., Yamagishi, I., and Nariai, H. "TEMPERATURE COMPENSATION OF CRYSTAL OSCILLATOR BY MICROPROCESSOR," Proc. 32nd Annual Frequency Control Symposium, 1978, pp.398-402.

9. Ong, D. G. "MODERN MOS TECHNOLOGY," McGraw-Hill, p.267, 1984

10. Uno, T., and Shimada, Y. "A NEW DIGITAL TCXO CIRCUIT USING A CAPACITOR - SWITCH ARRAY," Proc. 37th Annual Frequency Control Symposium, 1983, pp.434-441.

42nd ANNUAL FREQUENCY CONTROL SYMPOSIUM - 1988

# AN ACCELERATION COMPENSATED PRECISION QUARTZ OSCILLATOR

B.T. Milliren, D.W. Martin and D.A. Emmons
Frequency and Time Systems, Inc.
Beverly, MA 01915

## ABSTRACT

A dual-resonator compensation method for reducing acceleration sensitivity of precision quartz oscillators has been developed. Readily available 10 MHz SC-cut resonators are counteraligned to cancel the individual $\Gamma$ vectors. Phase noise under vibration is reduced to the level expected for the measured residual sensitivity. Performance characteristics of the newly developed FTS Model 1120 are described.

## I. INTRODUCTION

The need for high-precision oscillators having very low sensitivity to acceleration has grown rapidly as system demands have become more stringent. Requirements for improved oscillators are seen in position location and GPS receiver applications, coherent radar systems, deep space communications, and phase locked loop systems in which vibration induced phase noise adversely affects signal acquisition.[1]

A new acceleration compensated precision quartz crystal oscillator has been developed. The FTS Model 1120 employs two 10 MHz 3rd overtone SC-cut resonators which are counteraligned to provide compensation through cancellation of the individual acceleration sensitivity vectors. It is shown that this dual-resonator compensation technique reduces the g-sensitivity of an oscillator, utilizing readily available resonators which can be matched in pairs.

Compensation can be achieved over the vibration frequency range of DC to 2 kHz. The results for a prototype oscillator show that the achieved sensitivity along the principal axis is indeed equivalent to the vector difference of the individual resonator sensitivities.

Success of this technique depends on the fact that the g-sensitivity of resonators generally behaves as a vector quantity [2,3,4]. Thus, careful counter-alignment of two equal magnitude vectors should lead to a null sensitivity in all directions; several authors have discussed this method [5,6,7].

The concept is reminiscent of the pioneering work of John Harrison; his counterbalanced, counter-rotating pendulum assemblies resulted in the first high-precision marine chronometers 250 years ago.

The dual resonator technique is compared with other compensation methods involving feedback from an accelerometer. Limitations common to all methods are considered. Phase shifts which destroy the vector cancellation at certain frequencies are seen to require vibration isolation in order to attenuate high frequency vibration input.

Very careful alignment of resonators and sensitive measurement techniques are required. Alignment is checked under dynamic vibration conditions and these measurements are more complex than for alignment of a single resonator along the known axis of an accelerometer. In addition, it is important to perform the vector alignment under the condition of correct resonator operating temperature.

Residual effects common to any low g-sensitive oscillator are discussed. These include circuit board or stray capacitance vibration sensitivities. The FTS 1120 is implemented using hybrid and surface mount techniques to reduce to a minimum any phase modulation due to modulated circuit capacitances.

Performance data for the FTS 1120 oscillator is presented, showing phase noise for both quiescent and vibrating operation, time domain stability, and temperature sensitivity.

## II. DISCUSSION

The intrinsic acceleration sensitivity of resonators has several consequences: vibration induced phase noise and the associated degradation in system performance; dependence of oscillator frequency on orientation in the earth's gravitational field; and degradation of time-domain stability in noisy or even normal micro-g environments. A useful summary of these effects is given in reference 1.

The impact on phase noise performance in the case of sinusoidal vibration is:

$$L(f) = 20 \log [(\Gamma \cdot a) f_o / 2f]$$

where $\Gamma$ is the fractional frequency sensitivity of the resonator, a is the peak g input, and $\Gamma \cdot a$ is the vector product of the two. The oscillator frequency is $f_o$ and the vibrational perturbing frequency is f. Script L is the spectral density of phase noise on one side of a double-sided spectrum. Strictly speaking, for a sinusoidal input at f, L(f) is the level of a discrete sideband rather than a spectral density.

For random vibration, the expression is the same but "a" has the meaning:

$a^2$ = twice the power spectral density of acceleration in $g^2/Hz$.

Methods for reducing the response of an oscillator to vibration include: compensation of the acceleration sensitivity by means of an accelerometer feedback signal applied either to the OCXO voltage tuning point or as a DC signal on the electrodes of a doubly-rotated cut resonator; and a dual-resonator approach in which each resonator $\Gamma$ "compensates" the other, within the oscillating loop. Another method of compensation within the oscillator loop is the use of a g-sensitive reactive element to compensate a single resonator.

CH2588-2/88/0000-334 $1.00 C 1988 IEEE

The first approach has been successfully employed in the FTS 9110-120 oscillator.[4] An accelerometer signal having accurate DC response is fed back to the OCXO linearized tuning network for compensation over the frequency range 0 -100 Hz. Up to 60 Hz, the achieved compensation is equivalent to a ten-fold reduction in $\Gamma$ (20 dB in phase noise), which makes this system particularly useful for shipboard applications where vibration input frequencies are comparatively low.

Feedback signals may also be directly applied to the electrodes of an SC-cut resonator. This approach[8] has shown impressive (40 dB) reductions in phase noise over a frequency range of 5-250 Hz. As in all compensation schemes, this one requires exact alignment of the resonator $\Gamma$ to the effective accelerometer axis. A three-axis accelerometer was used, which permits final angle adjustment to be done electrically. However, application of an electrical potential across the resonator may be a problem from the standpoint of aging.

In each approach, the accelerometer should be in close proximity to the resonator, and both elements must be rigidly mounted to the same structure for good high frequency performance. For universal utility, accelerometers should have drift-free response to DC, and linear response throughout the bandwidth of interest.

The frequency response of the compensating element must be wider than the upper frequencies encountered in vibration. This is seen in Figure 1 which shows typical phase noise under vibration. Vibration induced noise for the example $\Gamma = 10^{-9}$ per g is shown in Curve a. Curve b shows the response when the equivalent $\Gamma$ is compensated to the $10^{-10}$ level. In the vicinity of resonance of the accelerometer, the associated mechanical phase shift causes loss of proper compensation. Thus, at resonance the phase reversal causes the overall effective $\Gamma$ to be twice what it is in the uncompensated case, resulting in a 6 dB enhancement of phase noise above the case of a single uncompensated resonator.

Curve c, of Figure 1 shows the effect of attenuating vibration input to the oscillator package, by means of mechanical isolators having a 130 Hz natural resonance frequency, and 40% damping (Q = 1.6). The unavoidable performance degradation of 20 log Q is seen at the isolator resonance. The lowest curve shows phase noise for quiescent operation of the 10 MHz oscillator (or, under vibration, if the $\Gamma$ compensation were to be perfect).

a) 1 g sinusoidal input, uncompensated, $\Gamma = 10^{-9}$, $f_o$ = 10 MHz

b) accelerometer compensated, but phase change at $f_n$

c) with vibration isolator system, Q = 1.6

d) $\Gamma$ = 0; or quiescent operation

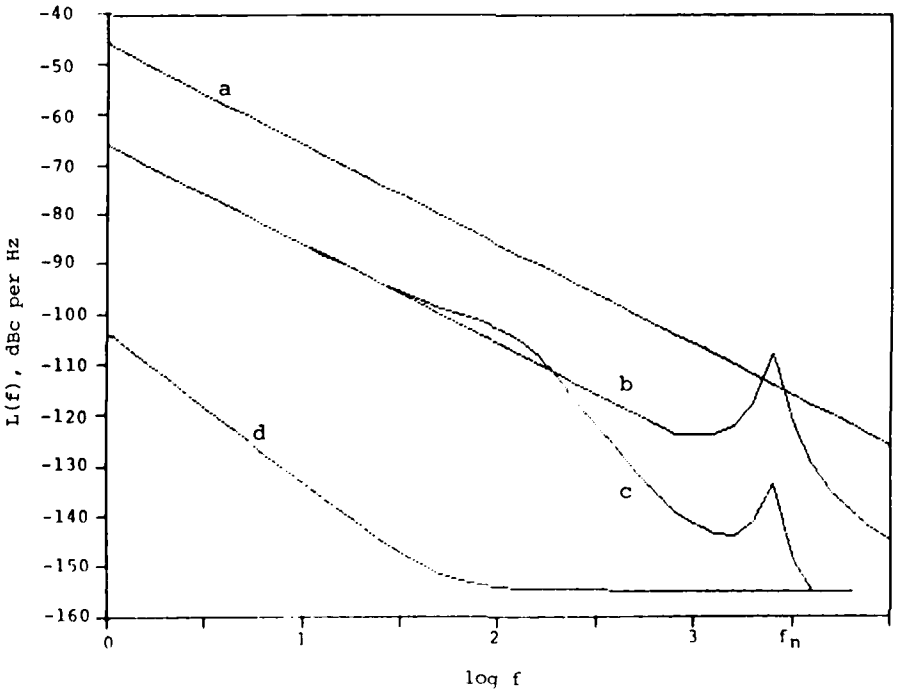



Figure 1. GENERALIZED PHASE NOISE UNDER VIBRATION

An additional effect to consider is that of centripetal acceleration arising from rotational motion of the oscillator package, both constant and varying. If resonator and compensator are separated by some radial distance, 2r, then frequency changes occur, of magnitude $2\Gamma r w^2$. Both resonator and accelerometer see forces which add in phase rather than cancelling. Angular rates of 1 radian/second, with r = 2 inches, cause an equivalent sensitivity of up to 1% of the individual $\Gamma$, depending on spin axis. Oscillatory rotations at a frequency f give rise to sidebands at 2 f.

An additional limitation of compensation schemes is sometimes caused by the temperature dependence of the direction of the resonator $\Gamma$.[9,10] Care must be taken to perform the resonator alignment at the temperature of operation. We have observed angular changes of as much as 0.1 to 0.2 degrees per degree C.

Aside from resonator $\Gamma$ considerations, mechanical resonances within the oscillator loop will degrade phase noise performance. Vibration sensitivity of a conductor or component, for example, can modulate the resonator tuning reactance. Also, vibration of the resonator plate within its enclosure, can modulate $C_o$.

Thus, for example, if a modulation sensitivity of $10^{-5}$ pF per g exists by virtue of a mechanical stray capacitance, this can cause an equivalent "$\Gamma$" sensitivity of $10^{-10}$ per g.

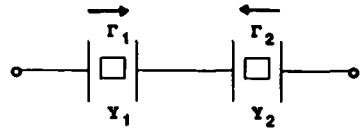
Modulation of capacitance outside the oscillating loop directly affects the phase of the output frequency, producing additive phase noise. Mechanical resonances generally produce narrow L(f) sidebands corresponding to the mechanical Q and energy input. Thus, mechanical resonances anywhere in the oscillator assembly have an impact on performance in a dynamic vibration environment.

FTS 1120

The FTS 1120 10 MHz oscillator uses the dual resonator approach to $\Gamma$ compensation (Figure 2). Resonators are first characterized for acceleration sensitivity; the resonators are then paired according to equality of their acceleration sensitivity. The resonators employed are standard 10 MHz 3rd overtone SC-cut resonators in T0-8 enclosures. Each resonator is aligned to a predetermined axis such that the sensitivity vectors cancel. The resultant achievable $\Gamma$ is a function of how well the resonators are matched. A mismatch of 10% yields a factor of 10 limit on $\Gamma$ improvement. Off-axis sensitivity is a function of angular alignment. A 1° mis-alignment results in a residual orthogonal axis response equal to 3% of $\Gamma$.

A 10 MHz resonator was used for the 1120 design after investigations of 5 MHz resonators showed twice the sensitivity of 10 MHz devices.

In the 1120, the resonators are rigidly mounted in close proximity to one another within a temperature controlled isothermal enclosure. Feedback compensation is implemented directly within the oscillating loop. Each resonator has a companion "accelerometer" of mechanically similar construction, and with similar frequency response to both static and dynamic forces. Figure 3 shows a functional block diagram of the 1120 oscillator.




- $\Gamma$ VECTORS ANTI-PARALLEL
  (1° MIS-ALIGNMENT = .03 $\Gamma$ PERPENDICULAR RESPONSE)
- $|\Gamma|$ MAGNITUDES MATCHED
  (RESULTANT $\Gamma = \Gamma_1 - \Gamma_2$)

Figure 2. DUAL RESONATOR

A rigid oven isotherm supports the quartz resonators as well as the oven control and oscillator electronics. The oven enclosure also functions as the alignment fixture, providing accurate reference planes for locating the resonator $\Gamma$'s. Because these reference planes are orthogonal to each other, it is a straightforward procedure to find the two null axes for a resonator and then rotate the oven 90 degrees to measure $\Gamma$.

Electronics are implemented using hybrid circuit technologies to reduce stray capacitance modulation. Circuit boards are rigidly mounted to avoid flexure-induced capacitance modulation. Heaters are arranged on four sides of the oven assembly, to reduce thermal gradients. The effective oven thermal gain is 70.

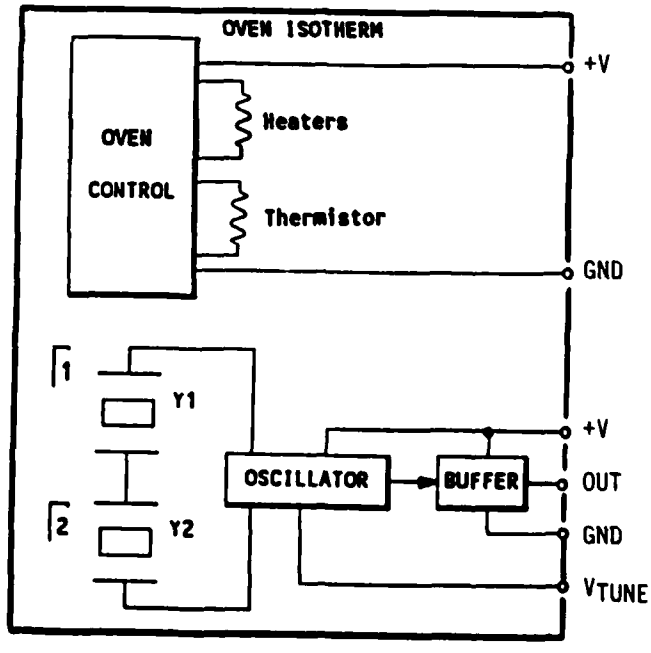



Figure 3. MODEL 1120 BLOCK DIAGRAM

Figure 4 is a view of the 1120 oscillator. The size is 3 x 3 x 4 inches, exclusive of the mounting brackets. Mechanical tuning is controlled via rugged switches in place of a conventional potentiometer to avoid vibration induced frequency shifts. Vibration isolation at high frequencies is afforded by the high-loss rubber grommet mounts, located on opposite corner edges so that the oscillator package is supported through its center of mass. The natural frequency and damping constant of the system were chosen to give attenuation above 180 Hz, with a low Q. Maximum excursions of this isolation system at resonance are minimized.

The theoretical behavior of this isolation system is shown in Figure 5. Resonance occurs at about 130 Hz. A damping ratio 40% of critical gives Q = 1.6. The vertical axis is displayed as 20 log of the transmissibility, versus normalized frequency; this is directly translatable to the expected effect on vibration induced phase noise.

The measured transmissibility of the mounting is shown in Figure 6. Acceleration measured at the package is compared to the vibration table input which was a white random spectrum at the level .001 $g^2$/Hz. Enhancement of 4 to 5 dB is seen at resonance and the attenuation at 1 kHz is 20 dB as expected. The vibration direction is along the Y axis, which is the axis of alignment for the resonator $\Gamma$ vectors.

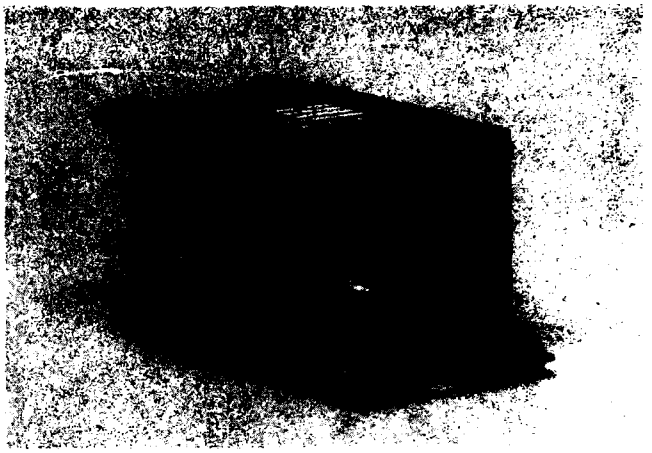


Figure 4. FTS MODEL 1120 OSCILLATOR

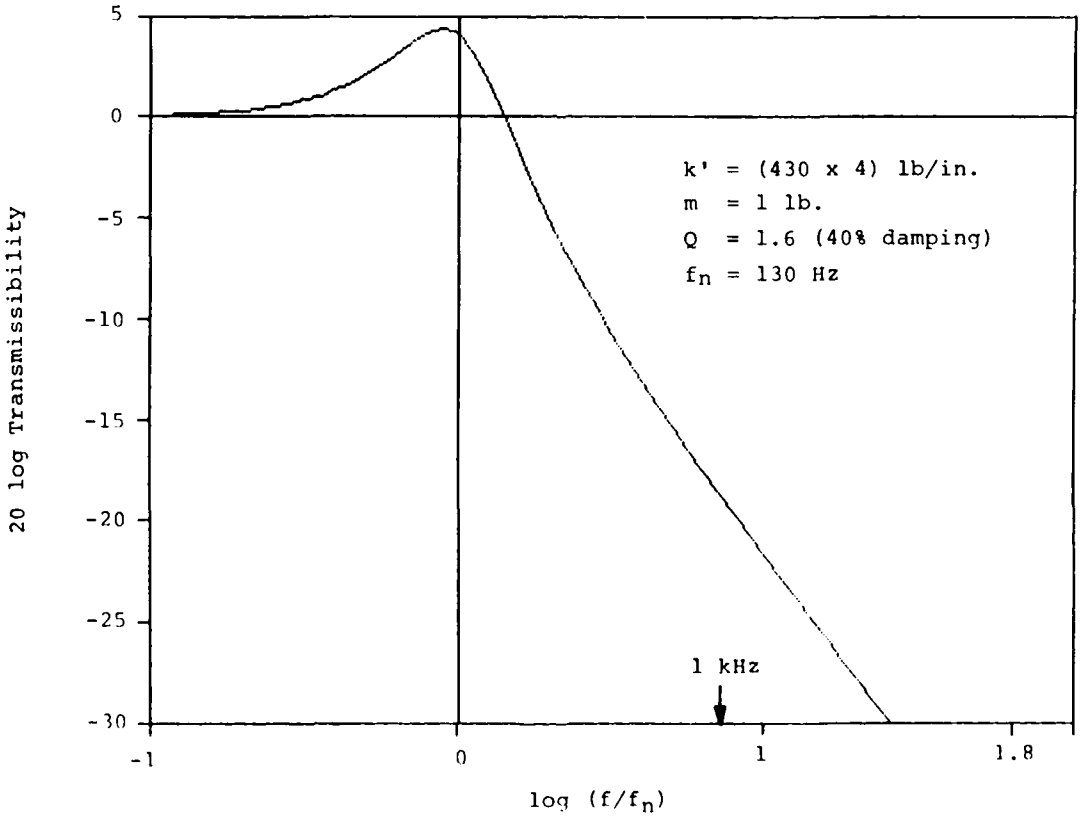



Figure 5. VIBRATION ISOLATION SYSTEM

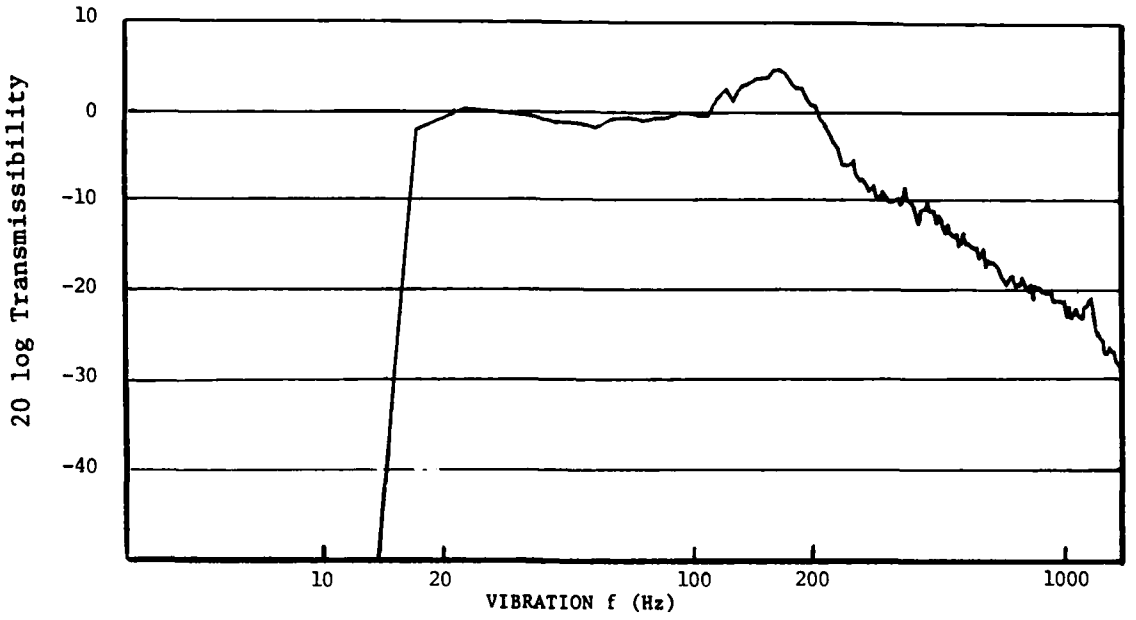



Figure 6. MEASURED TRANSMISSIBILITY FOR Y AXIS

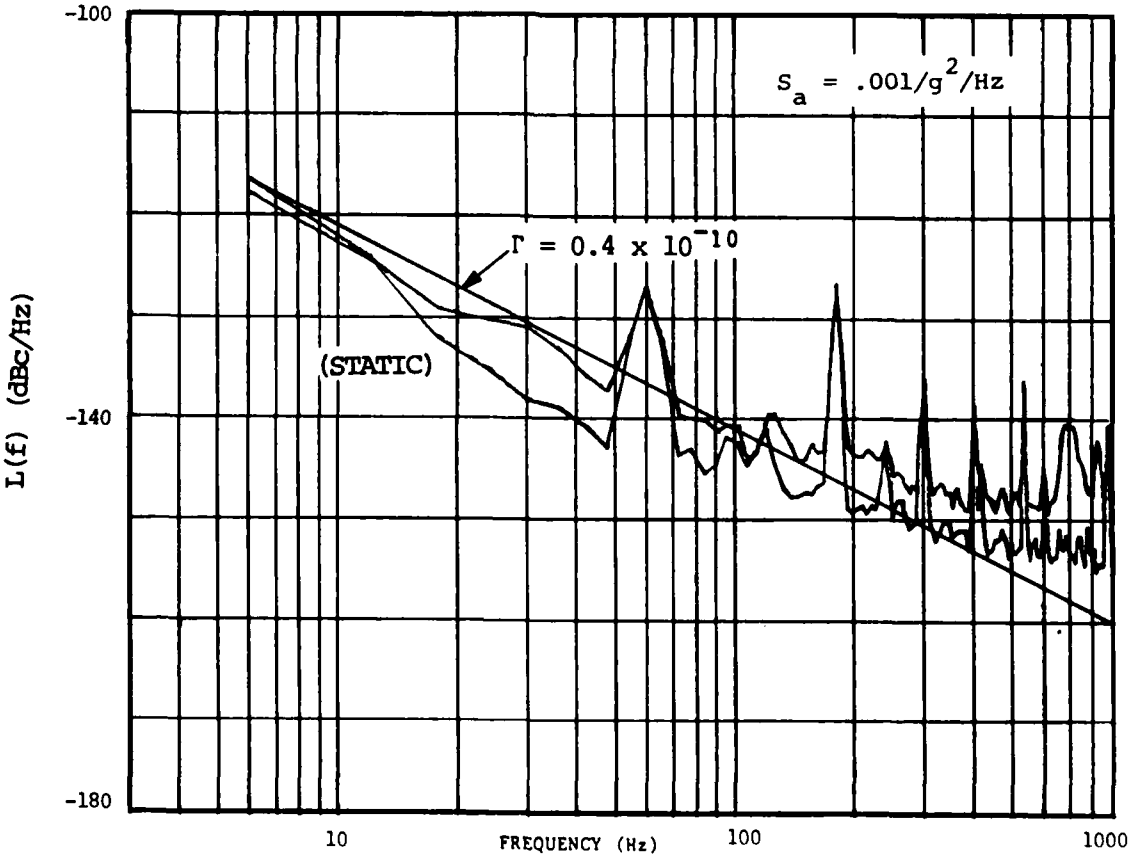



Figure 7. PROTOTYPE MODEL 1120 PHASE NOISE: X-AXIS

## III. PHASE NOISE RESULTS

Two 10 MHz SC-cut resonators were initially characterized under dynamic conditions. Individual $\Gamma$ magnitudes were measured by means of the frequency fluctuations spectrum under vibration, in a frequency locked loop. The results were $1 \times 10^{-9}$/g and $0.7 \times 10^{-9}$/g.

These resonators were then installed in the 1120 oven assembly and aligned so that their acceleration sensitivity vectors were opposed. The assembly was then vibrated in the direction parallel to the resultant vector, the Y axis, and residual sensitivity was found to be $3 \times 10^{-10}$/g.

After final adjustments, the prototype unit was foam potted to prevent lead motion or changes in alignment. This unit was then tested along each of the three axis. The static tip-over g sensitivity was also measured and found to be $3 \times 10^{-10}$ per g.

Figures 7 and 8 show results of testing along the X and Z axes, which are orthogonal to the $\Gamma$ vector direction. Vibration input level was .001 $g^2$/Hz, flat from 5 Hz to 1 kHz. Other prototypes have been tested to 2 kHz. They behave well, when account is taken of spurious vibration table response.

In the X axis response, the equivalent sensitivity is represented by a $f^{-2}$ line showing $\Gamma = 0.4 \times 10^{-10}$ per g. This is a measure of the degree to which the X axis represents a null response; i.e., the success of alignment, good mechanical behavior of the resonator pair, and control of additive vibration-induced noise.

The lower curve in each plot shows phase noise without vibration. The X axis response shows about 6 dB degradation of the quiescent phase noise.

In Figure 8, it is seen that the null alignment along the Z axis is less complete. The equivalent residual sensitivity is $0.75 \times 10^{-10}$ per g.

The axis of principal sensitivity is the Y axis, due to the residual mismatch. A near-perfect match of $\Gamma$ magnitudes would yield low sensitivity along any vibration direction.

Figure 9 shows results of vibration along the Y axis. The theoretical result for an effective $\Gamma = 3 \times 10^{-10}$ per g is shown as a $f^{-2}$ line. Enhancement at the vibration isolator resonance is seen to be 4 to 5 dB as expected. Beyond 200 Hz, the phase noise decreases at 40 dB per decade, reflecting the vibration isolator roll-off.

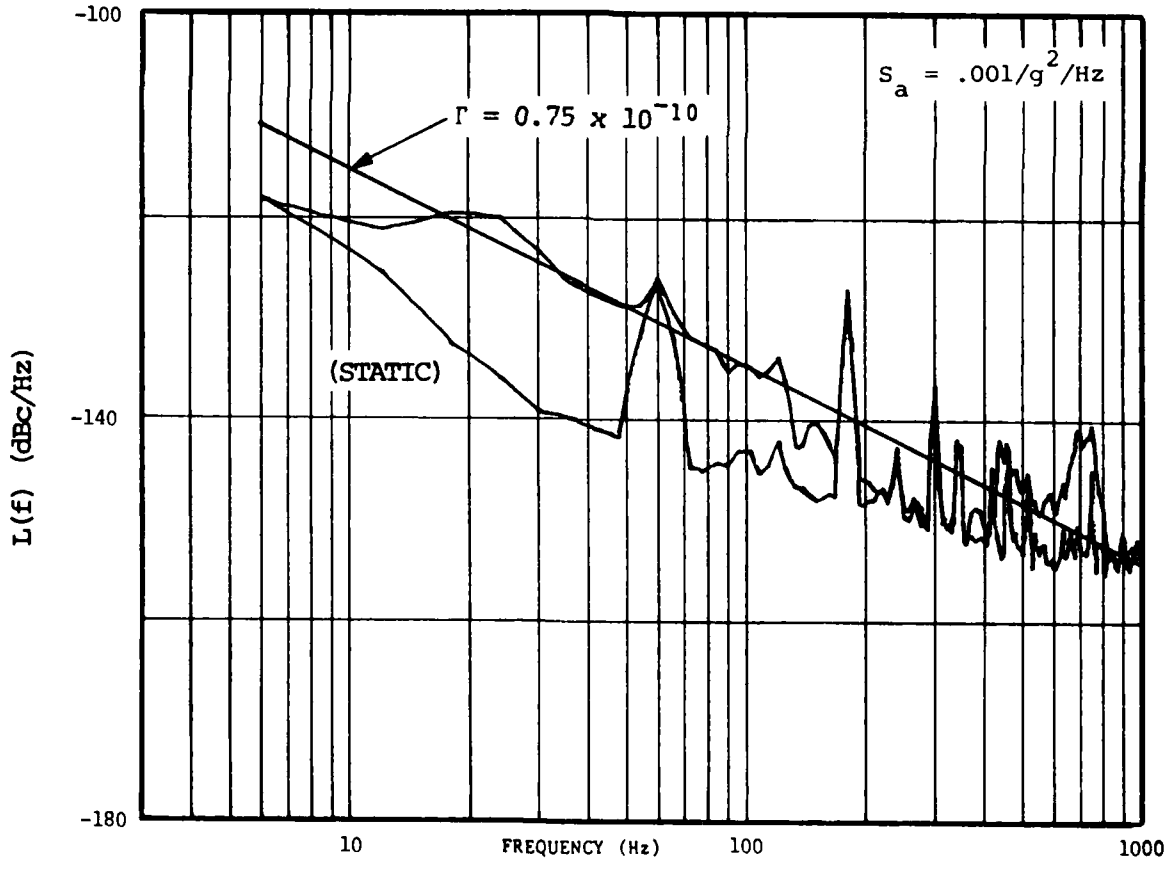



Figure 8. PROTOTYPE MODEL 1120 PHASE NOISE: Z-AXIS

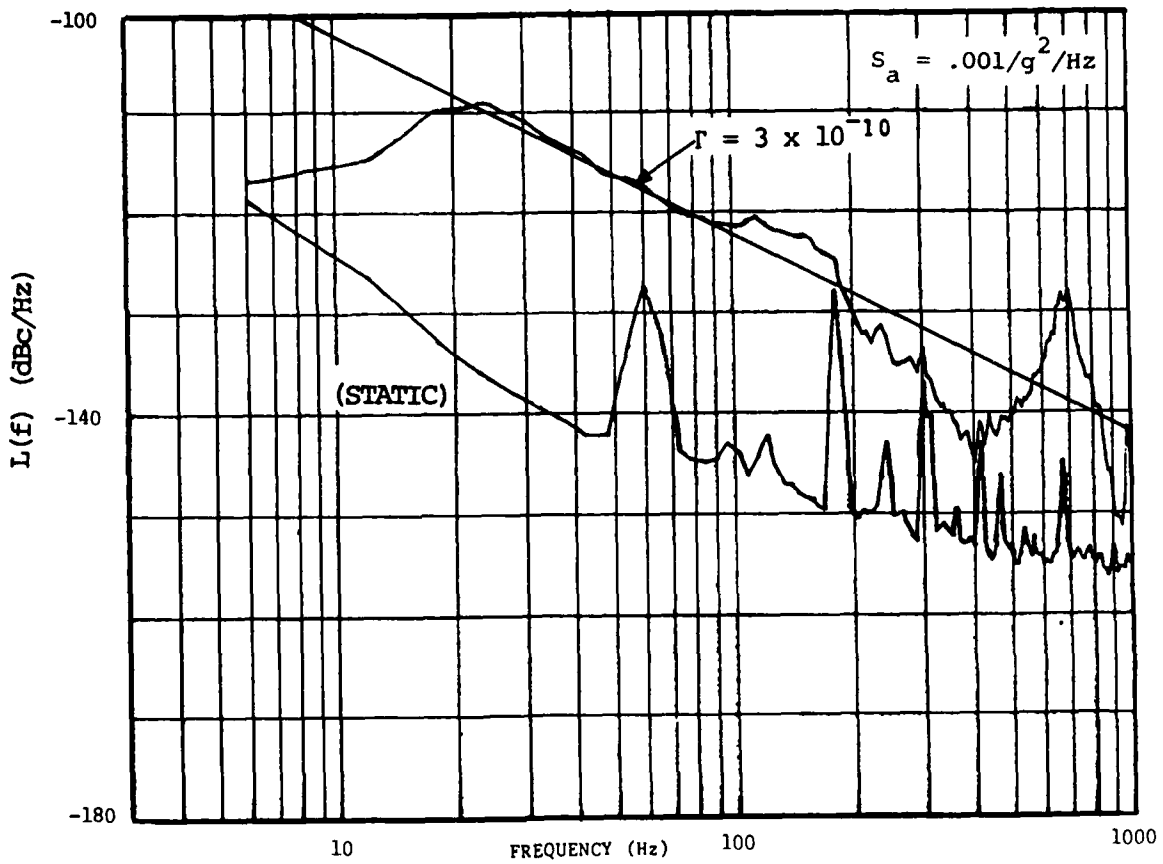



Figure 9. PROTOTYPE MODEL 1120 PHASE NOISE: Y-AXIS

| PARAMETER | PERFORMANCE |
|---|---|
| SIGNAL CHARACTERISTICS | |
| Frequency | 10 MHz |
| Amplitude | 1 Vrms/50 ohms |
| Phase Noise (BW - 1 Hz) for offset of: | |
| | $10^0$ Hz -103 dBc |
| | $10^1$ Hz -133 dBc |
| | $10^2$ Hz -140 dBc |
| | $10^3$ Hz -156 dBc |
| Short Term Stability for averaging time: | |
| | 1 s $1.3 \times 10^{-12}$ |
| | 10 s $5 \times 10^{-12}$ |
| | 100 s $1 \times 10^{-11}$ |
| Harmonic Distortion | -44 dBc |
| Frequency Sensitivity to: | |
| load variation | $2.5 \times 10^{-11}$/1% |
| supply voltage variation | $1 \times 10^{-11}$/1% |
| Electrical Tuning | $\pm 2.2 \times 10^{-7}$ |
| Mechanical Tuning | $\pm 2.7 \times 10^{-7}$ |
| Linearity of Electrical Tuning | 12% |
| Aging | $1 \times 10^{-10}$/day after 7 days |
| Warm-Up | $2 \times 10^{-9}$ in 15 minutes (30 W) |

| PARAMETER | PERFORMANCE |
|---|---|
| ENVIRONMENTAL | |
| Temperature (Operating) | $3.5 \times 10^{-9}$ (-18 to 71 °C) |
| g-sensitivity | Γ = $3.1 \times 10^{-10}$/g (see text) |
| POWER | |
| Oven | 3.1 W (at 25 °C, 21 - 31 V) |
| Warm-Up | 1 A, constant current |
| Oscillator Electronics | 50 mA, 16 - 20 V |
| MECHANICAL | |
| Weight | 0.54 kG |
| Dimensions | 3" x 3" x 4" |
| VIBRATION ISOLATION | |
| Natural Frequency | 130 Hz |
| Q | <2 |

FIGURE 10. FTS 1120 OSCILLATOR FOR HIGH VIBRATION ENVIRONMENTS

The resonant response at 700 Hz is attributable to the secondary frequency response of the two-crystal network as shown by Gagnepain and Walls[5]. This response appears at a frequency which is a function of circuit parameters, chosen to avoid having it occur, for example, in the one or two hundred Hz region. Note that this spurious $\Gamma$ sensitivity is not excited along the X or Z axes. It represents uncompensated $\Gamma$ sensitivity in the Y axis direction (the $\Gamma$ direction).

The net result of $\Gamma$ sensitivity for the prototype oscillator may be stated as: $\Gamma^2 = (0.4^2 + 3^2 + 0.75^2) = 3.1 \times 10^{-10}$ per g. It is dominated of course by the residual unmatched sensitivity along the alignment axis; this can be made arbitrary small, by resonator selection, while the off-axis sensitivities are also made small by alignment.

Figure 10 shows a summary of other performance data for the Model 1120 oscillator, in quiescent operation.

## IV. SUMMARY

An acceleration sensitivity compensation method which uses the similar properties of two resonators to cancel $\Gamma$ within the oscillating loop has been demonstrated. In a prototype FTS 1120 oscillator, using two readily available resonators, the phase noise under vibration is reduced to the level expected for the measured residual value of $3 \times 10^{-10}$. Off-axis sensitivity is seen to be a factor of 4 less than this.

The dual-resonator technique has been discussed in the context of limitations which are common to any type of acceleration sensitivity compensation. A major advantage of this technique is that the mechanical properties and resonant crystals complement each other in a way that other resonator-accelerometer systems do not.

Response of the compensation extends to DC and no external circuitry, frequency shaping, or tuning linearization is needed. In addition, the two resonators are mounted in close proximity within a highly rigid structure.

Because of additive phase noise at higher vibration frequencies, or unavoidable phase shifts in compensated response, most systems require some attenuation of vibrational energy at high frequencies. We have demonstrated a well-damped compact isolation system for the FTS 1120 which attenuates inputs above 180 Hz and has a Q less than 2.

Overall, results on a prototype oscillator having dual-resonator g-sensitivity compensation, show that readily available resonators can be matched in pairs to give an equivalent low value of $\Gamma$ over a broad frequency range.

Acknowledgement

The authors wish to express appreciation for support and encouragement from J. Vig and his group at LABCOM.

## REFERENCES

1. R.L. Filler, "Acceleration Sensitivity of Quartz Crystal Oscillators: A Review", Proc. 41st AFCS (Annual Frequency Control Symposium), 398-408 (1987).

2. M. Valdois, J. Besson and J.J. Gagnepain, "Influence of Environmental Conditions on a Quartz Resonator", Proc. 28th AFCS, 19-32 (1974).

3. J.M. Przyjemski, "Improvement in System Performance Using a Crystal Oscillator Compensated for Acceleration Sensitivity", Proc. 32nd AFCS, 426-431 (1978).

4. D.A. Emmons, R.M. Garvey and B.T. Milliren, "Acceleration Sensitivity Compensation and Symmetry in Quartz Oscillators", 38th AFCS, 80-85 (1984); D.A. Emmons, "Acceleration Sensitivity Compensation in High Performance Crystal Oscillators" Proc. 10th PTTI, NASA Tech. Memo. 80250, pp. 55-82 (1978).

5. J.J. Gagnepain and F.L. Walls, "Quartz Crystal Oscillators with Low Acceleration Sensitivity," Technical Report. NBSIR77-855, National Bureau of Standards (1977).

6. R. Besson, J.J. Gagnepain, D. Janiaud, and M. Valdois, "Design of a Bulk Wave Quartz Resonator Insensitive to Acceleration", Proc. 33rd AFCS, 337-345 (1979).

*7. A. Ballato, "Resonators Compensated for Acceleration Fields", Proc. 33rd AFCS, 332-329, (1979).

8. V.J. Rosati, "Suppression of Vibration Induced Phase Noise in Crystal Oscillators: An Update", Proc. 41st AFCS, 409-412, (1987).

9. C.R. Dauwalter, "The Temperature Dependence of the Force Sensitivity of AT-cut Quartz Crystals," Proc. 26th AFCS, p. 108. (1972).

10. E.P. Eernisse, "Temperature Dependence of the Force Frequency Effect for the Rotated X-cut", Proc. 33rd AFCS, 300-305 (1979).

*Ballato emphasized quartz axis symmetry but Filler established that the resonator sensitive axis is not necessarily correlated with the crystal axes: R.L. Filler, "Acceleration Resistant Crystal Resonators," US Patent No. 4,410, 822, 1983. Additional improvements were made: J.R. Vig and F.L Walls, "Acceleration Insensitive Oscillator, "US Patent No. 4,575,690, 1986.

42nd Annual Frequency Control Symposium - 1988

# THE RELAXATION QUARTZ CRYSTAL OSCILLATORS - DESIGN AND PERFORMANCES

Dragan M. Vasiljević

Faculty of Electrical Engineering, University of Belgrade
Bulevar Revolucije 73, 11000 Belgrade, Yugoslavia

## Summary

This paper describes the rules for analysis and design of relaxation quartz crystal oscillators. We first show two basic topologies of relaxation oscillators and the corresponding mathematical model describing their behaviour. By using the phase space approach we calculate the oscillating condition and frequency and flicker-frequency noise level. Additionally, the characteristics of the optimal oscillator circuit having loaded crystal Q-factor close to unloaded are determined. Based on the decribed analysis, the design rules for optimal relaxation circuit are derived. These rules are demonstrated through the design and realization of the new optimal relaxation crystal oscillator based on two-transistor negative resistance two-ports. Finally, the measured characteristics of the realized oscillator are presented and compared to known relaxation circuits. It was concluded that the optimal relaxation crystal oscillator is a small number of component circuit having low phase noise and small sensitivity on circuit and ambient parameter variations. Also, its design procedure is simple and clear. The main shortcoming of the relaxation quartz oscillator is a limited range of operating frequencies (up to few MHz).

## Introduction

Relaxation quartz crystal oscillators appear in many applications due to their simple design, small number of components and high output signal amplitude. These oscillators also known as quartz multivibrators, were intuitively and experimentally designed. The essential difference between the harmonic and the relaxation oscillators is the existence of voltage or current value jumps in relaxation circuits, happening in very short time intervals. The first analyses have shown[1] that quartz crystal in relaxation oscillator circuits is excited by voltage step function when electronic switch changes state. This generates the sinusoidal quartz current having exponentially decreasing amplitude and oscillating frequency near to the quartz crystal series resonance. The crystal current, in back, controls electronic switch which must have current controlled input characteristics. These results provided new realizations of quartz relaxation oscillators and better understanding of their behaviour. The measurements of quartz multivibrator performances[2,3] show very different results. From poor to very good, close to the best harmonic crystal oscillators. It was observed that various oscillator circuits give different performances using the same quartz crystal unit. It was a sign that more detailed and studious analyses of relaxation quartz oscillators and their better understanding are needed.

The next investigation step was the application of the general oscillating theory of relaxation systems on quartz multivibrator analysis. Many known oscillator circuits were analysed to find their convenient mathematical model. In addition, using phase space approach, the behaviour of the relaxation quartz oscillators was analytically investigated in both steady state and transient regimes. The conditions for oscillation starting, the oscillation building up and frequency and flicker-frequency noise level were calculated.[4,5] The main results of the described analysis are repeated in the first section of this paper.

In the next section we introduce the concept of the optimal relaxation oscillator electric circuit and derive and explain design rules for its realisation. The optimal relaxation oscillator is a circuit having frequency stability determined by quartz crystal unit parameters only.

The realisation of the almost optimal quartz multivibrator is presented in the third section. Starting from two-transistor negative resistance two-ports[6] and described design rules, the almost optimal relaxation oscillator was designed and tested.

In the last section the performances of quartz multivibrators and their fields of applications are discussed.

## The Relaxation Quartz Oscillator Analysis

The main parts of the relaxation quartz oscillator are: active part of the oscillator, its bias circuit and the quartz crystal unit acting as time network. Topologically, all quartz multivibrators can be devided into two classes: oscillators using quartz crystal outside the feedback loop, and circuits with crystal placed in feedback loop, Fig.1. Both circuits obey

CH2588-2/88/0000-342 $1.00 C 1988 IEEE

the same law of behaviour and analysis procedure, as is shown in references 4 and 5.

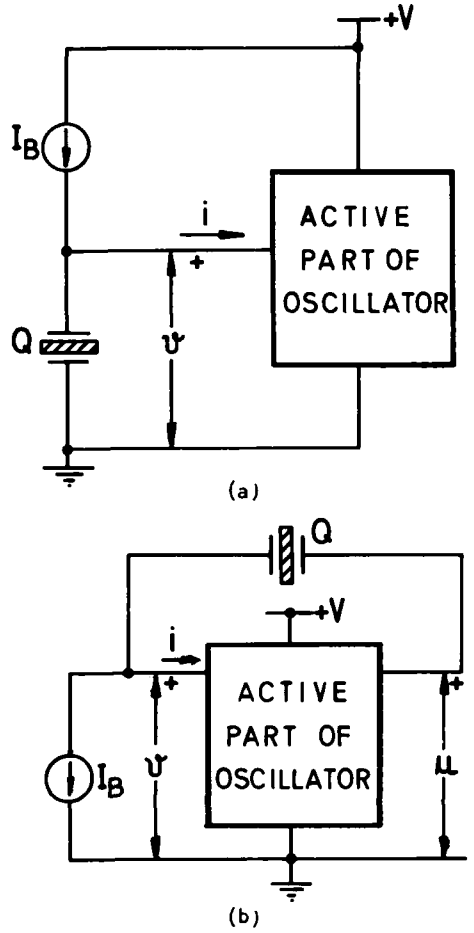



Figure 1. Block-diagram of crystal multivibrator having quartz crystal:(a) outside and (b) inside feedback loop.

We will now demonstrate the analysis procedure for oscillators having crystal unit outside feedback loop (Fig.1(a)). First, we need an electric model of the oscillator. The oscillator active part is modelled by its input static characteristic and by small parasitic inductors and capacitors determining its dynamical behaviour. The supposed input static oscillator characteristic for further elaboration is shown in Fig.2.

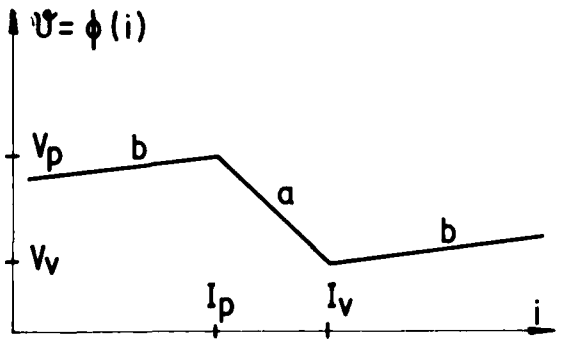



Figure 2. Supposed oscillator active circuit static characteristic.

The bias circuit determines the place of the oscillator singular point on the mentioned static characteristic. This is a direct current circuit, usually a current source. The quartz crystal unit is modelled by a well-known series resonant circuit. So, we obtain a relaxation oscillator electric model shown in Fig.3. The oscillator mathematical model can be easily derived using Kirchhoff's laws and the electric model in the form of the fourth order nonlinear differential equations:

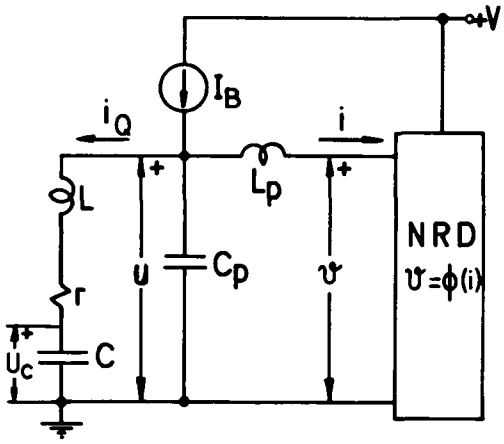



Figure 3. Electric model of oscillator from Fig.1(a).

$$L_P \frac{di}{dt} = u - \Phi(i)$$

$$C_P \frac{du}{dt} = -i_Q - i + I_B \qquad (1)$$

$$L \frac{di_Q}{dt} = u - ri_Q - u_C$$

$$C \frac{du_C}{dt} = i_Q ,$$

where: $L_P$ is parasitic oscillator active circuit input inductivity; $C_P$ is parasitic capacity including quartz crystal unit parallel capacity; $I_B$ is bias current; L is quartz inductivity; C is quartz series capacity and r is a series quartz resistivity. Using a piecewise linear approximation for oscillator input static characteristics, the described equations can be linearized on few segments and solved analytically through phase portrait approach. In relaxation systems the phase portrait contains phase trajectories representing rapid, jump change of the oscillator state and slow motion phase trajectories. For nonlinearity given in Fig.2, origins of fast trajectories are points $i=I_P$ and $i=I_V$. The end of the fast phase trajectory is determined by a jump condition depending on the circuit parameter and direct current operating conditions. For the characteristic in Fig.2, the jump from point $(I_P, V_P)$ finishes in $(i > I_V, V_P)$. Also, the jump from point $(I_V, V_V)$ finishes in $(I < I_P, V_V)$.

The slow motion phase trajectories exist inside interval $I_P < i < I_V$ when oscillations build up from the singular point placed in the same interval. Also, the slow phase trajectories exist in intervals $i < I_P$ and

$i > I_V$ during the steady state oscillations. The character of slow phase trajectories is determined by roots of characteristic equation of system (1) in the case of slow motions when $L_p$ and $C_p$ are zero:

$$\lambda^2 + \frac{\omega_r}{Q_o}\left(1 + \frac{\Phi_i'(i)}{r}\right)\lambda + \omega_r^2 = 0, \qquad (2)$$

where: $Q_o$ is unloaded quartz crystal unit Q factor and $\omega_r$ is series resonant frequency $\omega_r = 1/\sqrt{LC}$. The roots of the characteristic equation:

$$\lambda_{1,2} = -\frac{\omega_r}{2Q_o}\left(1+\frac{\Phi_i'(i)}{r}\right) \pm j\omega_r\sqrt{1 - \frac{\left(1+\frac{\Phi_i'(i)}{r}\right)^2}{4Q_o^2}}, \qquad (3)$$

depend on the slope of the oscillator active part input characteristic.

The steady state oscillations are represented in the phase portrait by close phase trajectory. This is a well-known limit cycle. For the relaxation oscillator under analysis, the limit cycle contains parts of slow phase trajectories in segments of static characteristic having slope b (Fig.2), and parts of rapid phase trajectories starting in points $(I_v, V_v)$ and $(I_p, V_p)$, Fig.2.

All previous results can be easily calculated or proved using the detailed analysis shown in references 4 and 5.

## The Oscillator Behaviour

The Oscillating Condition. For reliable oscillation starting, the oscillator singular point has to be unique and unstable. The number of singular points is determined by a bias circuit. The singular point character is determined by the roots of characteristic equation. For the analysed model in Fig.3, current source $I_B$ defines the singular point place in the middle of interval $I_p < i < I_V$ in Fig.2. This singular point is of the focus type due to the high value of the quartz crystal unit Q factor. To obtain unstable singular point, root of characteristic equation must have positive real part. This is proved when the absolute value of static characteristic slope, $|a|$, is larger than quartz crystal unit series resistance r.

The Oscillation Buildup. The character of the singular point determines the oscillator transient regime. For the analysed example, the singular point of the focus type means that oscillation builds up exponentially. It means that the amplitude increases slowly.

The Oscillation Period. We show that the limit cycle contains two segments of fast trajectories and two segments of slow trajectories. If we adopt the assumption that jumps are instantaneous, then time when system moves through slow trajectories of limit cycle, determines the oscillation period. The velocity of the change of the oscillator state depends on the imaginary part of the characteristic equation roots. Because, the root value depends on the oscillator input static characteristic slope, the velocities on various segments in Fig.2 are different. In the special case, when all segments of static characteristic containing slow motion parts of the limit cycle have the same slope, the oscillation period can be calculated from:

$$\frac{2\pi}{T} = \omega_r\sqrt{1 - \frac{\left(1+\frac{\Phi_i'(i)}{r}\right)^2}{4Q_o^2}} \qquad (4)$$

This oscillation period remains constant when the oscillator changes the time of residence on various static characteristic segments. It means that ambient and circuit parameters which change the shape of the limit cycle have no influence on the oscillation period. This is a case of the most practical interest in the oscillator design.

The Oscillator Loaded Q Factor. If we avoid short time intervals corresponding to jumps (rapid motions), the dynamical impedance between quartz crystal ports in the oscillator in Fig.3 is equal to the slope of the oscillator input static characteristic. For a constant slope in all segments, the loaded Q factor value is equal:

$$Q_L = \frac{Q_o}{1+\frac{\Phi_i'(i)}{r}} = \frac{Q_o}{1+\frac{b}{r}}. \qquad (5)$$

From the last equation it seems that low static characteristic slope gives minimal Q factor degradation and the best frequency stability.

Flicker-Frequency Noise Level. There is experimental law for flicker frequency oscillator noise level calculation[7]:

$$S_y(f) = \frac{10^{-12.75}}{Q_o^2 f} + \frac{a_L}{4Q_L^2 f}, \qquad (6)$$

where $a_L$ is an experimentally determined constant in the range $10^{-11} \div 10^{-13}$. It is clear that high $Q_L$ gives low noise.

## The Design of Relaxation Quartz Oscillators

We can now introduce concept of the optimal quartz relaxation oscillator as a circuit whose frequency stability is determined by quartz crystal unit, only. It means that the crystal unit resources are fully exploited in the optimal oscillators. From the discussion in the previous sections, we can see that the optimal relaxation quartz oscillator can be realized by using oscillator active circuit static characteristic having zero slope on segments containing limit cycle slow phase trajectories (slope b in Fig.2, Eqns (4), (5)). It is supposed also that rapid motions are instantaneous. The real oscillator active circuits have static characteristic slopes different from zero. In these circumstances, almost optimal oscillator can be designed if:

(a) all static characteristic segments containing limit cycle slow trajectories have the same slope $S_i = S$,

and,
(b) if slope S is as low as possible.
The condition (a) provides stable oscillation period for the circuit and ambient parameter variations. The condition (b) gives lowest crystal Q factor degradation in oscillator circuit.

The design rules can be formulated from previous discussion.
1) The oscillator active part static characteristic slopes $S_1$ must be equal and as low as possible in order to obtain the best frequency stability.
2) The oscillator singular point must be unique and unstable for reliable oscillation starting.
3) The shape of static characteristic must provide appropriate excitation power level for quartz crystal unit.
4) The active part oscillator circuit should have high switching speed for wide range of operating frequencies.
5) Also, standard design requests have to be fulfilled:
   - stable power supply,
   - load isolation,
   - temperature stability,
   - printed circuit layout and package.

## The Realization of Almost Optimal Relaxation Oscillator

The general procedure for two transistor two-ports negative resistance devices (NRD) is known in the literature[6]. Also, the tables of NRD realizations are published. Starting from these data we chose current controlled NRD shown in Fig. 4(a). The selected NRD is

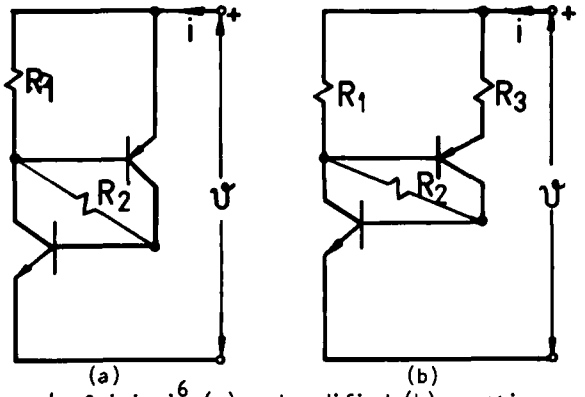



Figure 4. Original[6] (a) and modified (b) negative resistance two transistor two-ports (NRD): (a) $R_1$=680Ω, $R_2$=8,2kΩ; (b) $R_1$=100Ω, $R_2$=8.2kΩ, $R_3$=250Ω.

modified as in Fig.4(b) to obtain static characteristic shown in Fig.2. The resistor $R_1$ value is lowered to rise peak current value $I_p$. The resistor $R_3$ is introduced to match static characteristic slopes for $i>I_v$ and for $i<I_p$. Also, for the same purpose, npn transistor should have high current gain. The measured static characteristics of circuits in Fig.4 are shown in Fig.5. The obtained segment slopes for modified circuit are: +200Ω, -1000Ω,+100Ω, from the left to the right, Fig.2.

The number and position of singular points are determined by direct current NRD bias circuit. For NRD in Fig.4(b) the direct current constant current source provides only one singular point at any point of static characteristic. The unstable character of the singular point is ensured if absolute value of the negative slope of the static characteristic is larger than quartz crystal unit series resistance, Eqn (3). This condition can be fulfilled by the choice of $R_2$, Fig.4 (b), when singular point is placed on the segment of the static characteristic having negative resistance.

The quartz crystal unit power level depends on the difference of the peak and valley voltages in static characteristic in Fig.2. They can be adjusted by resistors $R_2$ and $R_3$ in circuit in Fig.4(b).

The switching speed of the adopted NRD depends on the transistors switching times, dominantly.

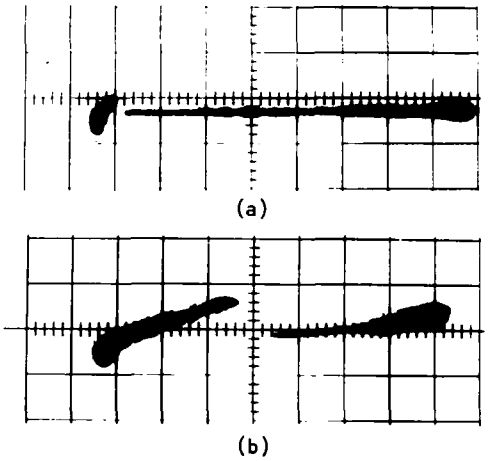



Figure 5. Static characteristic of NRD shown in Fig.4. Vertical: voltage 2V/div; horizontal: current 2mA/div.

Using NRD shown in Fig.4(b), transistor current source and quartz crystal unit, the new almost optimal relaxation oscillator is designed and shown in Fig.6.

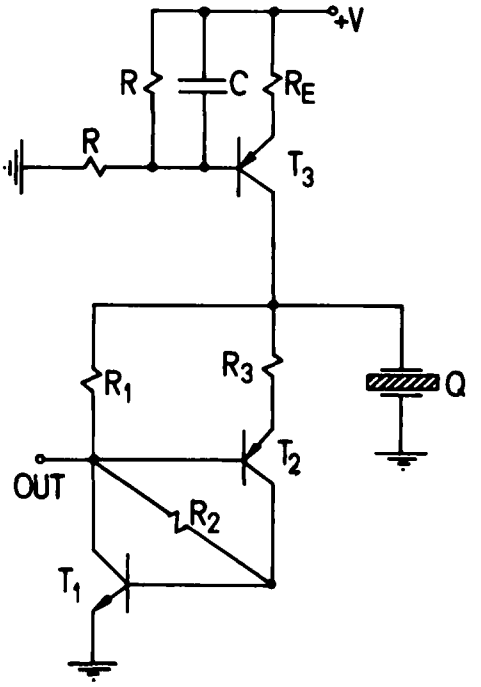



Figure 6. The realization of almost optimal relaxation quartz oscillator: $T_1$=$T_3$=BSX28, $T_2$=BSX29, R=1kΩ, $R_E$=0.5kΩ, C=4.7nF, $R_1$=100Ω, $R_2$=8.2kΩ, $R_3$=250Ω, crystal: r=141Ω, $f_r$=999 913Hz, $Q_o$=200 000.

Certainly, we made some compromise to fulfill many design requests using only three circuit elements: $R_1$, $R_2$, $R_3$.

## The Experimental Results

The circuit in Fig.6 is realized and tested. Its limit cycle shown in Fig.7 has almost equal shape in right and left half. It means that slow phase trajectories of limit cycle lie on static characteristic segments having the same slopes. The horizontal limit cycle segments represent rapid oscillator state changes. The designed oscillator has above average performances[2]. The power supply sensitivity is 0.3ppm/V and temperature sensitivity is 0.005ppm/°K where quartz crystal unit is held on constant temperature. The short-term

Figure 7. The oscillator limit cycle. Horizontal: i = 2.5 mA-pp ; vertical: $u_C$ = 2 V-pp. The voltage $u_C$ is obtained from capacitor C=0.5 nF connected in series with crystal.

frequency stability was measured, too . The results given in Fig.8 show low flicker-frequency noise level due to the low Q factor degradation. The first term in the right side of equation (6) represents flicker noise level introduced by quartz crystal unit, only. For selected quartz crystal unit Q factor $Q_0$=200 000, this noise level is L(f)= -115dB on 1 Hz offset frequency. The measured oscillator noise was about - 100 dB. It implies Q factor degradation on 20%, $Q_L = 0.2Q_0$ . The white phase noise level is emphasized due to the transistors operation in saturation.

## Discussion

The measurements of relaxation quartz oscillators give various results[2,3]. If we follow classification in reference 2, the designed almost optimal oscillator has above average performances.

The advantages of relaxation quartz oscillators are:

+ easy design ;
+ small number of components ;
+ low Q factor degradation ;
+ low sensitivity on ambient and circuit parameter variations ;
+ low close to carrier phase noise ;
+ high output signal level.

The main shortcomings of relaxation crystal oscillators are:

- limited range of operating frequencies ;
- emphasized white phase noise level.

Finally, the next question strikes: what is the key advantage of relaxation quartz crystal oscillators related to harmonic quartz crystal circuits? This is a possibility to obtain minimal Q factor degradation in relaxation circuits.

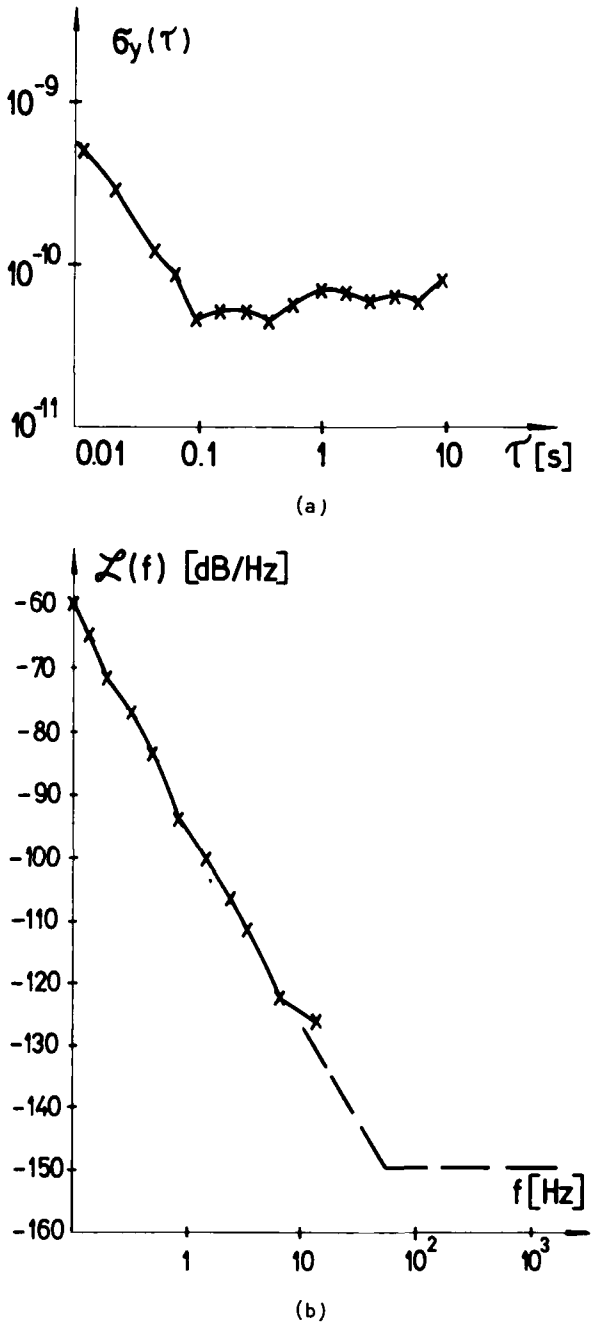



Figure 8. Phase noise of designed oscillator: (a) Allan deviate; (b) amplitude spectar.

## Conclusions

have shown analysis and design procedures for relaxation quartz crystal oscillators. Also, the concepts of optimal and almost optimal relaxation crystal oscillator was introduced. The new almost optimal oscillator was designed and tested to confirm described theory. The design of really optimal relaxation oscillator is open question for future elaboration based on the presented theory.

## Acknowledgements

The author wishes to thank Milan Stojanović for help in the experimental procedures.

## References

1. S.Tesic,D.Vasiljevic,"A quartz complementary multivibrator", Proc. IEE, Vol.123,pp.851-854,Sept.1976.

2. R.J.Matthys,"Crystal Oscillator Circuits".New York: Wiley, 1983, Ch.12.

3. D.M.Vasiljević,"Short-term frequency stability of relaxation crystal oscillators",IEEE Trans. Instr. Meas.,Vol.IM-33,pp.315-322,Dec.1984.

4. D.M.Vasiljević,"Analysis and design of the relaxation quartz crystal oscillators" in Proc.38th Ann. Freq. Control Symp.,Philadelphia, May 29-June 1, 1984.

5. D.M.Vasiljević,"Relaxation quartz oscillators with crystal placed in feedback loop: analysis and design", IEEE Trans.Circ.Syst., Vol.CAS-33,No.8, pp.897-904, August 1987.

6. L.O.Chua,J.B.Yu,Y.Y.Yu,"Negative resistance devices", Int.J.Circuit Theory and Applications,Vol.11, pp. 161-186,1983.

7. V.F.Kroupa," Noise in precision oscillators", in Proc. Conference on Electronic and Piezoelectric Elements,Liberec,Czechoslovakia,July 6-8,1983.

42nd Annual Frequency Control Symposium - 1988

# CLARIFICATION AND A GENERALIZED RESTATEMENT OF LEESON'S OSCILLATOR NOISE MODEL

Benjamin Parzen

consulting engineer
San Diego, CA 92103

## SUMMARY

Leeson's oscillator noise model and its limitations are examined. A modified model, without the limitations and based upon oscillator analysis in the immittance domain, is stated. Computer aided analysis is used to set up and demonstrate the oscillator operation in accordance with the model.

## 1. INTRODUCTION

This paper reports on some work on the investigation of oscillator phase noise performed with the aid of the theory and program described later in this paper.

Leeson's oscillator noise model (LONM), relating the circuit noise to the oscillator noise, has proved to be an extremely useful tool in oscillator work since its publication (1). Numerous other writings have been published on the LONM and its applications among which are Refs. 2 thru 4. One common application is the calculation of the oscillator operating Q, Qop, from the experimentally derived oscillator noise plot.

## 2. CLARIFICATION OF LEESON'S MODEL

2.1 Leeson's expression for the oscillator noise output voltage, Ve, is

term--1 term----------2

$$S_{\phi Ve}(f) = S_{\phi Vo}(f) \left[1 + \left[\frac{fo}{2\,Qop\,f}\right]^2\right] \quad (1)$$

where term 1 is the relative circuit noise density and term 2 is the oscillator relative noise output voltage density for unit white circuit noise.

Other writings (2,3,4) have derived from Eq 1, the more basic and simpler expression for the resonator relative noise current density, I, and voltages directly proportional to I.

term --1 term--------2

$$S_{\phi I}(f) = S_{\phi Vs}(f) = S_{\phi Vo}(f) \left[\frac{fo}{2\,Qop\,f}\right]^2 \quad (2)$$

Qop is calculated from Eq 1.

2.2 When comparing the program results to those obtained from Eqs 1 and 2, discrepancies were found. The derivation of LONM was then examined and it was discovered that the model had implicit assumptions not made in the program.

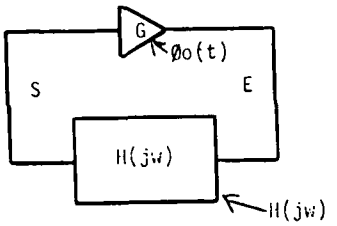
It also became evident that the use of Eq 1 for determining Qop is valid only for few practical output circuit configurations and it is likely that the values of Qop usually calculated from Eq 1 are too small, as implied in Ref 4.

Leeson published the model without formal proof. Sauvage (2) published the proof. Fig 1a is Sauvage's ocillator diagram based upon the feedback oscillator model. Sauvage points out that for Eq 1 to be true,

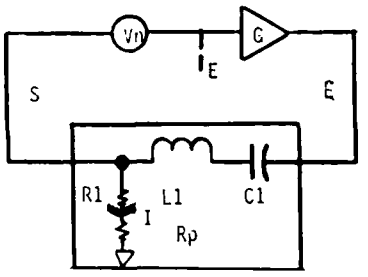
$$H(jw) = a / ( a + jw ) \quad (3)$$

where,

$$a = wo / 2Qop \quad (4)$$




a. General Model




b. Eq 3 H(jw) Model

Fig 1 Sauvage's Oscillator Model

Fig 1b is Fig 1a redrawn to show the H(jw), required by Eqs 3 and 4, as an idealized resonator, having components L1, C1, and R1. It should be noted that this resonator is not a good appproximation for a crystal resonator since an R1 terminal is not available in the crystal and Co is absent. Rp is provided to account for the difference between the resonator Q and Qop. Also, for simplicity, the noisy phase generator of Fig 1a has been replaced by an equivalent noisy signal source producing Vn (f).

CH2588-2/88/0000-348 $1.00 C 1988 IEEE

The phase noise spectrum at point E is that of Eq 1 and the phase noise spectrum of current I and at point S is that of Eq 2. The value of term 2 of Eq 2 drops 6 dB per octave while the value of term 2 of Eq 1 is initially that of term 2 of Eq 2 and approaches 1 as f increases. If the oscillator output is taken at point E or equivalent point then a good approximation of Qop can be calculated. However, an output taken at point S contains no information on Qop.

In practice, the oscillator output is taken at some intermediate point, a too small value of Qop results, and the apparent value obtained is strongly influenced by the secondary noise sources, including resonator noise.

2.3 Leeson's model has contributed greatly to oscillator design and analysis. However, improvements may be, now, practical and should be considered. Knowledge of the potential areas of improvement can be acquired from the following listing of the minor limitations of the present model.

It is too simplistic as it includes only 2 signal points.
It is not valid at f approaching 0.
It is resticted to one, but very useful, idealized filter (resonator).
The phase noise is always symmetrical with respect to f.
It ignores the limiting process present in every real oscillator.

## 3. A RESTATEMENT OF LEESON'S MODEL

3.1 During the past 10 years, the writer has been developing theory for oscillator analysis in the immittance domain. Effective and facile application of the theory to real problems also necessitated the creation of a PC computer circuit analysis program. The results of this effort are

A remarkably simple oscillator theory which fully describes the oscillator operation and which is readily translatable into the real world.
A very user friendly multipurpose circuit analysis program.

The theory and program are universal in that they apply to all oscillators, past, present, and future.

The theory and program are symbiotic, in that
To apply the program to oscillators, the theory must be used.
The program is very helpful and almost indispensable in applying the theory to real oscillator problems including that of improving the theory.

It should be noted that, as the program does not incorporate any information peculiar to oscillators, it follows that the theory may be used, in conjuction with other circuit analysis programs, to analyze oscillators, but with much greater, and perhaps prohibitive, difficulty.

3.2 The theory is based upon a combination of 2 mutually compatible oscillator models

The negative resistance model as the primary model.

The noise source, amplifier, filter, model (5), in which, positive feedback produces the necessary extremely high effective amplifier gain, as the secondary model.

At steady state, in any mesh, of a hypothetically noiseless oscillator,

$$Zt = \sum Z = 0 \qquad (5)$$

From which,

$$\sum X = 0, \qquad \sum R = 0 \qquad (6)$$

Defining, $RN = \sum R_{negative}$, $RT = \sum R_{positive}$ (7)

Then

$$RT = -RN \qquad (8)$$

3.3 The total frequency, F, is described by

$$F = fo + f$$

where fo is the nominal frequencuy
f is the offset or Fourier frequency

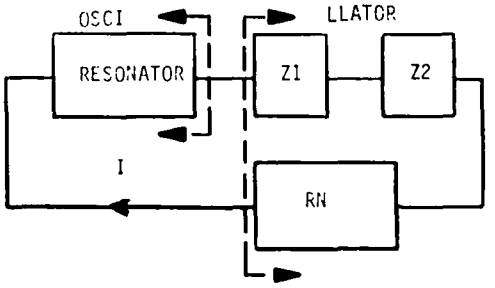



Fig. 2 Noiseless Oscillator Model

3.4 Fig 2 is the simplified diagram of the steady state oscillator resonator mesh. (The mesh containing the resonator is usually chosen for analysis as it has the most information.)

As shown in Fig. 2, that part of the oscillator containing the resonator circuitry is called the osci, symbol os. The remainder is called the llator, symbol LL.

from Eq 5, $Zos = -ZLL$ (9)
Also, $\Delta F = -\Delta XLL / (dXos/dF)$ (9a)

Eq 9a enable the easy determination of oscillator frequency shifts caused by llator changes. This is very useful for investigating the contribution of the llator, as influenced by the resonator charactristics, to the steady state oscillator frequency.

3.5 Fig 3 is the complete diagram for any real oscillator at steady state. (The real oscillator always includes at least 1 noise source) The noise source, Vn, represents the total equivalent noise contributions of all the noise sources in the oscillator. The real oscillator always contains a non physical resistance, dR. The value of dR is determined by the oscillator limiting system which sets the level of the oscillator output.

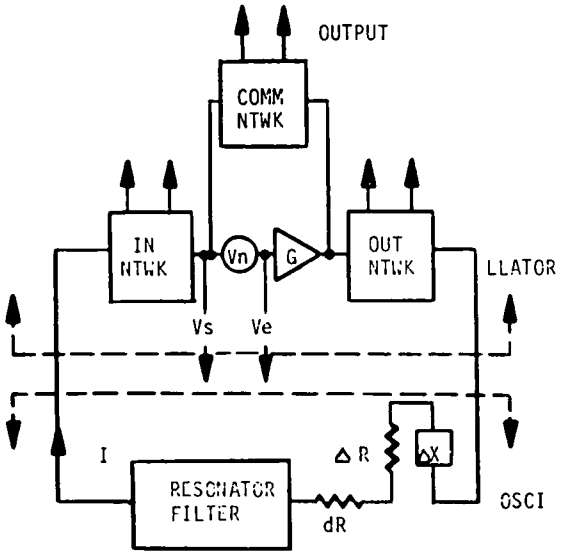



Fig 3. Real Oscillator Model

The oscillator, at steady state, satisfies Eqs 10 to 12.

$$Zt(0) = \sum Z = d\dot{R} \qquad (10)$$

From which,

$$\sum X = 0, \qquad \sum R = d\dot{R} \qquad (11)$$

Also,

$$\left[\frac{RN}{dR}\right]^2 = \left[\frac{Vs(0)}{Vn(0)}\right]^2 \qquad (12)$$

$$RT = -RN + dR \qquad (12a)$$

$$\text{Let } (RN/dR)^2 = (Ar)^2 \qquad (12b)$$

3.6 In the real oscillator,

a. The oscillator is constructed and power is applied.

b. The operator sets the output level by adjusting the limiting circuitry.

c. The oscillator determines fo, Vn, RT, dR and all other operating conditions.

d. The oscillator determines the phase noise as a function of f.

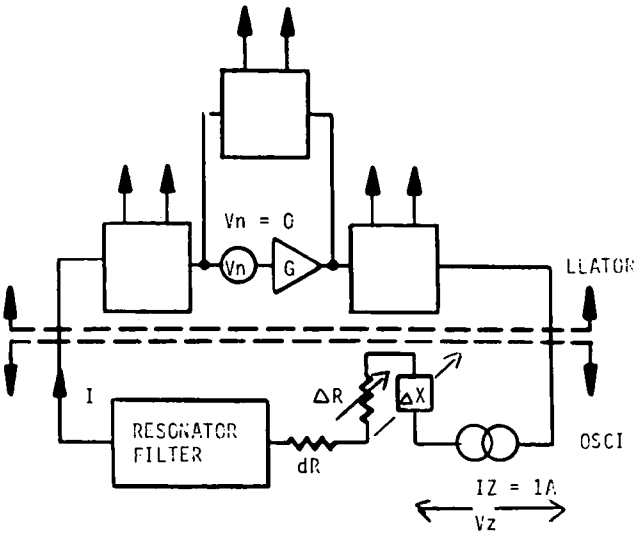



Fig.4 Real Oscillator ( Z(t) Config.)

3.7 In the program oscillator (See Figs. 3 and 4)

a. The oscillator is constructed by entering the circuit into the computer.

b. The program sets the effective gain of the active circuitry.

c. The operator sets fo.

d. The operator sets dR = 0, Vn = 0, and adjusts the circuit elements so that Eq 6 is satisfied for dR = 0 to the extent that in the mesh impedance, for Zt = $V_Z$ in Fig 4, Xt and Rt < $10^{-4}$ the dR of Step f. (The program includes provisions for facilitating this operation.)

e. The operator sets Vn in Fig 3 to Vn real.

f. The operator sets dR to the value where the desired oscillator output is obtained as measured by the ac current at any 1 point or ac voltage between any 2 points.

g. The program determines the voltages and currents at all other points.

h. The program determines the phase noise at and between all points.

i. The program determines Zt(f), (see Fig 4) when dR is at the value of Step f, and Vn is set to 0.

j. $Zt(f) = dR + \Delta Ros(f) + \Delta RLL(f) + j\,Xt(f)$
where, $\Delta Ros(f) = Ros(f) - Ros$
$\Delta RLL(f) = RLL(f) - RLL$

3.8 Steps 3.6d to f are greatly eased by providing the circuit with the trimming elements $\Delta X$ and $\Delta R$ and the 1 amp current source, Iz, shown in Fig 4.

$\Delta X$ and $\Delta R$ are for step d.

Iz is provided to yield Zt which is equal to the voltage across Iz. Iz is present only during steps d and i. It is shorted out at all other times.

3.9 The phase noise in Fig 3 is given by

$$\underset{I}{\mathcal{L}}(f) = \underset{Vs}{\mathcal{L}}(f) = \underset{Vn}{\mathcal{L}}(f) \left[\frac{RT}{Zt(f)}\right]^2 \quad \text{(13)}$$

(term----1 term--------2)

$$\underset{Ve}{\mathcal{L}}(f) = \underset{Vn}{\mathcal{L}}(f) \left[1 + \left[\frac{RT}{Zt(f)}\right]^2\right] \quad \text{(13a)}$$

where $$\underset{Vn}{\mathcal{L}}(f) = \frac{Sp\ Vn(f)}{(Vsc)^2} \quad \text{(14)}$$

where Sp indicates "power spectrum of" and Vsc indicates "carrier voltage at point S".

An additional useful expression is

$$Sp_{Vs}(f) = Sp_{Vn}(f) \left[\frac{RT}{Zt(f)}\right]^2 \quad \text{(15)}$$

Eq 13 is the equivalent of Eq 2 and in conjuction with Fig 3 is the restatement of Leesons's model. Also Eq 13a is the equivalent of Eq 1.

Since Zt(0)=dR, it follows that Eq 13 is valid for all values of f.

Eqs 1 and 2 are valid for only an idealized single mesh resonator. Eq 13 is valid for all resonators.

Eq 15 is useful for the case where Vn(f) has a flicker noise component which theoretically approaches infinity as f approaches 0. This equation states that the oscillator noise power spectrum, at f when Xt(f) >> dR, is independent of dR. Therefore, at these values of f,

$$\underset{Vs}{\mathcal{L}}(f) = Sp\ Vs(f) \ / \ (V'sc)^2 \quad \text{(16)}$$

where V'sc is the arbitrary desired carrier V.

3.10 For the idealized single mesh resonator,(e.g. crystal resonator less Co), Eq 13a becomes

$$\underset{Ve}{\mathcal{L}}(f) = \underset{Vn}{\mathcal{L}}(f) \left[1 + \frac{1}{(Ar)^{-2} + (2\ Qop\ f/fo)^2}\right] \quad \text{(17)}$$

Eq 17 differs from Eq 1 in the $(Ar)^{-2}$ term.

3.11 For larger values of f, Eq 17 can be approximated by

$$\underset{Ve}{\mathcal{L}}(f) = \underset{Vn}{\mathcal{L}}(f) \left[1 + \frac{1}{(2\ Qop\ f\ /\ fo)^2}\right] \quad \text{(18)}$$

which is identical to Eq 1.

3.12 When f approaches 0,

$$\underset{Ve}{\mathcal{L}}(f) = \underset{Vn}{\mathcal{L}}(f) \frac{1}{(Ar)^{-2} + (2\ Qop\ f\ /\ fo)^2} \quad \text{(19)}$$

which differs substantially from Eq 1. This equation also clearly yields the surprising result that, for white noise, the signal 3 dB bandwith, BW

$$BW = fo/(Ar * Qop). \quad \text{(20)}$$

3.13 Note that Eq 13 applies to a point with the best phase noise floor and Eqs 13a and 17 to a point with the worst phase noise floor. In practice, the output noise floor is usually somewhere between the two.

It will be noted that the limitations of Par 2.3 are absent in the restated model.

## 4. CONCLUSIONS

Leeson's Oscillator Noise Model, including its assumptions and limitations has been clariified.

Oscillator theory has been presented and a more correct and universal oscillator noise model has been stated.

The importance of Zt(f), RT, and dR and the tremendous power of analysis in the immittance domain complemented by computer aided analysis have been demonstrated. An important additional advantage is that this anlysis method provides considerably greater understanding of the operation of the real oscillator.

## 5. REFERENCES


1. Leeson, D. B., "A Simple Model of Feedback Oscillator Noise Spectrum.", Proc. I.E.E.E. 54, 329-330 (Feb. 1966)

2. Sauvage, G., "Phase Noise in Oscillators: A Mathematical Analysis of Leeson's Model", I.E.E.E. Trans. Instrum. Meas. IM-26(No. 4), 408-410 (Dec. 1977)

3. Parzen, B. "Design of Crystal and Other Harmonic Oscillators", Wiley, New York, 1983

4. Curtis, G.S., "The Relationship between Resonator Noise and Oscillator Noise and Resonator Noise Measurement Techniques", Proc. 41th Annual Frequency Control Symposium, pp. 420-428, May 1987

5. Robins, W.P. " Phase Noise in Signal Sources", Peter Peregrinus Ltd., London, 1982

6. Parzen, B. "Oscillator and Stability Analysis in the Immittance Domain", in preparation

42nd Annual Frequency Control Symposium - 1988

# PHASE NOISE IN DIRECT DIGITAL SYNTHESIZERS

Edward M. Mattison and Laurence M. Coyle
Smithsonian Astrophysical Observatory
Cambridge, Massachusetts, 02138

## Introduction

The receiver systems of most atomic clocks include a frequency synthesizer to adjust the frequency of a voltage controlled crystal oscillator that is phase locked to the atomic transition frequency. In the case of hydrogen maser receivers constructed by the Smithsonian Astrophysical Observatory (SAO), as with many other such systems, synthesizer phase variations contribute to the output of the voltage controlled crystal oscillator (VCXO) at the same level as phase noise from the maser signal; therefore a knowledge of the synthesizer phase noise is important for an understanding of the phase limitations of the instrument. We have designed and constructed a high resolution direct digital synthesizer (DDS) intended for hydrogen masers. We have investigated its phase noise by simulating its output signal and calculating its phase spectrum, and by measuring its phase spectrum directly. In addition we have compared its phase noise with that of other synthesizers used in hydrogen masers.

## Direct Digital Synthesizer Fundamentals

The basic components of the DDS[1,2] are shown in Fig. 1. At each cycle or pulse of a clock signal, a number that

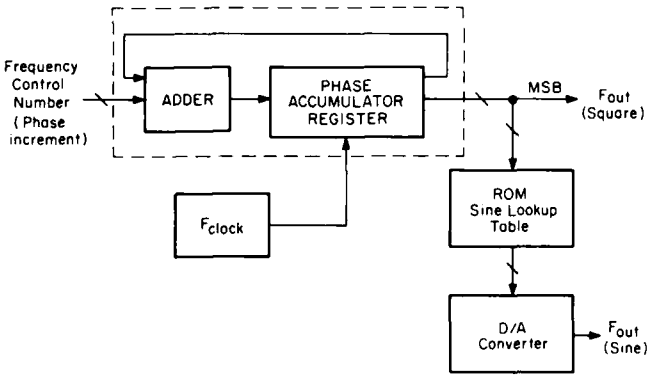



Figure 1. Block Diagram of Basic DDS Components

represents a phase increment is added to a digital accumulator whose content is proportional to the phase of the output signal. At each clock cycle the accumulator value is passed to a read-only sine lookup table (ROM) whose output goes to a digital-to-analog converter (DAC). The DAC produces a series of voltage steps that represent a sine wave sampled at the rate of the clock pulses. When the accumulator overflows as the result of the repeated additions, its most significant bit (MSB) changes from 1 to 0, and a new output cycle has begun. If only the MSB is observed, the DDS produces a square wave that changes sign coincidentally with the zero crossing of the sampled sine wave. Because the square wave can change sign only at a clock cycle, its phase is in general delayed by a varying amount relative to the ideal phase of the signal to be produced. The phase delay is proportional to the number remaining in the accumulator after an overflow.

If the accumulator capacity is L bits, and the phase increment number added at each clock cycle is P, the DDS's output frequency is given by

$$F_{out} = F_{clock}\frac{P}{2^L} \quad (1)$$

If $P$ is divisible by an integral power of 2, the ratio $P/2^L$ can be reduced to lowest terms, $M/2^N$, where the reduced phase increment, $M$, is odd. If, for example, $L$ = 8 bits and $P$ = 00111000 binary, $P$ and $2^L$ are each divisible by $2^3$ to give $M$ = 00111 (= 7 decimal) and $N$ = 5. Then Eq.1 can be written

$$F_{out}/F_{clock} = \frac{00111000}{100000000} = \frac{00111}{100000} = \frac{7}{32}(\text{decimal}) \quad (2)$$

On average, one output period for this example corresponds to 32/7 clock periods. Because in general a non-zero value remains in the accumulator after each overflow, the accumulator sequence is not the same for every output cycle; rather, the sequence repeats when the accumulator contains zero, which occurs after $M$ output cycles, or $2^N$ clock cycles. Thus the lowest Fourier frequency to be expected in the phase spectrum is

$$F_{\min} = \frac{F_{out}}{M} = \frac{F_{clock}}{2^N} \quad (3)$$

In the example above, the repetition time for the accumulator sequence, which represents the longest variation period of the output waveform, corresponds to 7 output cycles, or 32 clock cycles.

## SAO Direct Digital Synthesizer

The SAO maser synthesizer is part of a receiver system that phase locks a 100 MHz VCXO to the 1420.4 MHz hydrogen maser signal. Using a 5 MHz clock signal derived from the VCXO, the synthesizer produces an output signal of approximately 405 kHz. The DDS is controlled by a Z-80 microprocessor that is the central component of a digital control and monitoring system. This system permits local or remote control of the maser's tuning diode voltage, hydrogen pressure, and synthesizer frequency. In addition, it samples and stores the maser's analog operating parameters at selectable intervals, and transfers the data to an external computer upon request. The synthesizer, which is constructed on a single STD bus wire-wrap card (Fig. 2), is based upon a pair of 32-bit number controlled oscillators[3]. It has an accumulator length of 64 bits, and thus a maximum resolution of $2.7 \times 10^{-13}$ Hz, corresponding to a fractional VCXO frequency variation of $1.9 \times 10^{-22}$. While the synthesizer is capable of producing frequencies up to approximately 40% of the clock frequency, the microprocessor software has been written to allow synthesizer frequencies in the range from 405750 Hz to 405760 Hz (the frequency interval relevant for the maser) with a resolution of $10^{-6}$ Hz, corresponding to a VCXO fractional frequency variation of $7 \times 10^{-18}$.

## Simulation of DDS Phase Variation

Because the DDS can make output transitions only coincidentally with clock pulses, the zero crossings of its output do not in general coincide with the theoretical zero crossings of the ideal waveform to be approximated. (Only when the phase increment P is an integral power of 2 is there no accumulator overflow, and the output is exactly equivalent to the ideal waveform.) The accumulator remainder following each overflow is proportional to the phase delay between the ideal

CH2588-2/88/0000-352 $1.00 C 1988 IEEE

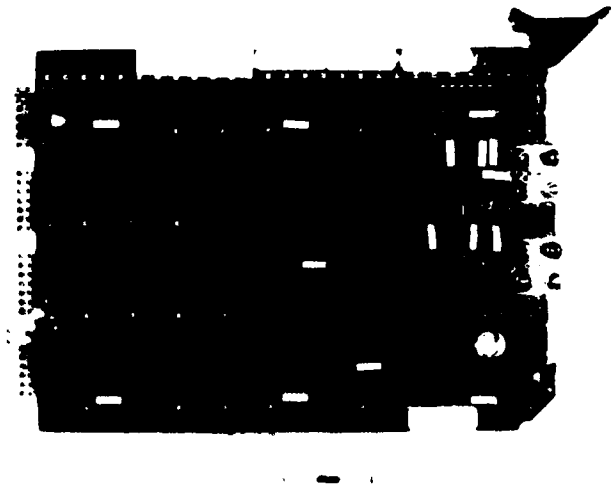


Fig. 2. SAO 64-bit Direct Digital Frequency Synthesizer

waveform and the actual zero crossing. The series of remainders, then, represents the phase variation of the square wave DDS output, sampled once each output cycle. We simulated the DDS phase spectrum by calculating the remainder series for particular phase increment numbers, and then calculating the discrete Fourier spectrum of the series by means of a fast Fourier transform software package.

We simulated the phase spectrum for two ranges of phase increment numbers: "simple" numbers with few non-zero bits, all within the eight most-significant bits; and the phase increment that produces the frequency 405750 Hz for a clock frequency of 5 MHz. Simple numbers produce phase spectra with a small number of well separated lines, making comparison with measured spectra relatively straightforward, while 405750 kHz is within the frequency range used in hydrogen masers. The simulated phase spectra for simple phase increment numbers display $M$ equally spaced spectral lines, separated by $(F_{out}/M)$ Hz, for Fourier frequencies f in the range $0 \le f < F_{out}$.

It has been pointed out[2] that a source of DDS phase variation is the phenomenon of "phase truncation". Although DDS accumulators generally are designed to store a large number of bits, hardware limitations make it practical to pass only a small number of the most significant bits to the sine look-up ROM. (In the STI 1172A, 8 bits are passed to the ROM.) If the number of bits in the reduced phase increment $M$ is larger than the number of bits sent to the ROM, the consequent truncation of the accumulator contents produces an error in the output phase. For the "simple" values of $M$ considered here, which contain at most 8 bits, phase truncation does not occur, and thus does not affect the correspondence between the simulated and measured spectra.

## Measurement of DDS Phase Spectra

The DDS phase spectra were measured by comparing the outputs of two synthesizers operated from a common clock reference. The signals were combined in a double-balanced mixer whose output was amplified and filtered, and the resulting phase spectrum was measured with a digital spectrum analyzer[4]. The mixer's output is proportional to the phase variation of the input signals when the average phase difference between the inputs is 90°.

Because the DDS signal is produced by a deterministic sequence of accumulator values beginning with zero, one expects the observed relative phase variation between two DDS's to depend upon the delay between the start of one DDS sequence and the start of the second. (If the DDS's are started simultaneously, no relative phase variation should be observed.) Using a circuit that reset the accumulators of two DDS's with a selectable time delay, we observed variations in the measured phase spectrum with initial delay. To eliminate this source of variability, we measured the phase spectra of a single DDS relative to a sine-wave analog frequency synthesizer (HP5100A), whose output is uniform in time aside from random phase noise. For Fourier frequencies less than 100 Hz the DDS phase noise is considerably lower than that of the HP5100A, and in that frequency range we measured the phase spectral densities of two DDS's relative to one another.

The DDS synthesizers used for these measurements were capable of producing either a square wave output from the accumulator MSB, or a sine wave from the DAC. The controlling software was modified to allow any frequency to be selected.

## Results

Simulated and measured phase spectra for phase increment $M$ = 00111 (=7 decimal) are shown in Fig. 3. The simulated square-wave spectrum, shown by triangles, was calculated for Fourier frequencies up to the Nyquist limit, or half the sampling frequency. (The sampling frequency equals $F_{out}$ since samples are calculated once each output cycle.) The open triangles represent the calculated lines reflected about the Nyquist frequency. For the measured spectra $F_{clock}$ = 400 kHz,

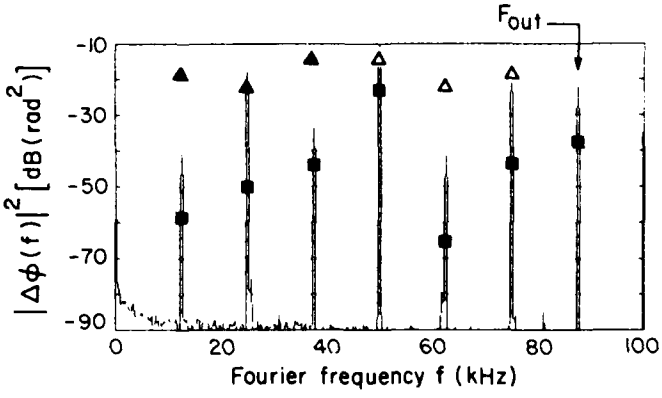



Fig. 3. DDS Phase Spectra for $M$ = 00111

giving $F_{out} = (7/32)F_{clock}$ = 87.5 kHz. The lines in the measured spectra correspond in number — 7 — and position to those predicted by the simulation. The magnitudes of the even lines in the measured square wave spectrum (2, 4, 6...$\times F_{min}$) agree within a few dB with those of the simulated spectrum, while the odd spectral lines (1, 3, 5...$\times F_{min}$) are lower than the predicted levels by as much as 30 dB. The levels of the sine wave phase spectrum are considerably below both the simulated and the measured square wave spectra.

Spectra for $M$ = 000111001 (=57 decimal) are shown in Figs. 4 and 5. Here $F_{clock}$ = 400 kHz and $F_{out} = (57/256)F_{clock}$ = 89062.5 Hz. The square wave spectrum (Fig. 4) clearly shows the predicted 57 lines between 0 Hz and $F_{out}$, and suggests the rapid increase in spectral complexity with increasing length of the phase increment. The dots represent the simulated spectral line strengths. This spectrum also has

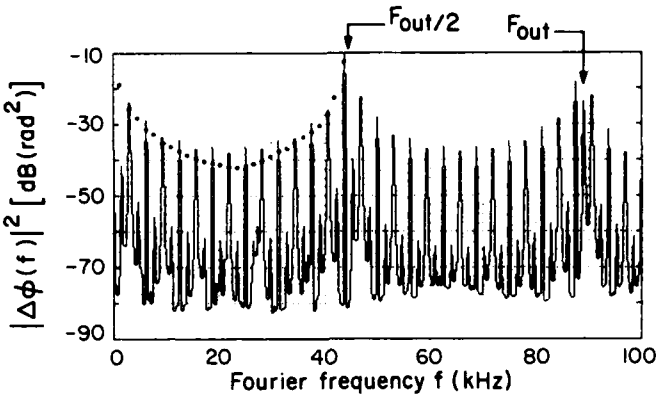



Fig. 4. DDS Square Wave Phase Spectra for $M$ = 00111001

even spectral components that agree in magnitude with the simulated spectrum, and odd components that are considerably lower. The sine wave phase spectrum (Fig. 5) is substantially lower in magnitude than the square wave spectrum.

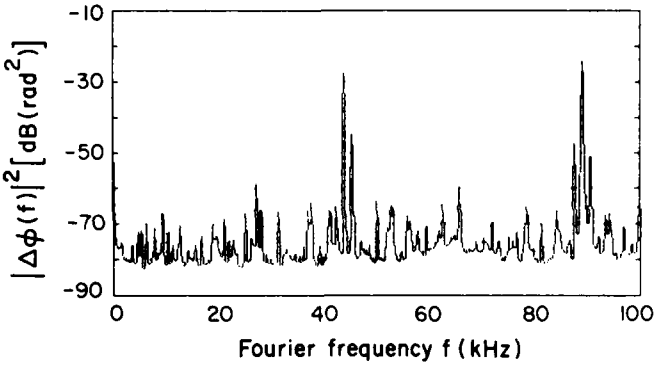



Fig. 5. Measured DDS Sine Wave Phase Spectrum for $M$ = 00111001

The improvement of the sine wave phase spectrum over the square wave spectrum can be understood by the following heuristic argument. Because the sine wave DAC changes its output coincidentally with clock pulses, and thus theoretically makes zero crossings at the same instant as the square wave output, one expects the phase variations of the sine wave to be identical to those of the square wave. However, the DAC circuitry has associated capacitances that cause slew-rate limiting, which is equivalent to low-pass filtering. Fig. 6 shows the effect of a low pass filter (LPF) on the DAC output. In both 6a and 6b the theoretical DAC zero transition, shown by a dashed line, coincide with the square wave zero crossing. However, the LPF causes the DAC output to decay exponentially from the previous level to the new one. In 6a the phase delay $\Delta\phi$ between the theoretical sine wave and the DAC transition is small, the DAC voltage goes from a relatively large positive value to a small negative value, and the exponential decay delays the actual zero crossing for a relatively long time, increasing the phase error by a large fraction. In 6b, by contrast, $\Delta\phi$ is large, the DAC voltage goes from a small positive to a large negative value, and the exponential decay delays the zero crossing by a relatively short time, making a small increase in the phase error. Therefore the net phase errors – theoretical phase error plus LPF delay – for the two cases tend toward equality, reducing the net phase variation from one cycle to the next.

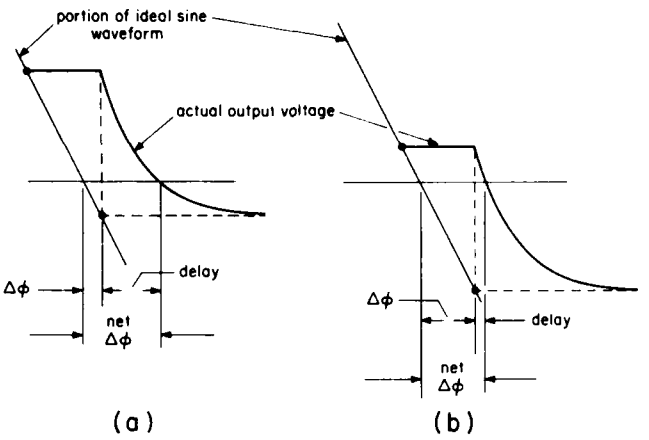



Fig. 6. Effect of Filtering on DDS Stepped Sine Output

The measured and simulated square wave phase spectra for $F_{out}$ = 405750 kHz are shown in Fig. 7. (For the measurement, values of $F_{clock}$ = 1 MHz and $F_{out}$ = 81150 Hz were used, in order to be able to display the spectum up to $F_{out}$ on the 100 kHz spectrum analyzer.) The number and positions of major lines in the simulated and measured spectra agree, although the measured spectrum has lower resolution than the simulation. For convenience, the major lines of the simulated spectrum are indicated on the measured spectrum by dots.

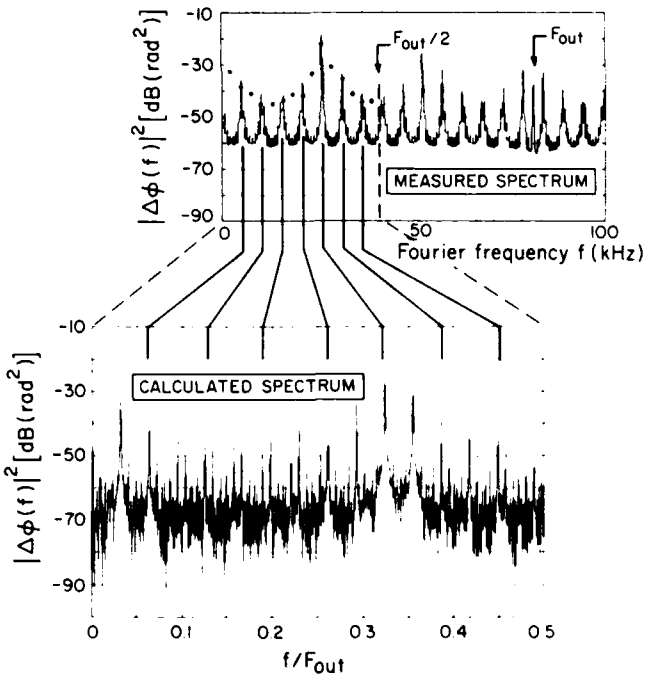



Fig. 7. Simulated and Measured DDS Square Wave Spectra $F_{out}/F_{clock}$ = 405750/5 000 000

(The phase increment for this output frequency has a 1 in its 64th, or least signficant, bit. For a 5 MHz clock clock signal, the longest repetition time of the accumulator sequence is $2^{64}$ clock cycles, or $1.2\times10^5$ years. It is clearly impossible to integrate a measurement long enough to observe this period, or to simulate the spectrum with enough accumulator values to resolve the minimum line separation.) As with the other measurements, the odd major components of the measured spectrum are suppressed relative to the predicted levels.

For evaluating synthesizer operation in the maser, Fourier frequencies lower than roughly 100 Hz are of interest. Because the lock-loop time constant for the maser receiver system is on the order of 10 Hz, higher frequency synthesizer phase variations are attenuated by the phaselock LPF. Fig. 8 shows the phase spectral density of two sine-wave DDS's for Fourier frequencies below 100 Hz, for $F_{clock}$ = 5 MHz and $F_{out}$ = 405750 Hz. (Because the synthesizers are equivalent, the spectral density of a single synthesizer is 3 dB lower than the levels shown.) Because the low-frequency spectrum is free of discrete lines, the spectral density (with units of dB relative to 1 radian$^2$/Hz) is the appropriate measure of phase noise. The spectral density at these Fourier frequencies was not noticeably affected by the starting delay between the two synthesizers.

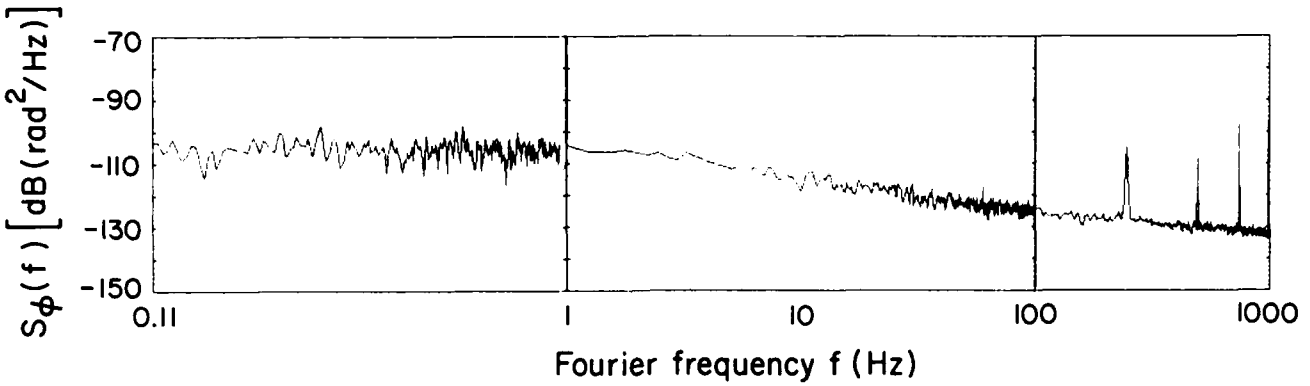



Fig. 8. Phase Spectral Density of DDS Sine Wave
$F_{clock}$ = 5 MHz; $F_{out}$ = 405750 kHz

In order to establish a basis for comparison of the DDS phase performance, we measured the phase spectra of two other types of frequency synthesizers used in hydrogen masers, a phase-lock loop (PLL) synthesizer made by SAO, and a Dana Laboratories Corp. model 7010 synthesizer. The results, shown in Table 1, indicate that the phase noise of the DDS synthesizer is less than or at most equal to the other synthesizers for Fourier frequencies below 1 kHz.

Table 1
Comparison of Phase Noise Among Synthesizers
$F_{clock}$ = 5 MHz; $F_{out}$ = 405750 Hz
Units = dB(rad$^2$/Hz)

| Freq (Hz) | 0.1 | 1.0 | 10 | 100 | 1000 |
|---|---|---|---|---|---|
| Oscillators: | | | | | |
| DDS Sq/Sq | — | −103 | −113 | −120 | −127 |
| DDS Sin/Sin | −105 | −104 | −115 | −124 | −132 |
| PLL | — | −95 | −112 to −95 | −120 to −110 | −135 to −120 |
| Dana | −75 | −90 | −110 | −118 | −115 |

## Summary and Discussion

With increasing use of digital frequency synthesizers in electronic systems, an understanding of their phase noise performance is increasingly important. We have demonstrated the ability to calculate accurately the locations of phase spectral lines, and have compared simulated spectra with measured values. The measured line magnitudes agree with the calculated values for the square wave spectra, with the exception of the suppression of alternate (odd) lines. Because we observed this suppression with an analog (swept-filter) spectrum analyzer as well as with the digital analyzer, we conclude that it is not related to spectrum analyzer performance. We are currently investigating the phenomenon by means of high-resolution calculations of the process of multiplying a DDS square wave with a square wave of uniform period.

Our spectral measurements clearly show that the use of a sine-wave output with low pass filtering improves the DDS phase spectrum. It would be valuable to investigate quantitatively the effects of LPF cutoff frequency and slope on the phase spectrum.

It has been shown[2] that the power spectrum of the DDS can be very sensitive to specific values of the phase increment number. A similar sensitivity is to be expected for the phase spectrum. In any particular application, it is important to investigate the phase behavior of a DDS in the frequency range of interest. A large accumulator capacity makes possible an extremely large number of discrete output frequencies within any macroscope frequency range; thus it is a practical impossibility to examine the phase spectrum for every possible frequency. However, useful conclusions may be drawn by studying the trend of the phase spectrum with frequency and by identifying frequency ranges associated with particularly large values of phase variation. Further work along these lines is in process at SAO.

## Acknowledgements

We are pleased acknowledge the support of the Jet Propulsion Laboratory for providing the impetus and funding for this work. For interesting and helpful discussions we are pleased to thank Leonard Cutler, Paul Kuhnle, Henry Nicholas, III, Donald Percival, Richard Sydnor, Robert Vessot, and Fred L. Walls.

## References

1. V. Reinhardt, K. Gould, K. McNab, and M. Bustamante, "A short survey of frequency synthesizer techniques." Proc. 40th Annual Frequency Control Symp., 1986, pp. 355-365.

2. Henry T. Nicholas, III, and Henry Samueli, "An analysis of the output spectrum of direct digital frequency synthesizers in the presence of phase-accumulator truncation." Proc. 41st Annual Frequency Control Symp., 1987, pp. 495-502.

3. Model 1172A number controlled oscillator, Stanford Telecommunications, Inc., Santa Clara, CA 95054

4. Mixing and amplification were performed by a Femtosecond Systems, Inc., model 600 phase noise detection unit. The spectrum analyzer was an HP model 3561A.

42nd Annual Frequency Control Symposium - 1988

# The Optimization of Direct Digital Frequency Synthesizer Performance in the Presence of Finite Word Length Effects

**Henry T. Nicholas, III**
TRW Digital Applications Laboratory
One Space Park, M5/1017
Redondo Beach, CA 90278

**Henry Samueli**
Electrical Engineering Department
University of California, Los Angeles
Los Angeles, CA 90024

**Bruce Kim**
Intel Corp.

## Abstract

Techniques for optimizing the performance of Direct Digital Frequency Synthesizers (DDFS) that use the Tierney, Rader, Gold architecture [1] are presented. These techniques permit the optimal partitioning of DDFS hardware to minimize the effects of finite phase resolution and finite output word-length. Using these results, given constraints on the phase resolution and output word-length, DDFS performance can be optimized under either a worst case spurious response criterion or under a total signal to noise ratio criterion. These optimization techniques are also used to analyze the errors introduced by the compression of the required look-up table storage size. Two look-up table compression techniques are examined: the Sunderland technique [3][4], and a proposed modification of the Sunderland technique. It is shown that by using a different optimization technique to choose the look-up table samples, the algorithmic error incurred by the Sunderland architecture can be improved by approximately 12 dB. Furthermore, a new compression technique is proposed, based upon an extension of the Sunderland technique, which provides additional improvements in the look-up table storage compression. Finally an actual DDFS design is presented that utilizes this new compression technique and which provides a simulated digital spurious rejection of 90.2 dB, a 14 bit digital output, and a frequency resolution of .023 Hz at the maximum simulated clock frequency of 100 Mhz. This design, which will be fabricated in a radiation hardened 1.25μM CMOS process is presently in the layout phase of design and will enter fabrication in November '88.

## Introduction

Direct Digital Frequency Synthesis is an indispensable technique for generating reference frequencies whenever extremely precise frequency resolution and fast switching speed are required. The most common DDFS architecture is the Tierney, Rader, Gold, architecture [1]. This architecture, which has been well documented in the literature [5][6][7], synthesizes a sine wave by using a periodically overflowing 2's complement phase accumulator to generate and store phase information, and uses a Read-Only Memory (ROM) based look-up table to compute the sine function. There are three sources of noise which are inherent to all DDFS implementations, in addition to the noise generated in the D/A conversion process. The first source of noise is P(n), the distortion due to phase truncation at the input to the sine function computing hardware, which is usually a ROM. The second is g(•), which is a nonlinear distortion that is usually present when methods of compressing the storage requirements of the look-up table are employed, and the third is A(n), the noise introduced by the finite precision of the sine samples stored in the look-up table. These noise sources are depicted symbolically in Figure 1. This paper will be concerned with the characterization of these sources and their relation to the hardware requirements.

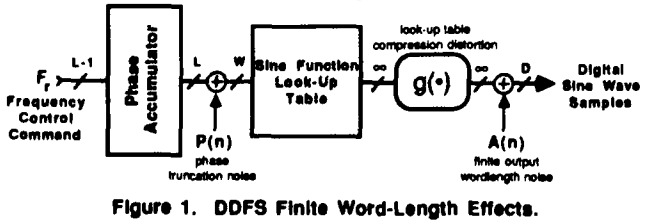



**Figure 1. DDFS Finite Word-Length Effects.**

## Effects of Finite Phase Word-Length

There are two finite word-length effects in the phase accumulator that effect DDFS performance: the word-length of the phase accumulator register, denoted as L in Figure 1, and the truncation of the phase accumulator output, which is used to index the sine function look-up table, which is denoted by B. An understanding of these two finite word-length effects is extremely important because they have a profound impact upon the spectral characteristics of all other system nonlinearities. An exact spectral analysis of DDFS spurious performance was presented in [2], which provided an algorithm for determining the magnitude and position of all spurs generated by phase accumulator truncation. This algorithm is restated in Figure 2 to emphasize some of the subtleties of phase accumulator operation that impact the overall DDFS spurious spectrum.

Number of spurs = $2\Lambda - 1$

Magnitude of all spurs $\zeta_K = \dfrac{\pi\, 2^{-L}\,(F_r,2^B)}{\sin\left(\dfrac{\pi\, K\,(F_r,2^B)}{2^B}\right)}$

Position of the spurs in the spectrum (between 0 and $\dfrac{2^L}{(F_r,2^L)}$)

$K = \left\langle \dfrac{F_n - \Gamma}{2^{L-B}}\,\Gamma\Lambda^{-1} \right\rangle_{2\Lambda}$ for $F_n - \Lambda$ divisible by 2

$K = \left\langle \dfrac{-F_n - \Gamma}{2^{L-B}}\,\Gamma\Lambda^{-1} \right\rangle_{2\Lambda}$ for $F_n - \Lambda$ divisible by 2

$\zeta_K = 0$ otherwise

B = number of phase bits truncated

L = phase accumulator word length

(a,b) = greatest common divisor of a and b

$\Gamma = \dfrac{F_r}{(F_r,2^B)}$ $\quad \Lambda = \dfrac{2^B-1}{(F_r,2^B)}$

**Figure 2. Exact Spectrum of Phase Truncation Spurs**

This algorithm provides an expression for the number of spurs generated, their amplitude, and their position in the spectrum as a function of B, the number of bits truncated from the phase accumulator, L, the phase accumulator word-length, and the frequency control word, $F_r$. These results illuminate some important properties of phase accumulator operation that effect the overall DDFS error spectrum. The most important property to note is that the number of spurs and the magnitude of the spurs depend on $F_r$ only through the greatest common divisor function of $F_r$ and $2^B$, which is denoted by $(F_r,2^B)$. As a result of this, for input values of $F_r$ that have the same value of $(F_r,2^B)$, the number of spurs and their respective amplitudes remain the same. Only the position of each spur in the spectrum changes. This phenomenon is a result of certain number theoretic properties of the phase accumulator. These properties are most easily demonstrated if we view the sequence of numbers formed by one numerical period of the phase accumulator as a single time series vector, $\mathbf{X}$. In this representation the length of one numerical period of the phase accumulator is defined as the length of the sequence of numbers that may be output by the phase accumulator before the first number is repeated. As was shown in [2], the length of this

CH2588-2/88/0000-357 $1.00 C 1988 IEEE

sequence is $2^L/(F_r,2^L)$. We then define ${}_1\mathbf{X}$ as the vector of length $2^L/(F_r,2^L)$, which is formed by the consecutive output samples of the phase accumulator when the input is $F_r = 1$. The elements of this vector are then defined using the notation:

$$ {}_1\mathbf{X} = [\ {}_1X_0\ ,\ {}_1X_1\ ,\ {}_1X_2\ ,\ {}_1X_3\ ,\ \bullet\bullet\bullet\ ,\ {}_1X_n\ ]^T \qquad (1) $$

This vector is printed as the transpose of a row vector for typographical convenience. Using this notation, the vector ${}_{F_r}\mathbf{X}$ can be formed by using:

$$ {}_1X_n = \langle F_r n \rangle_{2^L} \qquad (2) $$

to denote the value of each element of the vector from n = 0 to $2^L/(F_r,2^L)$ - 1. To illustrate this concept, the example of a 3 bit accumulator is shown in figure 3. For this example, if $F_r$ is odd then $(F_r,2^L)$ will be 1, thereby generating a time vector of length 8. The column vector on the left represents the time vector that results when $F_r$=1. For this simple example, the 3 bit accumulator starting with an initial phase accumulator register value of zero, simply adds "1" to the register contents every clock cycle until the register overflows, and returns to zero. This result agrees with equation (2) with the elements ${}_1X_0 = 0$, ${}_1X_1 = 1$, ${}_1X_2 = 2$, ect. The column vector to the right in figure 3 represents the time vector formed when $F_r = 3$. Note that although the phase accumulator accumulates and overflows three times faster for $F_r = 3$ than for $F_r = 1$, the numeric period of the two output sequences are the same.

An extremely important "side effect" of having the same numerical periodicity is that the column vector for $F_r = 3$ can be formed from a permutation of the values of the $F_r = 1$ vector. This permutation is represented by the arrows in Figure 3. It can be shown from the definition of the time vector in equation (2) that any phase accumulator output vector, ${}_a\mathbf{X}$, can be formed from the permutation of another output vector, ${}_b\mathbf{X}$, if $(a,2^L) = (b,2^L)$. The set of permutations that relate these output vectors belongs to a very special class of number theoretic relationships called affine permutations. Equation (2) thus defines the set of affine permutations that relate all time output vectors to the basic output vector, ${}_1\mathbf{X}$, when $(F_r,2^L) = 1$.

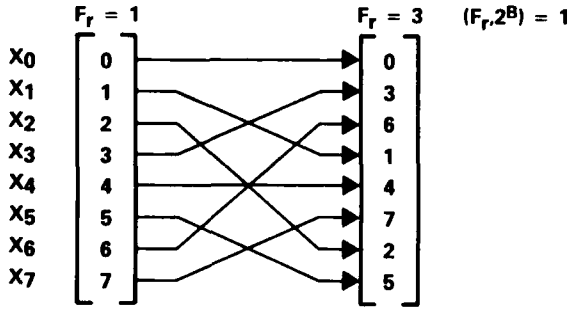



**Figure 3. Time Series Vectors for a 3-Bit Accumulator for $F_r$ = 1 and $F_r$ = 3.**

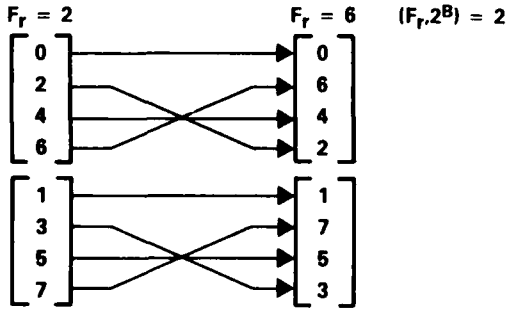



**Figure 4. Time Series Vectors for a 3-Bit Accumulator for $F_r$ = 2 and $F_r$ = 6.**

In Figure 4 we see the time vectors formed from values of $F_r = 2$ and $F_r = 6$. From the figure it is evident that the phase accumulator is now characterized by having two different sets of possible output vectors depending upon the initial contents of the phase accumulator. The top set of vectors in the figure represent the sequence of numbers output from the accumulator with an initial accumulator contents of 0. The bottom set of vectors represent the case for an initial accumulator contents of 1. In general, the phase accumulator will be characterized as having $(F_r,2^L)$ different "families" of output sequences to characterize its response, where output sequences belonging to the same "family" are defined as those having the same initial phase accumulator contents. As shown in Figure 4, time vectors belonging to the same family are related by a set of affine permutations of the same form as defined in equation (2).

There are several fundamental implications of these number theoretic properties of the phase accumulator. First, the worst case spurious response of the DDFS as a result of all memoryless system error effects will have exactly $(F_r,2^L)$ different values, one for each initial phase accumulator content from 0 to $(F_r,2^L)$ - 1. This means that simulations of the spectrum of DDFS nonlinearities are only valid for initial phase accumulator contents belonging to the same time series vector.

Another important implication is associated with the affine permutation relationship that links time output vectors of the same family. In equation (3) we see that the time output vector for any value of $F_r$ can be formed from a permutation of the individual elements of the vector for $F_r = 1$ by permuting the vector indices using $i' = \langle F_r i \rangle_{2^L}$.

$$ {}_1\mathbf{X} = [\ {}_1X_0\ ,\ {}_1X_1\ ,\ {}_1X_2\ ,\ {}_1X_3\ ,\ \bullet\bullet\bullet\ ,\ {}_1X_n\ ]^T \quad \text{vector } F_r = 1 \qquad (3) $$

$$ {}_{F_r}\mathbf{X} = [\ {}_1X_{\langle F_r 0\rangle_{2^L}}\ ,\ {}_1X_{\langle F_r 1\rangle_{2^L}}\ ,\ {}_1X_{\langle F_r 2\rangle_{2^L}}\ ,\ \bullet\bullet\bullet\ ,\ {}_1X_{\langle F_r n\rangle_{2^L}}\ ]^T \quad \text{vector for arbitrary } F_r $$

It can be shown that the discrete Fourier transform operator is invariant under permutations of this class applied in both the time and frequency domain. If we assume that in general the DDFS operates by applying some memoryless nonlinear function $S\{\bullet\}$ to the phase accumulator output to produce the sine function, then the DDFS output at every point in time for a given $F_r$ may be represented as function of one of the indices of ${}_{F_r}\mathbf{X}$, $S\{{}_{F_r}X_n\}$. Therefore the spectrum of the DDFS output may be represented as:

$$ N = \frac{2^L}{(F_r,2^L)} \qquad \mathbb{S}\{K\} = \sum_{n=0}^{N-1} S\{{}_{F_r}X_n\}\, e^{-j2\pi Kn/N} \qquad (4) $$

As was shown in equation (3), all input time vectors may be formed from a permutation of another time vector by permuting the indices. If this is applied to the definition of the DDFS output spectrum in (), it can be shown that the following identity holds for affine permutations of the class $\langle F_r i \rangle_{2^L}$:

$$ \mathbb{S}\{\langle F_r K\rangle_{2^L}\} = \sum_{n=0}^{N-1} S\{{}_{F_r}X_{\langle F_r n\rangle_{2^L}}\}\, e^{-j2\pi Kn/N} \qquad (5) $$

Therefore a permutation of the samples in the time domain results in an identical permutation of the DFT samples in the frequency domain. In practical terms this means that the spurious spectrum due to all system nonlinearities can be generated from a permutation of another spectrum of the same number theoretic class. Thus, if we restrict the allowable values of $F_r$ such that $(F_r,2^L)$ is always the same (such as by only allowing $F_r$ to be odd)

then all spurious spectrums may be generated from permutations of a single basic spectrum. If it is desired to optimize the worst case spurious response, then only one spectrum need be calculated for all values of $F_r$ with common values of $(F_r, 2^L)$, since each spectrum will differ only in the position of the spurs and not in the magnitudes.

A simple modification to the basic phase accumulator structure was introduced in [2] that causes all phase accumulator output sequence to belong to the number theoretic class $(F_r, 2^{L+1}) = 1$, regardless of the value of $F_r$. By using this phase accumulator modification, which is depicted in Figure 5, only one simulation need be performed to determine the value of the worst case spurious response or the total signal to noise ratio due to memoryless system nonlinearities. In addition, this modification has the added advantages of eliminating any dependence of the output spectrum on the initial phase accumulator contents, and of reducing the magnitude of the worst case spurious response by 3.92 dB.

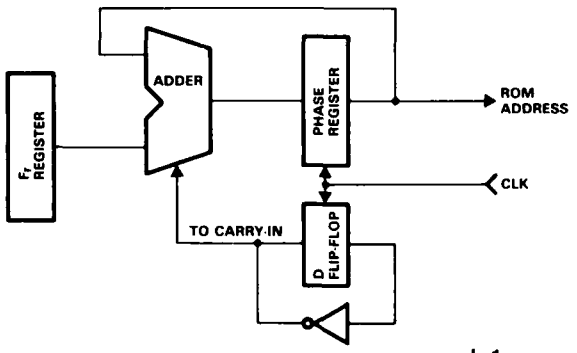



*Figure 5. Modified Phase Accumulator For $(F_r, 2^{L+1}) = 1$.*

## Finite Amplitude Word-Length Effects

The results relating the number theoretic behavior of the phase accumulator finite word-length effects to the DDFS output spectrum may now be applied to help bring about the optimization of the sine sample computing hardware. Two error effects are introduced in computing the sine samples from the phase argument. These are the error from the storage compression nonlinearity, and the error from the amplitude quantization of the sine samples, which are depicted in Figure 1. In the past these error effects have been treated as random phenomena. However, using the results from the previous section, it is now computationally feasible to treat them as a deterministic process.

## Look-up Table Storage Compression

### Exploitation of Sine Function Symmetry

The most elementary method of sine storage compression is to exploit the symmetry of the sine function about $\pi$ and $\pi/2$. By properly inverting the phase and the amplitude of the sine function, look-up table samples need only be stored for phase values between 0 and $\pi/2$. The details of this method are shown in Figure 6. As shown in the figure, the sine function between 0 and $\pi$ may be synthesized from the samples between 0 and $\pi/2$ by taking the phase modulo $\pi/2$ and then taking the absolute value of the phase. This is easily implemented in hardware by truncating the phase MSB and then using the second MSB to full wave rectify the magnitude of the phase. As shown in Figure 6, the sampled waveform at the output of the look-up table is a full wave rectified version of the desired sine wave. The final output sine wave is then generated by multiplying the full wave rectified version by -1 when the phase is between $\pi$ and $2\pi$. This is accomplished simply by multiplying by the negative of the phase accumulator MSB.

In most practical DDFS digital implementations, numbers are represented in 2's complement format. Therefore the 2's complementor must be used to take the absolute value of the phase and multiply the output of the look-up table by -1. However, it can be shown that if a 1/2 LSB offset is introduced into the number that is to be complemented, then a 1's complementor may be used in place of the 2's complementor without introducing an error. This provides a considerable savings in hardware since a 1's complementor may be implemented as a simple exclusive-or gate. This 1/2 LSB offset is provided by choosing the look-up table samples such that there is a 1/2 LSB offset in both the phase and the amplitude of the samples.

### Compression of Sin(θ) for $0 < \theta < \pi/2$

Prior to the publication of this paper, the most effective method of compressing the look-up table storage requirements for $0 < \theta < \pi/2$ was introduced by Sunderland et. al. [3][4]. This architecture, which is shown in Figure 7, provides a reduction in the look-up table storage requirements by replacing the storage requirements of one large ROM of size $2^{A+B+C}$ words with two smaller ROM's of sizes $2^{A+B}$ words and $2^{A+C}$ words, whose outputs are added together to reconstruct the sine function. This provides a storage compression ratio of 45,056 bits to 3,840 bits, or 11.7:1. The assignment of look-up table samples in this architecture is based upon several trigonometric approximations. First, the bits representing the phase argument, θ, of one quarter period of the sine function are decomposed into the sum of three functions: $\alpha < (\pi/2)$, $\beta < (\pi/2)(2^{-A})$, and $\chi < (\pi/2)(2^{-(A+B)})$ such that

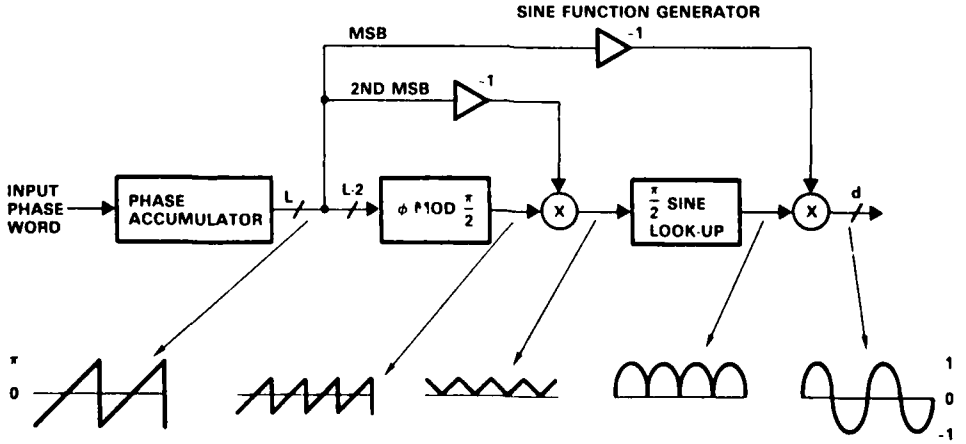



**Figure 6. Logic to Exploit Quarter Wave Symmetry.**

$\theta = \alpha + \beta + \chi$ [3]. Using trigonometric identities, the sine function is given by

$$\sin(\alpha+\beta+\chi) = \sin(\alpha+\beta)\cos(\chi) + \cos(\alpha)\cos(\beta)\sin(\chi) - \sin(\alpha)\sin(\beta)\sin(\chi) \quad (6)$$

using the relative magnitudes of $\alpha$, $\beta$, and $\chi$, (6) may be approximated by

$$\sin(\alpha+\beta+\chi) \approx \sin(\alpha+\beta) + \cos(\alpha)\sin(\chi) \quad (7)$$

The contents of the first ROM is then chosen to be $\sin(\alpha+\beta)$ and the second ROM stores the result of $\cos(\alpha)\sin(\chi)$. The error resulting from this approximation for a 12 bit output with a phase word segmentation of A = 4, B = 4, and C = 4, which will be denoted as a (4,4,4) segmentation, can be shown to be approximately -72.2 dB.

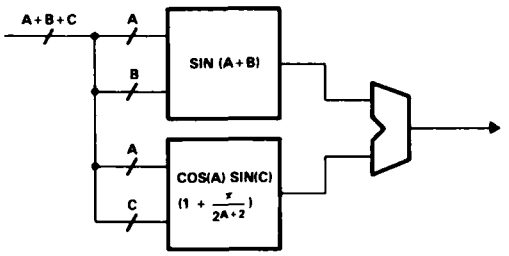



**Figure 7. Sunderland Architecture.**

However, significant improvements can be achieved, over the former results by using an alternate methodology for choosing the samples to be stored in the ROM's. Improvements in the spurious performance can be gained by modeling the segmentation of the phase arguments into the two ROM's as a low phase resolution course ROM, with additional phase resolution added by providing interpolation between the course phase samples using the fine ROM. We define as before, $\theta = \alpha + \beta + \chi$, with the word-length of the variable $\alpha$ to be A, the word-length of $\beta$ to be B, and $\chi$ to be C. The sine function between 0 and $\pi/2$ is divided into $2^B$ different regions. In each region the same $2^B$ fine ROM interpolation samples are used for interpolation between course ROM samples given the same value for the variable $\chi$. An example of this type of coarse-fine interpolation is shown in Figure 8 for the case of a 6 bit phase argument using a segmentation of A = 2, B = 2, and C = 2.

In this example, the course ROM samples are represented by the "heavy dots" along the dashed line, and the fine ROM samples are chosen to be the difference between the value of the sine function along the dashed line and the value value of the "error bars" directly below or above that point on the dashed line. In the example, the sine function is divided into $2^\alpha$ quadrants, correspond to $\alpha$ = 00, 01, 10, and 11. Within each quadrant, only one correction value may be used between the error bars and the dashed line for all $\chi$ offsets from a coarse sample. The value of the offset used for each value of $\chi$ is chosen to minimize either the mean square or the maximum error of all offsets for each value of $\beta$ within the quadrant. The actual algorithm for choosing the fine ROM samples to minimize the mean square error is shown in (8).

$$F_c(\alpha,\beta) = \sin\left(\frac{\pi}{2}\left(\frac{\alpha 2^{B}+\beta}{2^{A+B}} + \frac{1}{2^{A+B+C}}\right)\right) \quad (8)$$

$$F_f(\alpha,\chi) = \sum_{n=0}^{N-1} \frac{1}{2^B}\left[\sin\left(\frac{\pi}{2}\left(\frac{\alpha 2^{B+C}+\beta 2^{C}+\chi}{2^{A+B+C}} + \frac{1}{2^{A+B+C+1}}\right)\right) - F_c(\alpha,\beta)\right]$$

Where in (8) $F_c(\alpha,\beta)$ is defined to be the coarse ROM samples and $F_f(\alpha,\chi)$ is defined to be the fine ROM samples. The optimization criterion for choosing the samples to minimize the maximum absolute error is shown in (9).

$$F_f(\alpha,\chi) = \quad (9)$$
$$\frac{1}{2}\mathrm{MAX}\left\{\sin\left(\frac{\pi}{2}\left(\frac{\alpha 2^{B+C}+\beta 2^{C}+\chi}{2^{A+B+C}} + \frac{1}{2^{A+B+C+1}}\right)\right) - F_c(\alpha,\beta)\right\}$$
$$- \frac{1}{2}\mathrm{MIN}\left\{\sin\left(\frac{\pi}{2}\left(\frac{\alpha 2^{B+C}+\beta 2^{C}+\chi}{2^{A+B+C}} + \frac{1}{2^{A+B+C+1}}\right)\right) - F_c(\alpha,\beta)\right\}$$

Based upon simulation results, it was determined that the minimum-maximum error criterion tended to result in lower values of the maximum spur, although this result has not been explicitly proven. The simulations also showed that the mean square error criterion, provides the lowest total spur energy, in agreement with theory.

This methodology of choosing the coarse and fine ROM samples provides two advantages over previous methods. First, it provides superior spurious performance because the fine ROM samples may be chosen by computer optimization rather than by linear interpolation. Secondly, because this method of storage compression is not derived from trigonometric identities, it may be used to compress the storage of any arbitrary function, not just the sine function.

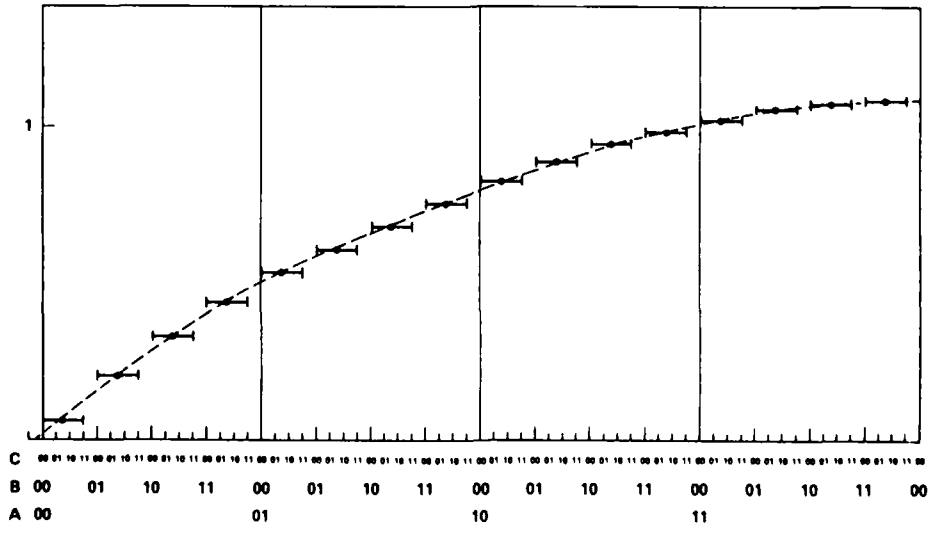



**Figure 8. Coarse-Fine Interpolation for (2,2,2) Segmentation of the Phase.**

A comparison was performed between the two different methodologies of choosing the ROM samples for the case of a hypothetical design with a phase precision of 12 bits. The segmentation was (4,4,4) for both cases. In the trigonometric derivation, the coarse ROM size was 256 x 11 bits and the fine ROM size was 256 x 4 bits. For the coarse/fine derivation, the coarse ROM size was 256 x 9 bits and the fine ROM size was 256 x 4 bits. The result of the simulations showed a worst case spurious performance of -72.2 dB for the trigonometric generation of the ROM samples and -84.2 dB for the coarse-fine generation of ROM samples. This demonstrates an improvement of 12 dB for the coarse-fine technique while simultaneously requiring less storage.

### Sine-Phase Difference Algorithm

Since the coarse-fine algorithm can be used to store any arbitrary function, another technique may be employed to achieve even greater reduction in ROM size. This may be accomplished by storing the the function

$$f(x) = \sin(\pi x/2) - \pi x/2 \qquad (10)$$

instead of the sine function in the look-up table. The sine function is then generated by adding the phase argument from the phase accumulator to the output of the ROM's, as shown in Figure 9. The advantage of this technique is that it reduces the maximum amplitude of the function to be stored in the coarse ROM. Therefore the coarse ROM output word-length may be reduced and the output adder may be simplified. This modification reduces the look-up table storage requirements by at least of $2^{A+B+1}$ bits. The application of this technique to the hypothetical (4,4,4) segmented 12 bit output DDFS design considered earlier results in a savings of 512 bits in the coarse ROM and an increase in the compression ratio from 11.7:1 to 13.5:1. The penalty for this reduction in ROM storage is the addition of another adder. However, this is an advantageous tradeoff in the high speed VLSI implementation of a DDFS. The ROM propagation delay, which cannot be pipelined, is reduced, increasing the maximum clock frequency of the DDFS. The expense of another adder, which is readily pipelineable, is inconsequential in the full custom VLSI implementation.

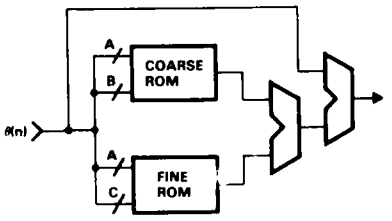



**Figure 9. Sine-Phase Difference Algorithm.**

### Exploitation of Fine ROM Symmetry

The fine ROM size may also be reduced by exploiting the symmetry in the samples. As shown in Figure 10, if the coarse ROM samples are chosen in the middle of the interpolation region, then the fine ROM samples will be approximately symmetric about the $\chi = (2^C - 1)/2$ axis. Thus by using an adder/subtracter instead of and adder to sum the coarse and fine ROM values, the size of the fine ROM may be halved. Some additional complexity must be added to the adder/subtracter control logic if this technique is to be used with the sine-phase difference algorithm, since the slope of the function in equation (10) changes sign at a non symmetric point between 0 and $\pi/2$ on the x-axis. Fortunately, the digital logic required to perform this can be accomplished with less than four logic gates for the 12-bit phase case considered previously.

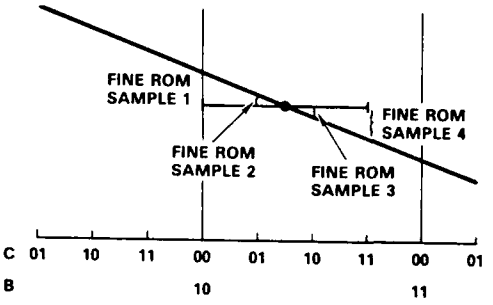



**Figure 10. Exploitation of Fine ROM Symmetry.**

Since the fine ROM is not generally in the critical speed path, the effective resolution of the fine ROM may be doubled at no extra hardware penalty, rather than halving the ROM. This is an extremely powerful technique because it allows the segmentation of the compression algorithm to be changed, effectively adding an extra bit of phase resolution to the look-up table, which thereby reduces the magnitude of the worst case spur due to phase accumulator truncation.

### 100 MHz DDFS Architecture Demonstration Vehicle

These architectural techniques were used to design a DDFS to meet a performance goal of better than -90 dB of spurious rejection with the potential of being pipelined to achieve a 100 Mhz clock rate in a standard 1.25 μM CMOS process. A 15-bit phase precision in the look-up table address was chosen, which from Figure 2 is shown to provide -90.3 dB of rejection in the magnitude worst case spur due to phase accumulator truncation. Next, the 15-bit phase must be segmented to address the coarse and fine ROM's. In this design it is assumed that the sine phase difference algorithm will be used, exploiting the symmetry of the fine ROM samples to double the effective fine ROM resolution. It is also assumed that the phase accumulator modification shown in Figure 5 is used, and that the quarter wave symmetry of the sinewave will be exploited, reducing the effective look-up table address requirements to 13 bits of phase precision for $0 < \theta < \pi/2$. To obtain the optimal segmentation of the 13-bit phase address to the coarse-fine ROM's, infinite output word-lengths were assumed in the simulations. Furthermore, it was heuristically determined that optimal segmentations of the phase argument have values of B and C that differ by at most 1, and that segmentations with values of $B > C$ perform better than $B < C$. Therefore, the optimization space of the word-lengths A, B, and C, were constrained such that $B = C$ or $B = C + 1$. Subject to these constraints, and that $A + B + C = 13$, the spurious response of the possible segmentations of the 13-bit phase were simulated. The results show the magnitude of the worst case spur in the spectrum plotted against the word-length variable A, and are pictured in Figure 11. A key concept to note in the graph is that the amount of storage required increases exponentially with increasing A, since the total ROM storage is given by $2^A(2^B + 2^C)$. Therefore the minimum value of A that meets the spurious rejection requirement should be chosen. The value of $A = 4$ corresponding to -95.33 dB meets the -90 dB requirement, which results in a phase segmentation of $A = 4$, $B = 4$, and $C = 5$. One anomaly in the graph in Figure 11 is the decreasing spurious performance as A is increased past 6 bits. This is a result of the fine ROM symmetry assumptions breaking down for small values of B. A similar set of simulations for a design without the exploitation of fine ROM symmetry show a monotonically increasing spurious performance with increasing values of A, as expected from theory.

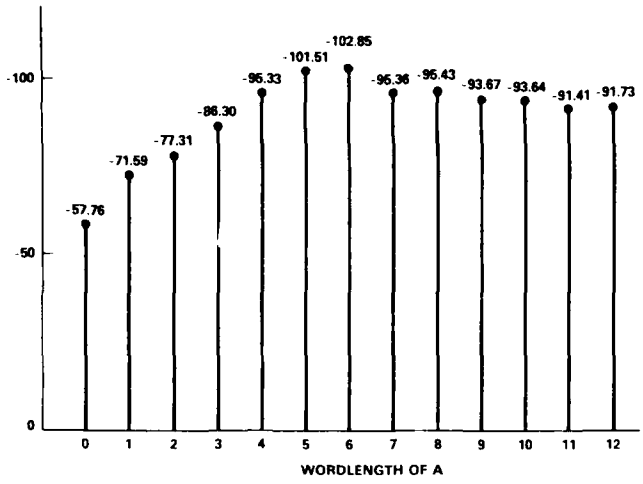



**Figure 11. Worst Case Spurious Response for Different Phase Segmentations.**

For the (4,4,5) segmentation of the phase, the effect of various output quantizations of the coarse ROM were simulated to determine the minimum word-length required to achieve the -90dB performance goal. The results of these simulations are shown in Figure 12. From these simulations, an output word-length of 9 bits was chosen for the coarse ROM, providing -89.9dB of spurious rejection including the effects of fine ROM quantization. The fine ROM word-length is provided as part of the simulation algorithm, where it is derived from the maximum amplitude of the samples calculated using equation (8). This resulted in a fine ROM word-length of 4 bits for a coarse ROM word-length of 9 bits.

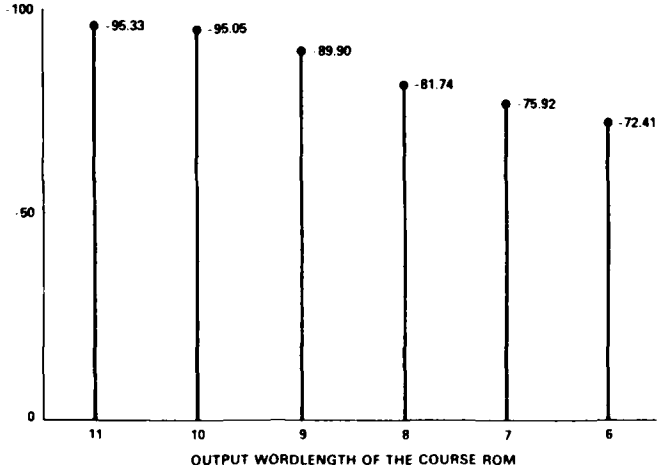



**Figure 12. Spurious Response for Different Coarse ROM Output Word-Length Quantizations.**

As a final optimization step, the amplitude values of the function to be stored in the look-up table can be scaled to provide improved performance in the presence of amplitude quantization. To accomplish this, the coarse and fine ROM samples were regenerated after multiplying the pre-quantized sinewave by a scaling constant that was varied in very small increments in the neighborhood of unity. The results of these simulations, which are graphed in Figure 13, show that a 1.75 dB improvement in the worst case spurious response can be ac ved by using a scaling constant of .999994 instead of 1.0 to ₂ rate the coarse and fine ROM samples. This results in a negligible reduction in the amplitude of the synthesizer output, while providing a worst case spurious that is now limited by the spurs due to phase accumulator truncation rather than amplitude quantization.

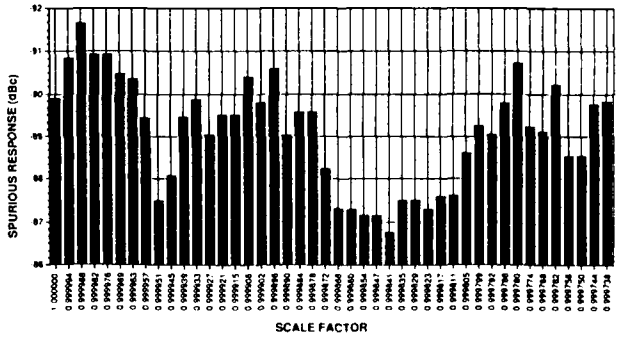



**Figure 13. Spurious Improvement From Scaling Sinewave Amplitude.**

## DDFS Architecture Demonstration Vehicle

Figure 14 shows the detailed logic block diagram of the DDFS design exploiting all architectural innovations discussed in this paper. The design has a simulated worst case spurious response of -90.3 dB using only 3,072 bits of ROM storage, providing a total storage compression ratio of 37:1. The simulated output spectrum of this design is depicted in Figure 15, showing the high degree of digital spurious rejection attained using this architecture. This design provides a 14-bit sinewave output, where the most significant 12 bits represent the ideal 12-bit rounded samples of a sinewave. This design is presently being implemented as a single VLSI chip in TRW's VHSIC 1.25μM CMOS process. Detailed circuit simulations predict the chip will be capable of operating at a clock rate of 100 Mhz, providing a .023 Hz frequency resolution over a 33 MHz bandwidth. The key circuit structures have already been fabricated in TRW's CMOS process as a test chip, shown in Figure 16, and have validated all logic and circuit simulations. The final design is scheduled to enter fabrication in November '88.

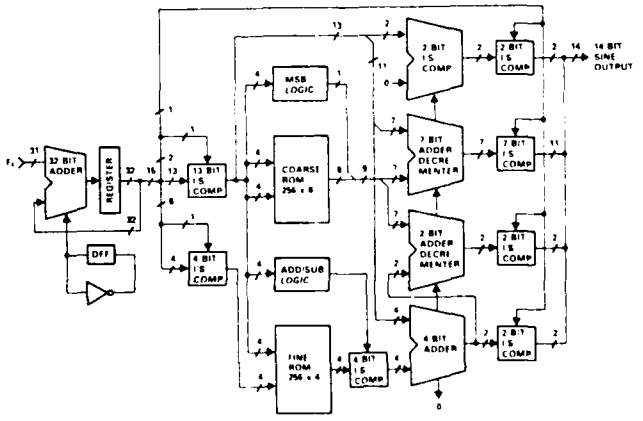



**Figure 14. -90.3 dB Spurious DDFS Architecture.**

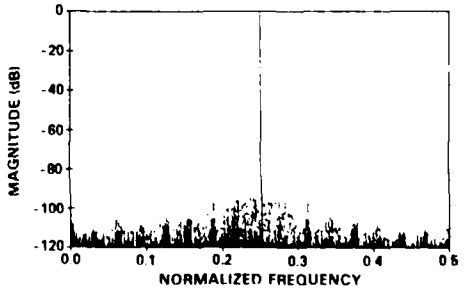



**Figure 15. Simulated DDFS Output Spectrum.**

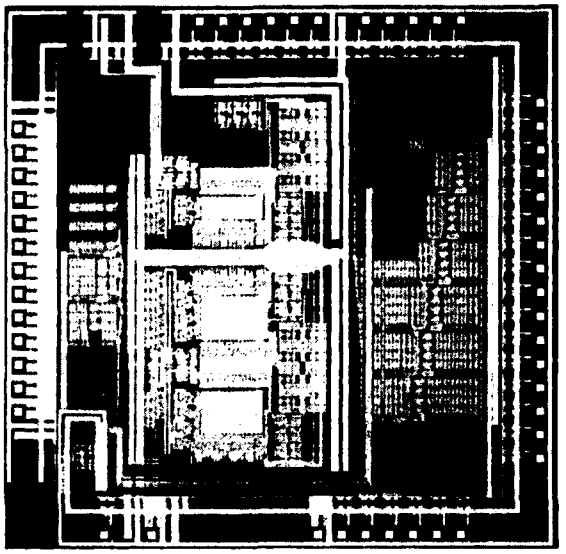


**Figure 16. 1.25 μM CMOS DDFS Test Chip.**

## Summary

Techniques for the design of VLSI architectures for direct digital frequency synthesis have been introduced, which allow for the optimization of the spurious response in the presence of finite word length effects. These optimization techniques exploit certain number theoretic properties of the phase accumulator to make the exhaustive simulation of DDFS performance in the presence of different system nonlinearities computationally feasible. These design techniques have been applied to design a 14-bit output DDFS with a simulated spurious performance of -90.3 dB and a level of pipelining capable of allowing a 100 MHz clock rate in a 1.25 μM CMOS process.

## Acknowledgements

The authors wish to thank M. Miscione of TRW Microelectronics Center for providing VLSI fabrication services and partial funding for this research. P. Hadinger and R. Smith of TRW Defense Communications Division also provided partial financial sponsorship and many valuable technical discussions on the systems applications of DDFS's.

## References

[1] J. Tierney, C.M. Rader, and B. Gold, "A Digital Frequency Synthesizer," IEEE Trans. Audio Electroacoust., vol. AU-19, pp. 48-57, 1971

[2] H.T. Nicholas III and H. Samueli, "An Analysis of the Output Spectrum of Direct Digital Frequency Synthesizers in the Presence of Phase-Accumulator Truncation," Proc. 41st annual Frequency Control Symposium USERACOM, Ft. Monmouth, N.J., May 1987, pp.495-502

[3] D.A. Sunderland, R.A. Strauch, S.S. Wharfield, H.T. Peterson, and C.R. Cole, "CMOS/SOS Frequency Synthesizer LSI Circuit for Spread Spectrum Communications" IEEE Journal of Solid State Circuits, vol. SC-19, pp.497-505, Aug 1984

[4] R.E. Lundgren, V. S. Reinhardt, and K. W. Martin, "Designs and Architectures for EW/Communication Direct Digital Synthesizers", Research and Development Technical Report SLCET-TR-85-0424-1, U.S. Army Laboratory Command, Fort Monmouth, New Jersey, Aug 1986

[5] Albert L. Bramble, "Direct Digital Frequency Synthesis", Proc. 35th annual Frequency Control Symposium, USERACOM, Ft. Monmouth, N.J., May 1981, pp.406-414

[6] J. Gorski-Popiel, Frequency Synthesis: Techniques and Applications, New York: IEEE press, 1975

[7] V. Manassewitsch, Frequency Synthesizers, Theory and Design, New York: John Wiley, 2nd Ed., 1979

[8] H.T. Nicholas, "The Determination of the Output Spectrum of Direct Digital Frequency Synthesizers in the Presence of Phase Accumulator Truncation", Thesis, Masters of Science in Electrical Engineering, UCLA, 1985.

42nd Annual Frequency Control Symposium - 1988

# A PHASE-LOCKED 4-6 GHz LOCAL OSCILLATOR USING MICROWAVE PRESCALER

Pramode C. Kandpal

Rockwell International
Dallas, TX 75207

## Summary

A phase-locked broadband local oscillator using a divide-by-8 microwave prescaler is described. The measured phase noise characteristics of the prescaler, the reference source, and the phase-locked source are discussed. Threshold bit error rate (BER) curve of a 6-GHz 64-QAM microwave digital radio using this direct division source is given and compared with a source using a multiplier and mixer configuration.

## Introduction

Microwave radios generally employ a phase-locked VCO locked to a harmonic of a crystal reference as a local oscillator. With the commercial availability of the high speed prescalers[1] up to 10 GHz, it is now feasible to phase-lock a fundamental frequency microwave oscillator to a low frequency stable source by dividing the microwave output. This approach yields a simplified microwave frequency source with good phase noise characteristics and low power dissipation and costs less than the multiplier and mixer method. This paper presents the performance of a direct division source using a microwave prescaler for local oscillator applications in the 4- and 6-GHz common carrier bands.

The block diagram of the circuit used is shown in figure 1. An electronically tunable dielectric resonator oscillator[2] is used as the fundamental microwave source. A buffer amplifier with a gain of 8 dB and isolation better than 20 dB is used at the output of the oscillator to provide isolation from the prescaler and increase the output to 17 dBm. The output from the directional coupler is applied to the divide-by-8 microwave prescaler, followed by two cascaded divide-by-4 ECL silicon prescalers. This brings the divided frequency to 27-51 MHz for the operational bandwidth of 3.5-6.5 GHz.

The reference oscillator uses a quartz crystal on the third overtone of the resonator. A buffer amplifier and two gain stages follow the oscillator. A Motorola MC12040 phase-frequency detector is used to determine the phase/frequency error between the divided and the reference signal. The output of the detector is summed up in a loop amplifier/filter, which determines the phase locking loop bandwidth. The error voltage from the amplifier is fed to the varactor on the microwave oscillator.

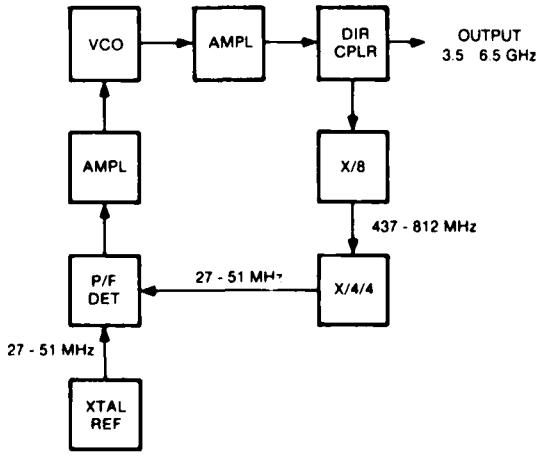
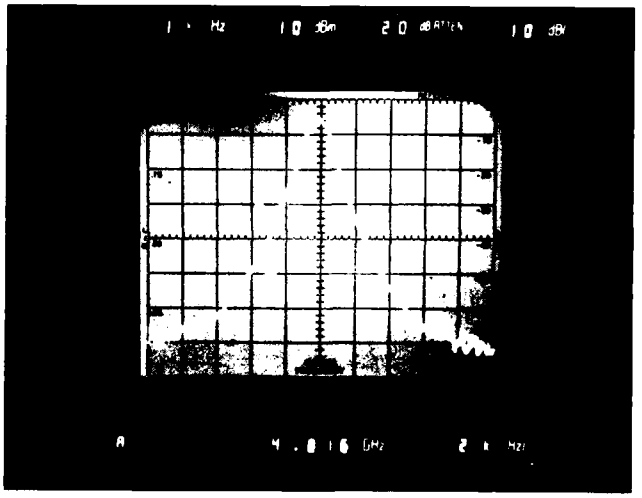


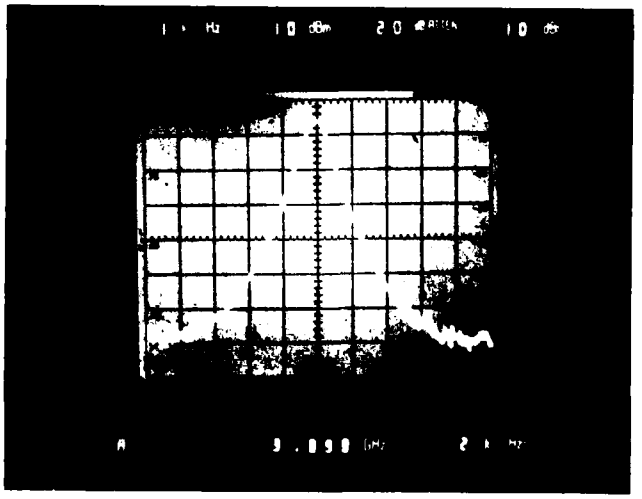

**FIGURE 1. BLOCK DIAGRAM OF THE PHASE-LOCKED VCO.**

## Experimental Results

Figure 2 shows the output of the 4-GHz VCO in the free-running and the phase-locked modes. It can be noticed that the free-running oscillator has only minor phase perturbations because of the high-Q tank circuit, which are further suppressed by the phase lock loop. The phase locking loop bandwidth is about 10 kHz which can be seen better in the phase noise curves of figures 3 and 4. The locking bandwidth (for pulling a drifting oscillator to the set frequency) was set at 25 MHz to compensate for the frequency drifts caused by load impedance, temperature changes, and aging. The phase noise was measured using the HP3047A and is shown in figures 3 and 4 at 4 and 6 GHz, respectively. The curves show the phase noise of the crystal reference source, free-running microwave source, and the phase-locked source. The prescaler ratio of 128 corresponds to a 42-dB (20 log 128 = 42 dB) increase in the phase noise of the phase-locked source compared to the



**(A) FREE RUNNING**



**(B) PHASE-LOCKED**

**FIGURE 2. SPECTRUM DISPLAY OF 4-GHz VCO IN FREE-RUNNING AND PHASE-LOCKED MODES.**



CH2588-2/88/0000-364 $1.00 C 1988 IEEE

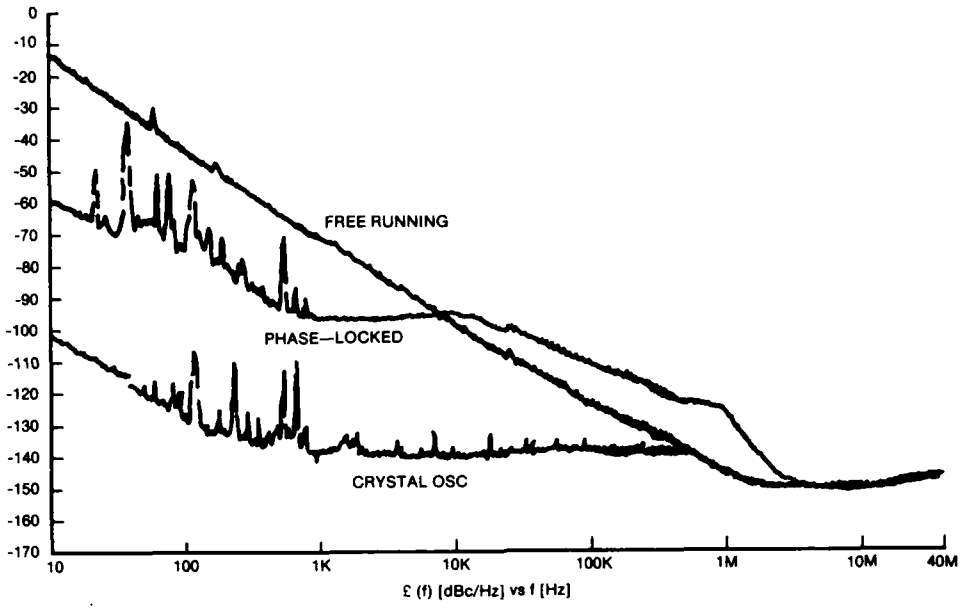



**FIGURE 3. PHASE NOISE CHARACTERISTICS AT 4 GHz.**

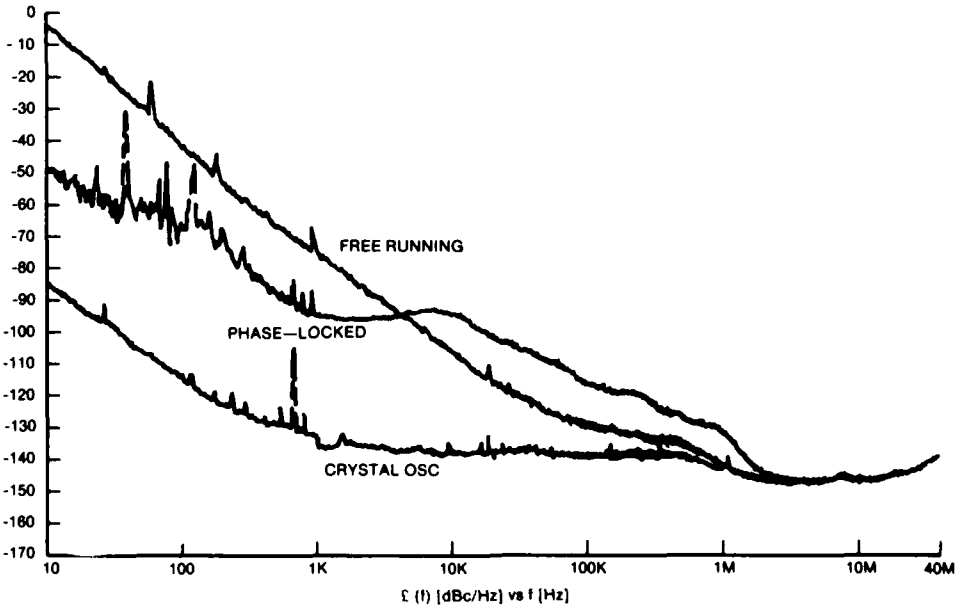



**FIGURE 4. PHASE NOISE CHARACTERISTICS AT 6 GHz.**

crystal reference, as evident from figures 3 and 4. The phase noise slope of the locked source follows the multiplied crystal source up to the phase lock loop bandwidth of 10 kHz. Beyond that, the curve follows the slope of the free-running source. The results obtained at 6 GHz are slightly inferior to that at 4 GHz because the crystal used in the reference oscillator had a relatively poor phase noise. The overall performance, however, is good enough for use in a digital microwave radio, as discussed later.

A low-noise synthesizer was used as the source of input to the prescaler to evaluate the noise contribution of the prescaler itself. The data is shown in figures 5 and 6 at 4 and 6 GHz, respectively. The phase noise of the output is reduced by 18 dB as expected. Near the carrier, however, phase noise is much higher because of the 1/f noise from the FETs. This data only demonstrates the level of phase noise generated by the prescaler relative to the phase noise of the synthesizer. The actual noise level from the prescaler is less than that shown in figures 5 and 6 unless the synthesizer is noise free. In any case, the overall noise from the prescaler is much less than that of the crystal multiplied by the prescaler ratio and thus does not contribute to the noise sidebands of the phase-locked microwave source.

Figure 7 shows the division threshold sensitivity for the prescaler as a function of the frequency. The input power requirements for the prescaler are less than −5 dBm for 4-6 GHz. Figure 8 shows the input/output of the prescaler at 6 GHz using a sampling oscilloscope.

The 6 GHz phase-locked source was evaluated in a 6-GHz 64-QAM digital radio in the down-converter. In this application, the single-sideband noise of the local oscillator will degrade the threshold performance. The performance of the receiver and thus the source, is measured in terms of bit error rate (BER). Figure 9 shows the data taken on a 6-GHz receiver using the direct division source versus a source using a multiplier/mixer. The performance obtained was satisfactory in both cases, which also indicates that it ultimately depends upon the LNA used.

## Conclusion

A broadband 3.5-6.5 GHz phase-locked source using prescaler has been described. It is found that phase noise of the microwave source is improved drastically in the phase lock loop bandwidth. Outside, the phase noise remains almost the same as that of a free running source. The microwave prescaler is found to contribute no additional noise to the phase-locked microwave source. The measured data is in excellent agreement with the theoretical predictions. The scheme is simpler and more cost effective than a mixer/multiplier configuration. The circuit has the advantage of requiring only that the crystal be replaced to set the local oscillator to a different frequency.

## References


1. Fujitsu GaAs Microwave Frequency Divider Data Sheet, FMM106HG.

2. P. C. Kandpal and C. Ho, "A Broadband VCO Using Dielectric Resonators," 1988 IEEE MTT-S International Microwave Symposium Digest, pp. 609-612.

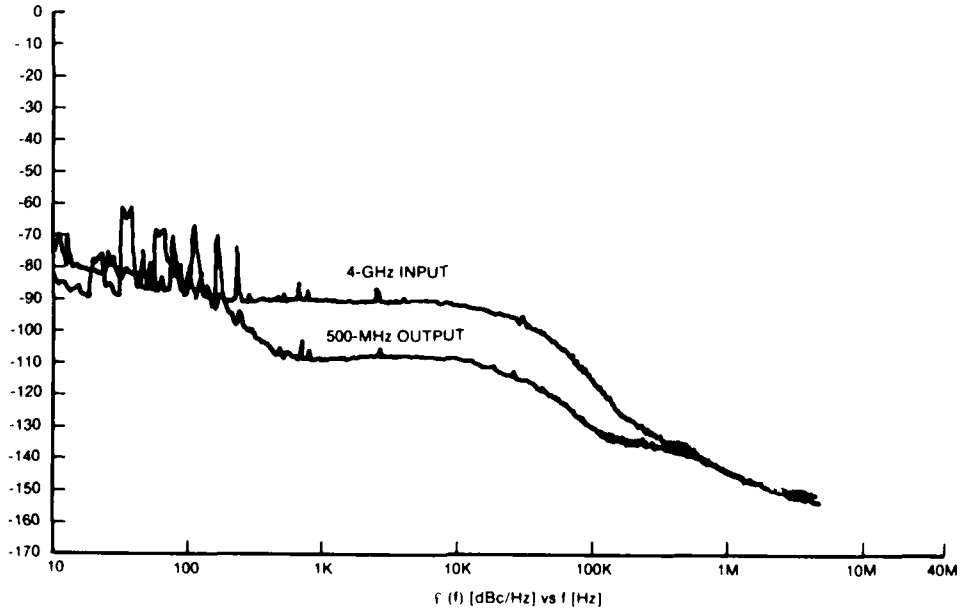




**FIGURE 5. PHASE NOISE CHARACTERISTICS OF THE PRESCALER AT 4 GHz.**

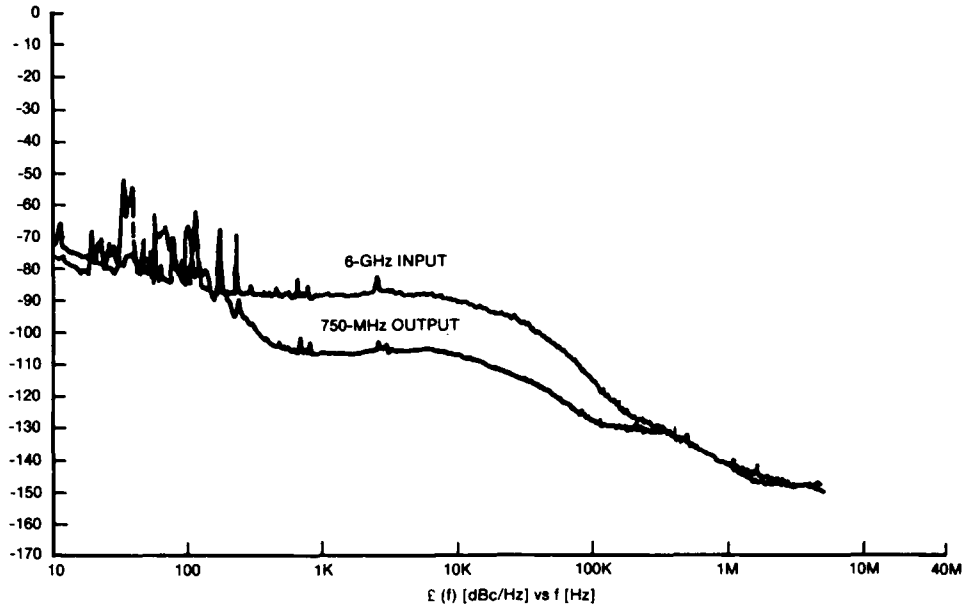



**FIGURE 6. PHASE NOISE CHARACTERISTICS OF THE PRESCALER AT 6 GHz.**

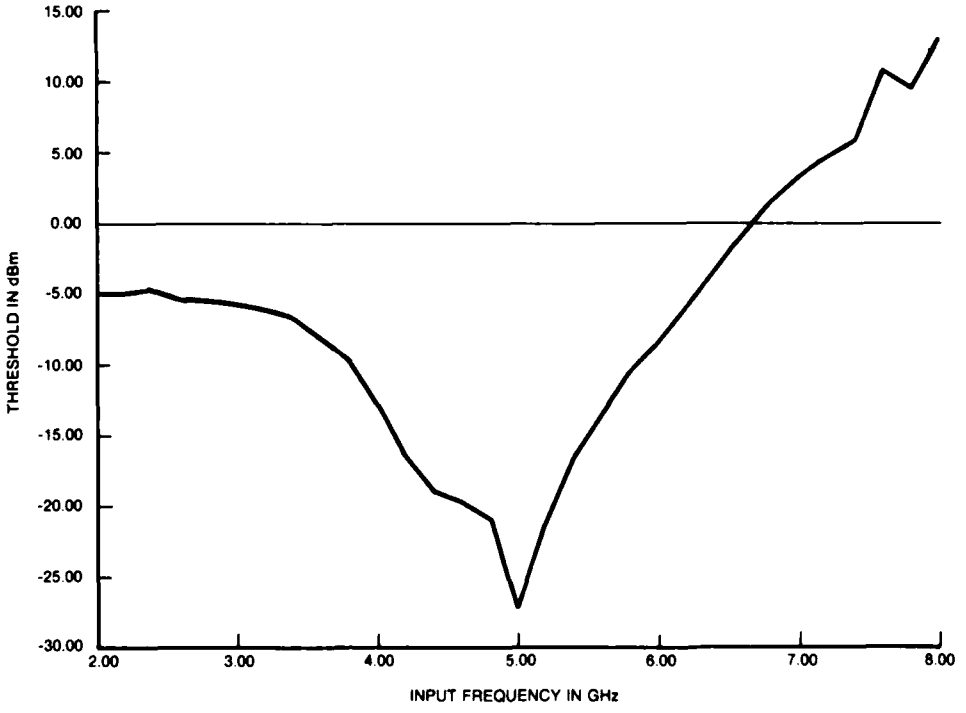



**FIGURE 7. DIVISION THRESHOLD SENSITIVITY VS FREQUENCY.**

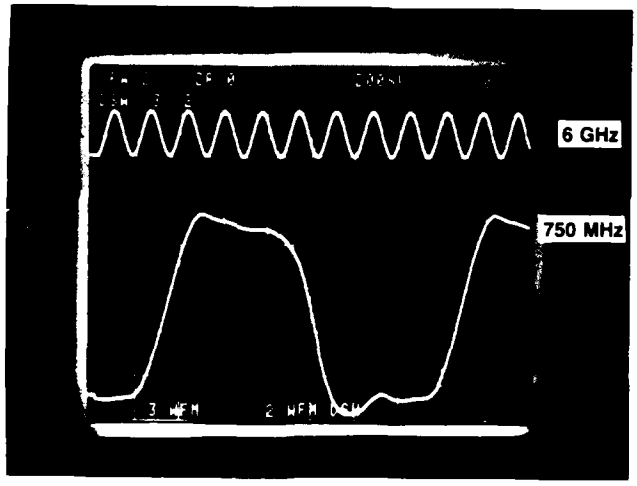



**FIGURE 8. INPUT/OUTPUT OF THE X/8 PRESCALER.**

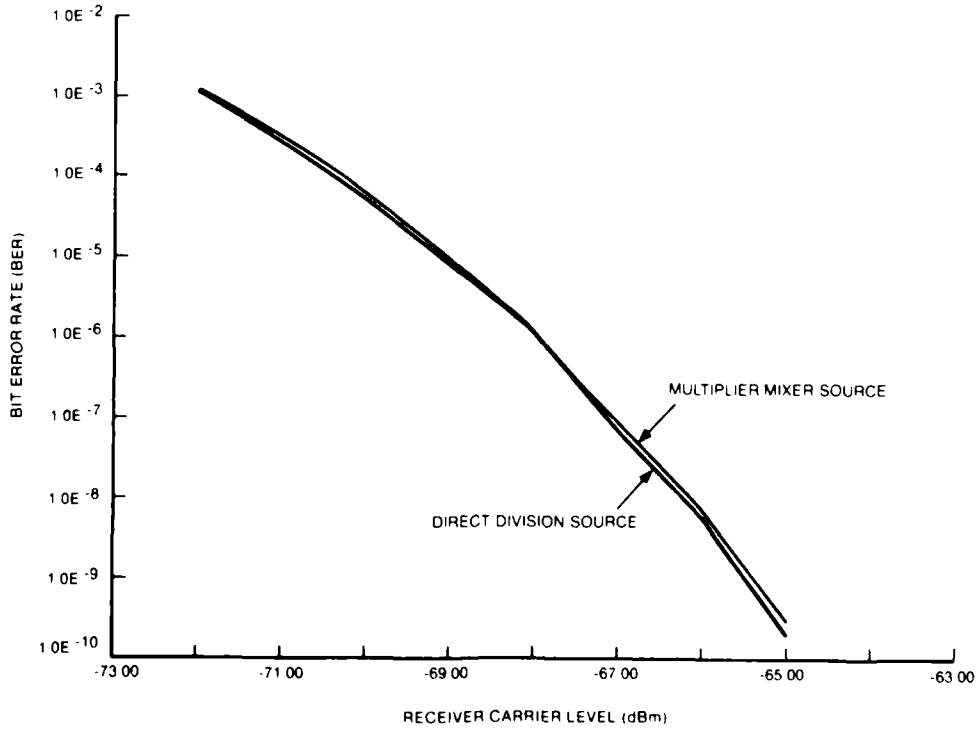



**FIGURE 9. BER CURVE FOR A 6-GHz 64-QAM DIGITAL RADIO.**

42nd Annual Frequency Control Symposium - 1988

# LOW NOISE, MICROWAVE SIGNAL GENERATION USING BULK AND SURFACE ACOUSTIC WAVE RESONATORS

M. M. Driscoll
Westinghouse Electric Corporation
Electronic Systems Group
Baltimore, Maryland 21203

## Summary

The use of high Q, bulk and surface acoustic wave resonators provides a means for generation of low noise, microwave, carrier signals.

The resulting signal phase noise spectrum consists of two distinct regions. In the near carrier region, the spectrum is characterized by flicker of frequency noise usually due to short term instability in the oscillator acoustic resonator itself. The spectral density of the fractional frequency fluctuations, $S_y(f)$, provides a convenient measure for comparing the short term frequency stability of resonators operating at different frequency ranges. In general, acoustic resonator $S_y(f)$ appears to vary as the inverse of the square of the device unloaded Q. Therefore, lowest near-carrier noise levels are obtained via frequency multiplication of a lower frequency oscillator using a correspondingly higher Q resonator.

Frequency multiplication, however, results in higher relative noise floor levels at high carrier offset frequencies. Techniques for obtaining low noise floor levels include:

1) Utilization of the oscillator resonator out-of-band impedance for reduction of sustaining stage transistor collector noise current;

2) Extraction of the oscillator output signal through the resonator;

3) Use of narrowband, passive and active (phaselocked oscillator) spectral cleanup filters using high Q acoustic resonators;

4) Use of low flicker of phase noise circuitry in the oscillator sustaining stage, frequency multipliers, and amplifiers; and,

5) Operation of circuits at acceptably high RF drive levels.

Methods for obtaining further microwave signal short-term frequency stability improvement include reducing the effects of conventional bulk and surface wave resonator self noise and use of unconventional resonator configurations exhibiting significantly higher Q than obtainable with quartz. In connection with these efforts, recent significant progress will be reported. This includes achievement in 40 MHz prototype oscillators, of output signal Sy(100 Hz) = $3\times10^{-27}$ [ℒ(100 Hz) = -156 dB/Hz] using conventional AT-cut quartz resonators exhibiting Sy(100 Hz) = $1\times10^{-26}$ to $1.5\times10^{-26}$ flicker noise levels.

In addition, HBAR (high overtone, bulk acoustic resonator) and UHF HBAR oscillator performance will be described in which resonator loaded Q = 160,000 at 320 MHz and self-noise equivalent to Sy(100 Hz) = $1\times10^{-24}$ have been achieved. The oscillator employs low noise 50 ohm modular amplifiers, a tunable (1 pole) helical filter for selecting desired HBAR resonance, and a lumped element delay line used to set required oscillator closed loop phase shift.

## Introduction

Requirements for improved radar and communication system performance, in terms of increased dynamic range and decreased signal detection thresholds, necessitate the achievement of improved spectral purity in the microwave transmitter and receiver local oscillator signals.

Generation of highest stability microwave signals is accomplished using high Q bulk and surface acoustic wave (BAW and SAW) resonators for oscillator stabilization and narrowband 'spectral cleanup' filters. The choice of resonator type, operating frequency, drive level, and circuit utilization scheme significantly impact the degree of output signal frequency stability obtained. In this paper, the emphasis is on reduction of signal phase noise due to electrical effects where the region of interest includes carrier offset frequencies in excess of several Hz.

It is very important to note that the (resonant frequency) sensitivity of the acoustic resonator to environmental stress, such as short-term temperature fluctuations and vibration, is such that these effects can degrade the signal spectra to a far greater degree than that due to electrical noise. This is especially true in the case of vibration encountered in airborne platforms. The solution to problems associated with vibration-induced FM sidebands in acoustic resonator-stabilized, signal generation hardware constitutes an active area of study in the frequency control community.

Unfortunately, there are no hard circuit design rules that can be applied to the solution of the (electrical) noise problem. For example, techniques yielding superior phase noise floor performance at high carrier offset frequencies normally do so at the expense of relatively poorer performance at low offset frequencies. Also, tradeoffs always need to be made with regard to obtaining required signal spectral performance margin verses overall circuit complexity, reliability, power consumption, etc.

In this paper, an attempt is made to acquaint the reader with some of the more important resonator utilization and associated circuit design considerations involved in the generation of low noise microwave signals.

CH2588-2/88/0000-369 $1.00 C 1988 IEEE

## Signal Spectral Relationships in Acoustic Resonator-Controlled Oscillators

Figure 1 shows two alternative methods for designing (and analyzing) acoustic resonator - stabilized, harmonic oscillator circuits. As shown in the figure, the sustaining stage portion of the circuit can be considered as: (a) a feedforward amplifier with the resonator included in the positive feedback circuit, or (b) a two terminal, negative resistance generator to which the resonator is attached.

The steady state conditions for oscillation are: unity closed loop gain and 2Nπ closed loop phase shift [figure 1(A)] or resonator impedance equal in magnitude and opposite in sign compared to that of the negative resistance generator [figure 1(B)]. The input impedance of the negative resistance generator sustaining stage can contain a substantial reactive component, depending on the type of oscillator circuit utilized.

Similarly, the conditions necessary to initiate oscillation in the circuits of figure 1(A) and 1(B), respectively, are: (a) greater than unity closed loop gain and 2Nπ closed loop phase shift; and (b) net negative resistance and zero reactance in the sum (series connection) of the resonator and negative resistance generator impedances. Steady state conditions are achieved via sustaining stage gain or negative resistance reduction at a prescribed circuit RF drive level by employing auxiliary AGC or ALC circuitry, or by biasing the sustaining stage transistors to current limit at the desired quiescent signal level. The transistors should not be biased so that voltage limiting (saturation) occurs since severe degradation in output signal stability results.

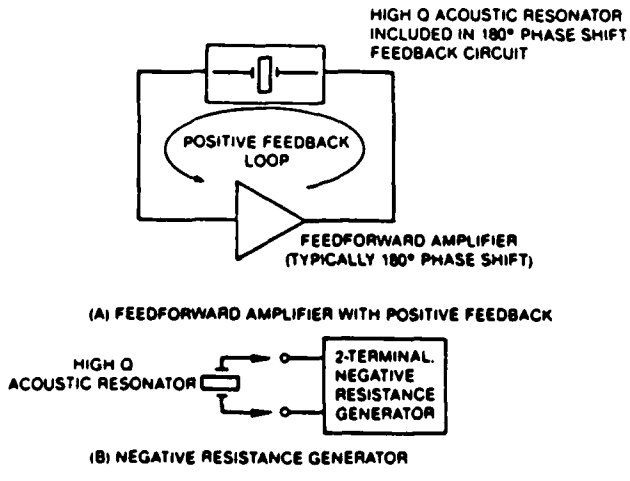



Figure 1. Basic Oscillator Configurations

Figure 2 shows how the sustaining circuit open loop phase noise sideband spectrum is related to that of the (closed loop) oscillator output signal. In the oscillator, there is a conversion of open loop circuit signal phase perturbations to closed loop signal frequency perturbations. This conversion is a consequence of required maintenance of the steady-state, quiescent signal phase relationships, and it is proportional to the reciprocal of the closed loop signal group delay. The delay is primarily determined by the resonator loaded Q [Ref. 1].

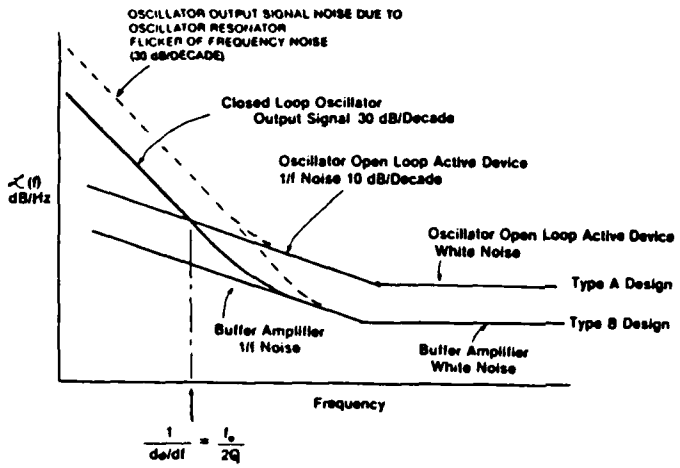



Figure 2. Typical Phase-Noise Sideband Spectra for VHF Crystal Oscillators

For a single resonator:

$$\frac{d\phi}{df} = \frac{2Q_L}{f_o}, \text{ where } Q_L = \text{loaded Q} \qquad (1)$$

As shown by the solid curves of figure 2, the effect of the conversion is a (20 dB/decade) increase in output signal phase noise sideband level at carrier offset frequencies less than the resonator half-bandwidth. The reason SAW delay line oscillators cannot provide levels of spectral purity obtainable using SAW resonators is that intentionally lower group delay must be used in the delay line devices to prevent multiple (360-degree phase shift) oscillator operating points from occurring in the delay line passband.

In many cases, the near-carrier (flicker of frequency) noise level exhibited by acoustic-resonator-stabilized oscillators is significantly higher than that predicted by measurement of sustaining stage phase noise and resonator loaded Q. It has been found that the "excess" flicker noise (shown by the dashed curve in figure 2) is due to short term frequency instability in the resonator itself [Ref. 2 thru 8]. In addition, variations in excess of 30 dB have been observed in the flicker of frequency noise level exhibited by "identically" fabricated bulk and surface acoustic wave resonators [Ref. 7,8]. Determination of the physical mechanisms responsible for resonator flicker noise, together with methods for increasing the yield of low noise devices, constitutes an area of active investigation in the quartz crystal industry. In general, resonator-induced flicker noise is usually dominant at carrier offset frequencies extending to (and sometimes several octaves beyond) the resonator half-bandwidth.

At offset frequencies in excess of the resonator half-bandwidth, the oscillator signal "white" phase noise sideband (or noise floor) level is determined by the circuit noise sources acting in combination with the out-of-band impedance exhibited by the resonator. In this respect, the way in which the resonator is embedded in the oscillator circuit plays a significant role in achievable output signal noise floor levels. This is depicted in curves A and B in figure 2.

In general, the architecture chosen in the design of a microwave signal source exhibiting specified

levels of short-term frequency stability involves determination of the following:

A. Oscillator (resonator) frequency of operation with respect to the desired final output frequency. This includes selection of resonator type (bulk or surface wave) and estimation of obtainable resonator flicker noise level and loaded Q.

B. Type of sustaining stage used. This includes consideration of circuit noise sources and methods for minimizing their effect.

C. Possible utilization of narrowband, acoustic resonator, spectral cleanup filters. This includes use of either passive filters or phaselocked, acoustic resonator-stabilized, voltage-controlled oscillators (VCOs).

## Oscillator (Resonator) Frequency of Operation

In the frequency multiplication process, the carrier signal frequency and/or phase fluctuations are increased. The signal phase noise sideband level increases as 20 log M, where M is the multiplication factor. As a result, lower microwave signal phase noise floor levels can be obtained by operating the acoustic resonator-stabilized oscillator at the highest possible frequency (lowest multiplication factor). Unfortunately however, the reverse is true with regard to the flicker-of-frequency portion of the carrier signal spectrum.

As was explained earlier, low flicker noise stage (amplifier) circuitry is readily available such that the resulting oscillator signal flicker of frequency noise is due to short-term frequency instability in the acoustic resonator itself.

The spectral density of the carrier signal fractional frequency fluctuations, $S_y(f)$, is a convenient comparative measure of resonator flicker noise level because $S_y(f)$ is unaffected by (noiseless) frequency ultiplication. Thus, a microwave signal generated directly at 1 GHz using a SAW resonator-stabilized oscillator would exhibit a near-carrier (flicker-of-frequency) spectra identical to that obtained via times ten multiplication of a 100 MHz (bulk wave) quartz resonator-stabilized oscillator if both resonators exhibited identical $S_y(f)$ spectral densities.

Although large variations in resonator flicker noise levels have been measured, a "nominal" relationship between resonator noise level and unloaded Q has been observed.

In general, it has been found that resonator flicker noise, expressed as $S_y(f)$, varies as the inverse of the square of the unloaded Q [Ref. 6]. Because a given material (such a quartz) exhibits a fairly constant Q-f product, higher frequency devices exhibit lower unloaded Q and correspondingly higher flicker noise levels. For example, SAW quartz resonator noise measurements of 800 MHz devices having $Qu = 8 \times 10^3$ indicate a nominal level for $S_y(f = 1\ Hz) = 3 \times 10^{-21}$ [Ref. 7]. Similar measurements conducted at Westinghouse for 80 MHz, bulk wave quartz resonators having $Qu = 1.2 \times 10^5$ indicate a nominal level for $S_y(f = 1\ Hz) = 3 \times 10^{-23}$ [Ref. 8]. More recently, $S_y(f = 1\ Hz) = 3 \times 10^{-24}$ has been demonstrated at Westinghouse for 40 MHz, bulk wave resonators having $Qu = 2.5 \times 10^5$, and $Sy(f = 1HZ) = 8 \times 10^{-23}$ has been achieved at Raytheon for 500 MHz SAW resonators having $Qu = 1.8 \times 10^4$ [Ref. 9].

It can be seen then, that requirements for low carrier signal flicker-of-frequency noise demand use of lower frequency oscillators followed by frequency multiplication circuitry. However, the resulting (low flicker noise) signal will also exhibit a higher relative noise floor level due to the use of frequency multiplication.

A common technique for achieving both low flicker noise and noise floor levels involves use of narrowband, acoustic-resonator, spectral cleanup filters, as discussed in a following portion of this paper.

## Sustaining Stage Configuration

Figure 3 shows the simplified electrical equivalent circuits for bulk and surface acoustic wave resonators. The circuit of figure 3(a) consists of a series resonant motional impedance in parallel with a inter-electrode, or static capacitance. For commonly used quartz resonators, the static capacitance to series resistance impedance ratio (figure of merit) may be small enough to necessitate use of parallel inductance to tune out the static capacitance in order to achieve desired device impedance selectivity.

The two-port SAW resonator has the advantage that the static capacitance(s) do not appear in parallel with the motional impedance. However, two-port SAW resonator motional impedance levels are typically four times larger than for one-port configurations, resulting in increased device insertion loss and reduced tuning range using series-connected variable circuit reactance.

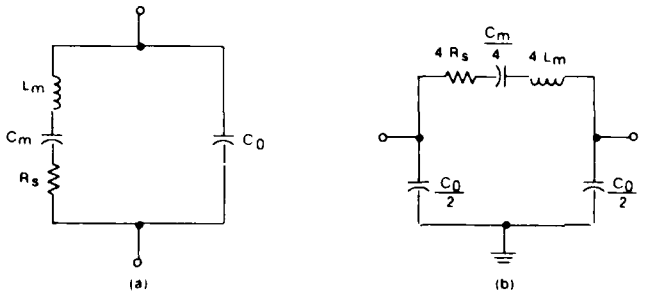



**Figure 3. Electrical Equivalent Circuits for Conventional Bulk and Surface Acoustic Wave Resonators**

Figure 4(a), 4(b) shows two single transistor sustaining stage configurations commonly utilized in BAW and 1-port SAW resonator-stabilized oscillators. As shown in the figure, the two circuits are quite similar when one replaces the resonator by its equivalent electrical impedance at the oscillator operating frequency. This is especially true if unbypassed emitter resistance is used in the figure 4(a) circuit. This is often the case, since the use of local negative feedback has been shown as a method for reducing transistor RF signal flicker of phase noise [Ref. 10].

Although there are numerous other types of single and multiple transistor sustaining stage configurations, it is useful to examine the effect of the acoustic resonator out-of-band impedance in determining the comparative noise floor performance in each of the figure 4 circuits.

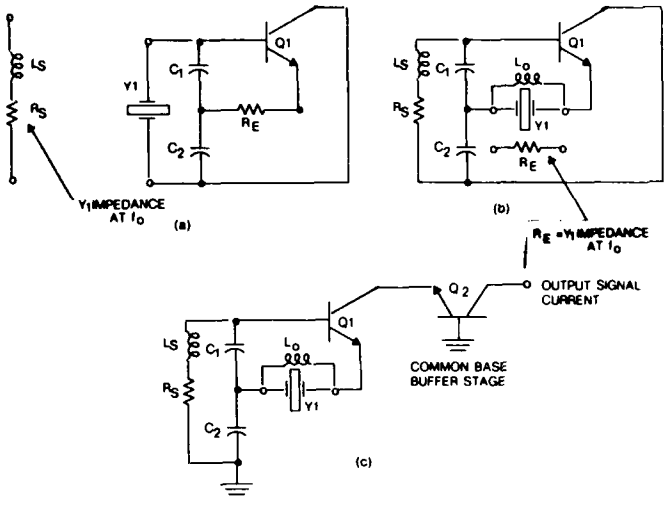
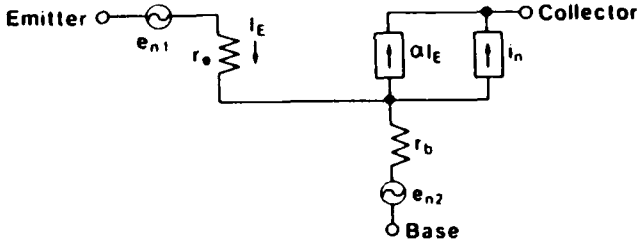



Figure 4. Single Transistor, Acoustic Resonator-Stabilized Oscillator Configurations

In the circuit of figure 4(a), the resonator is operated above series resonance and furnishes the required circuit inductive reactance. The absence of a requirement for inductors makes the figure 4(a) circuit especially attractive with regard to sustaining stage implementation in monolithic form.

In the circuit of figure 4(b), the resonator is operated at series resonance where it exhibits a resistive impedance. The figure 4(b) circuit is normally used for overtone mode resonator operation, and the selectivity of the L-C tank circuit inhibits the possibility of oscillation at the resonator fundamental or undesired overtone resonant modes. In figure 4(b), parallel inductance is normally used to anti-resonate the resonator static capacitance, and the resonator/inductor impedance magnitude increases very rapidly (and symmetrically) to relatively large reactive values at carrier offset frequencies in excess of the resonator bandwidth.

Figure 5 shows a simplified transistor model with (white) noise sources [Ref. 11]. When one considers the effect of a large (resonator out-of-band) impedance placed in series with the transistor emitter in figure 5, it becomes obvious that the effect is to suppress the noise current flowing in the collector due to both emitter shot noise and base resistance Johnson noise. In fact, the method used to determine the value of the (uncorrellated) collector shot noise (in figure 5) is to place a large impedance in series with the emitter.




$$\bar{e}_{n1}^2 \approx 2qI_E r_e^2 \Delta f \approx 2KTr_e \Delta f$$

$$\bar{e}_{n2}^2 \approx 4KTr_b \Delta f$$

$$\bar{i}_n^2 \approx 2q\left[\alpha_F I_E (1-\alpha_F) + I_{co}\right] \Delta f$$

At High Frequency

$$\bar{i}_n^2 \approx 2q\left[\frac{\alpha_F I_E (1-\alpha_F + f^2/f_\alpha^2)}{(1+f^2/f_\alpha^2)} + I_{co}\right] \Delta f$$

Where

$f_\alpha$ = alpha cutoff frequency

Figure 5. Simplified Transistor Model with (White) Noise Sources

Further, if the transistor is operated well below the alpha cutoff frequency, the magnitude of $i_n^{-2}$ is considerably reduced. This fact can and has been used to advantage, and figure 4(b) type oscillators have been constructed using high frequency transistors in which very low noise floor levels have been achieved at carrier offset frequencies where the resonator out of band impedance s large.

By contrast, the resonator impedance in the figure 4(a) circuit does not increase symmetrically from the operating frequency value. Also, at very high offset frequencies (where the resonator impedance is that of its static capacitance), substantial transistor noise source suppression does not occur because the resonator is not connected in series with the transistor emitter.

Figure 6 shows a typical result obtained for a 20 MHz voltage controlled crystal oscillator (VCXO). As shown in the figure, a -171 dB/Hz noise floor level (-168 dB/Hz for two, identical, phaselocked VCXOs) was achieved for only 200 microwatt crystal drive utilizing the sustaining stage configuration of figure 4(c). In the figure 4(c) circuit, useful output collector signal current is extracted through a common base buffer stage, and buffer transistor noise sources are similarly suppressed due to the large (sustaining stage collector) impedance in series with the buffer transistor emitter.

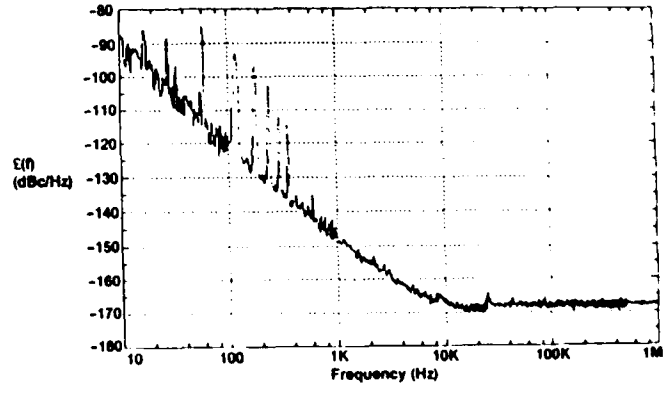



Figure 6. VCO Short-Term Frequency Stability for 200-Microwatt Crystal Dissipation

An additional method for achieving low noise floor levels is to extract the oscillator output signal current through the resonator itself, as shown in figure 7. In this type of configuration, the resonator performs the dual function of the oscillator stabilization element and the output signal spectral filter. Noise floor levels of -176 dB/Hz and -181 dB/Hz, respectively, have been obtained for 80 MHz, single and dual SC-cut, crystal controlled oscillators using figure 7 type circuit configurations [Ref. 12].

There are instances where use of an oscillator sustaining stage employing self-contained, 50 ohm, modular amplifiers may be preferable to the discrete component configurations such as those of figures 4 and 7.

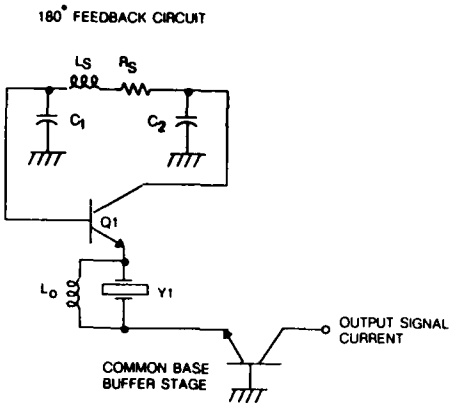



Figure 7. Oscillator Signal Extraction Through the Acoustic Resonator

As explained, the discrete component oscillator circuits like those of figures 4 and 7 result in lowest noise signal generation when the operating frequency is well below transistor cutoff frequency. In oscillators operating at high VHF and UHF frequencies using silicon bipolar transistors having $f_T$ in the 3 to 6 GHz region, circuit instability or marginal stability can occur as a result of circuit interconnecting lead inductance and stray capacitance.

Circuitry added to suppress these instabilities can degrade operating frequency circuit performance, and sustaining stage flicker of phase noise levels become more difficult to predict and evaluate.

High $f_T$, GaAs transistor circuitry cannot be used, owing to significantly higher device flicker noise levels.

Use of 50 ohm, modular amplifiers in acoustic resonator-stabilized oscillators allows "building block" design (and evaluation) methods to be used that are often preferable in the VHF/UHF frequency range. An advantage of this approach is that accurate, independent (open loop) characterization of sustaining stage amplifier performance, in terms of device small signal scattering parameters and operating point drive level phase noise levels, is easily accomplished using 50 ohm measurement apparatus.

These measurements indicate that amplifier flicker-of-phase noise levels of -160 dB/Hz at 100 Hz carrier offset can be reliably obtained. As a result, the contribution of amplifier 1/f noise on VHF and UHF oscillator output signal flicker-of-frequency noise is typically 10 dB below that resulting from short term instability in the acoustic resonators themselves. At higher modulation rates, the output signal noise floor level is determined by the amplifier effective noise figure and input signal drive level. Noise floor levels of -173 dB/Hz have been obtained for 80 MHz oscillators using 3rd overtone SC-cut quartz resonators and 4 dB noise figure, sustaining stage amplifiers [Ref. 13]. The circuit schematic diagram is shown in figure 8. The multiple decade bandwidth available using 50 ohm, hybrid and MMIC amplifiers indicates this approach will work quite satisfactorily with UHF bulk and SAW resonators as well.

## Narrow Bandwidth, Passive, Spectral Cleanup Filters

As implied in the previous discussion, a straightforward method for generation of microwave signals simultaneously exhibiting low (near-carrier) flicker

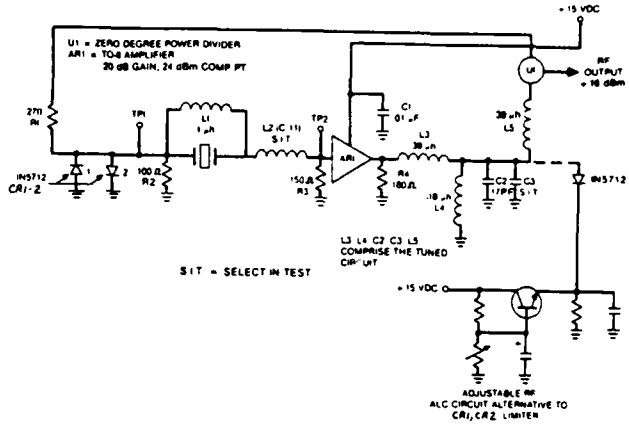



Figure 8. 80 MHz, SC Cut - Crystal Controlled Oscillator Using 50 Ohm Sustaining Stage Amplifier

of frequency noise and white phase noise (floor) levels involves use of a bulk wave, quartz crystal-controlled oscillator (typically in the low VHF range) followed by successive stages of low order, low noise frequency multiplication circuitry in combination with a narrowband spectral cleanup filter (typically operating in the low UHF band and incorporating quartz SAW resonators). Figure 9 shows the configuration in block diagram form.

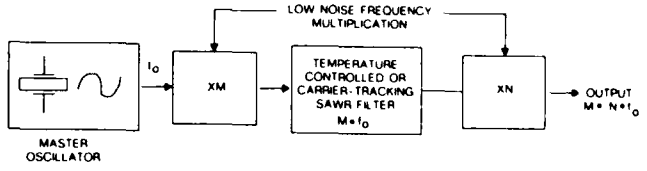



Figure 9. Microwave Signal Noise Floor Reduction Using a Narrow Bandwidth, Acoustic Resonator Filter

There are several advantages associated with use of passive spectral bandlimiting (compared to use of a phase-locked VCO). The passive filter approach is simpler to implement. Also, when the high Q acoustic resonator is used as a filter, resonator frequency instability due to device self-noise (as well as vibration sensitivity) result in carrier signal phase modulation. The phase modulation produces a lower relative sideband level at decreasing carrier offset frequencies than the frequency modulation that would occur if the resonator were used in an oscillator. Use of a single resonator (one pole filter) usually provides adequate selectivity since the filtered signal is a "zero bandwidth" CW one, and only modest filter shape factor and stopband attenuation is required. Quartz bulk or surface wave acoustic resonators can be used, depending on filter operating frequency. Since only modest levels of signal noise floor reduction (filter stopband attenuation) are usually required, resonator anharmonic spurious resonances do not degrade spectral improvements. Series reactance can be used for fine frequency set-on, and a simple, self contained, TO-5 or TO-8 component temperature controller can adequately maintain the resonator in its turnover temperature range. Circuit modification for electronic tuning allows use of AFC feedback so that the filter tracks the carrier

signal. This eliminates requirements for temperature control, fine frequency set-on, and minimal aging for the resonator. Figure 10 shows a block diagram of tracking filter AFC loop mechanization. Normally the AFC loop bandwidth is on the order of tens of Hz or less and is chosen for adequate tracking under conditions of maximum expected resonator drift rate (due to temperature fluctuation rate, for example). Figure 11 shows an example of the substantial reduction in a 320 MHz signal noise floor level obtainable using a passive SAW resonator filter. The filter of figure 10 was designed with AFC carrier track capability [Ref. 14].

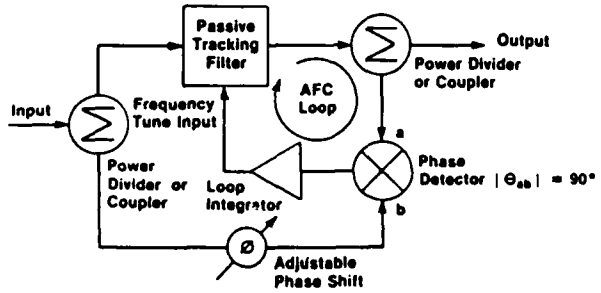



Figure 10. Block Diagram of the Passive, SAWR (Carrier) Tracking Filter

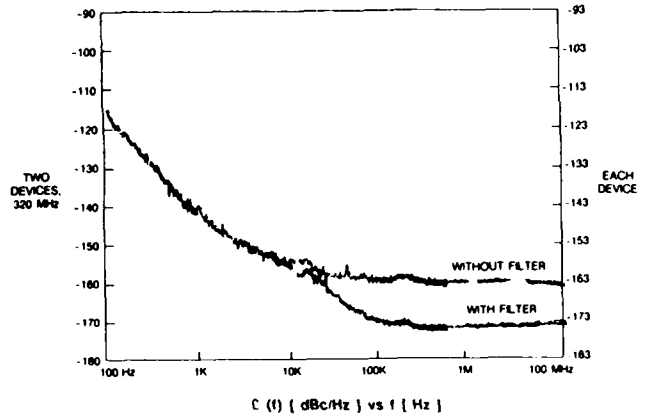



Figure 11. Measured 320 MHz Signal Spectral Improvement Using a SAW Resonator Filter

In cases where narrowest possible filter bandwidth is desired, resonator loading can be minimized (without the usual penalty of high resultant insertion loss) by driving/loading the resonator with low impedance, transistor amplifier stages. Figure 12 shows a simplified schematic diagram for two such mechanizations.

When passive, narrowband filters are used for spectral cleanup, filter passband disymmetry results in FM to AM noise conversion in the carrier signal. This effect can be made negligible by employing post-filter signal amplitude limiting.

## Spectral Improvement Using Phaselock Loops

An alternative, commonly used method for reducing carrier signal noise floor level is to use a phaselock loop. Maximum effectiveness, in terms of carrier signal spectral improvement, requires use of a low noise loop amplifier, low noise phase detector

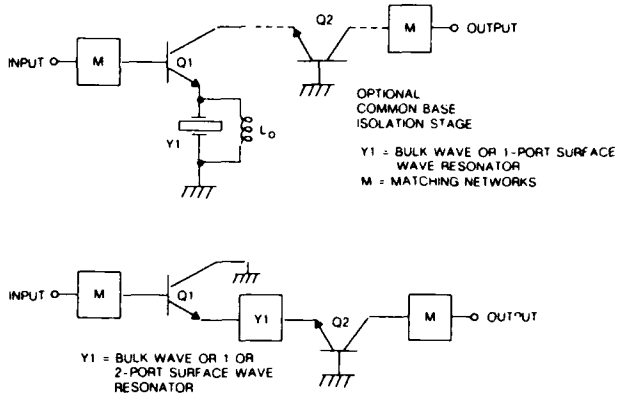



Figure 12. Methods for Reducing Acoustic Resonator Filter Loading and Insertion Loss

(balanced, Schottky diode mixer operated at adequately high RF drive level), and a low noise voltage controlled oscillator (VCO). VCO short-term frequency stability requirements are such that the VCO must utilize a high Q acoustic resonator as does the passive spectral cleanup filter previously described. Figure 13 shows the signal spec..al relationships existing in a phaselocked VCO employed as a spectral cleanup filter. The primary reason that the flicker noise of the free running VCO is poorer than that of the reference signal is that VCO acoustic resonator is higher in frequency, exhibits lower unloaded Q, and is less stable ($S_y(f)$ basis) than the reference signal crystal oscillator resonator.

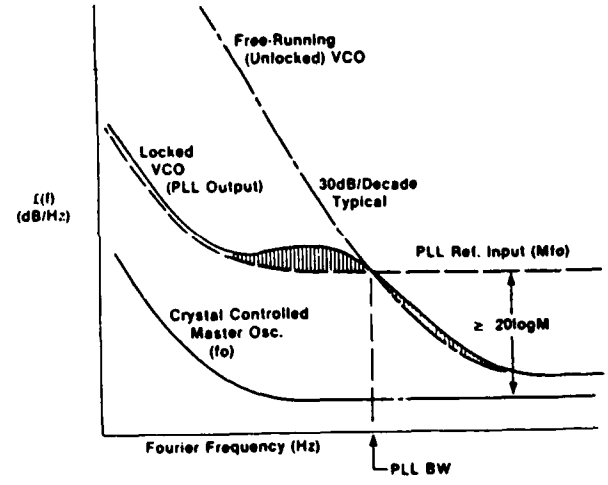



Figure 13. Phase Noise Spectra for PLL Signals

The expression relating the phase of the loop output signal, $\theta_o(s)$ (locked VCO) to the input reference signal, $\theta_i(s)$ and that of the free-running VCO, $\theta_v(s)$ is:

$$\theta_o(s) = \theta_i(s) \times H(s) + \theta_v(s) \times [1 - H(s)] \quad (1)$$

where

$$H(s) = \frac{G(s)}{1 + G(s)} \quad (2)$$

and G(s) is the open loop gain [Ref. 15].

Since the input signal and free running oscillator signal phase noise spectra are uncorrelated, the phase noise spectral density relationship may be written

$$S_{\Phi_o}(\omega) = S_{\Phi_i}(\omega) \mid H(j\omega) \mid^2 + S_{\Phi_v}(1) \mid 1-H(j\omega)^2 \mid \quad (3)$$

In equation (3), $S_{\Phi_i}(\omega)$ includes the phase noise of the phase detector and detector and loop amplifier, referred to the phase detector output (loop amplifier input) terminals.

Examination of equation (3) and figure 13 shows that in order to achieve lowest phase noise output signal, the loop bandwidth should be near the carrier offset frequency where the reference input and free-running VCO phase noise levels are identical. In the ideal case, then, the phase noise spectrum of the loop output signal is identical to that of the input signal at offset frequencies less than the loop bandwidth and is identical to that of the free-running VCO at frequencies in excess of the loop bandwidth.

In the passive tracking filter, AFC loop bandwidth variations are usually non-critical. In the phaselock loop, maintenance of optimum loop bandwidth is important and usually must be measured and set (on a unit-to-unit basis) in accordance with the somewhat variable VCO noise spectra, which also must usually be measured on a unit-to-unit basis. The shaded areas of figure 13 indicate deviations from ideal spectral filtering that can occur as a result of variation in loop bandwidth and shallow (mandatory 20 dB/decade) G(s) response in the vicinity of the loop unity gain frequency. In practice, the degree of spectral improvement obtainable using a phaselocked VCO is usually matched using the passive filter mechanization, which is generally less complex and less costly to produce.

The phase lock loop approach is, however, advantageous in some respects.

The phaselock loop spectral selectivity is very symmetrical, and signal amplitude limiting occurs in the VCO. Hence, FM to AM conversion is not usually a problem. The effective phaselock loop spectral selectivity (bandwidth) is easily adjusted to the optimum value. Also, the phaselock loop can be used to introduce intentional carrier signal modulation.

## Possible Methods for Providing Additional Signal Short Term Stability Improvement

In the generation of low noise signals using bulk and surface acoustic wave frequency control devices, signal near-carrier (flicker of frequency) noise results from short term instability in the resonator itself in oscillators employing low 1/f noise sustaining stage designs.

Lower relative resonator flicker noise levels require use of higher Q devices due to the generally observed relationship between resonator short-term frequency stability and unloaded Q factor.

Additional signal flicker of frequency noise improvement will depend on identification and reduction of the causes of and large variations in resonator 1/f noise. Studies are now underway at Westinghouse and elsewhere to evaluate the effects of resonator bulk defect density, surface quality and electrode metalization on resultant device short term frequency stability.

In connection with these efforts, a novel crystal oscillator circuit has recently been developed at Westinghouse in which it has been possible to actually reduce signal flicker-of-frequency noise to levels below that of the quartz crystal. Figure 14 shows a typical measurement result for two, phase-locked 40MHz oscillators incorporating low noise 5th overtone, AT-cut crystals manufactured by Piezo Crystal Company. The results obtained [£(100 Hz) = -156 dB/Hz per oscillator] verify achievement of excellent spectral performance.

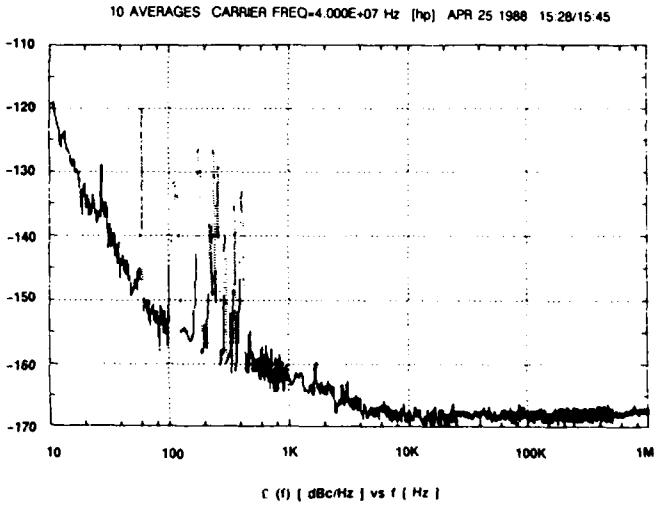



**Figure 14. Measured Phase Noise Spectra for Two, Phase-Locked, 40 MHz Crystal Oscillators.**

Obtaining additional increases in UHF resonator Q (together with low 1/f noise and moderate device drive level capability) is also important in connection with achieving improved oscillator signal spectra and also for use in narrowest possible bandwidth in UHF spectral cleanup filters employed to reduce signal noise floor levels. Use of non-conventional, bulkwave resonators exhibiting significantly higher Q than quartz offers a means of providing additional significant signal spectral improvement.

Figure 15 shows the structure of a High Overtone Bulk Acoustic Resonator (HBAR) currently under development at Westinghouse [Ref. 16].

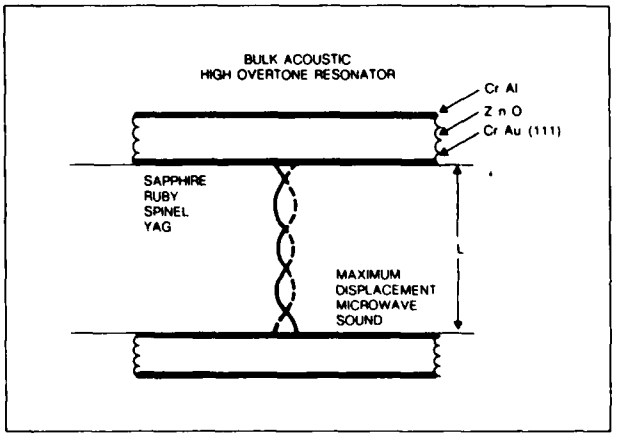



**Figure 15. High Overtone Bulk Acoustic Resonator (HBAR)**

The HBAR is similar to the traditional quartz bulk wave resonator, except for the method of transducing the electrical energy into and out of the resonating crystal. As shown in figure 15, the piezoelectric transducers (sputtered ZnO films) are separate from the resonating crystal. The operation of the resonator is as follows. If opposite surfaces of a crystal are parallel, a resonance response will occur at all frequencies for which the parallel surface separation is an integral multiple of acoustic half wavelengths. The Q displayed by the resonances is a function of crystalline losses, parallelism of the surfaces, and any loading due to the surfaces. Utilization of separate transducers allows injection and extraction of signals at microwave frequencies so the very high-order harmonics of the crystal fundamental resonance can be used.

The unique HBAR transducer has the added advantage that it is no longer necessary that the crystal display piezoelectric properties when excited. Thus the HBAR is not limited to piezoelectric crystals or orientations for which the piezoelectricity is significant. This has opened the way for consideration of many crystals with Q's at microwave frequencies an order of magnitude greater than quartz. These crystals include Yitrium Aluminum Garnet (YAG), Sapphire, Lithium Niobate, Lithium Tantalate and others. This technology results in a resonator with an fQ product at microwave frequencies of $8 \times 10^{13}$, representing the highest available Q in the VHF through microwave frequency range. Another unique and very important device characteristic is its low vibration sensitivity. Prototype resonators have been fabricated exhibiting frequency changes on the order of $1 \times 10^{-11}$ per g. This is 100 times lower than typical values for conventional quartz resonators.

Initial, prototype devices utilized longitudinal wave propagation and exhibited large frequency-temperature coefficient. Efforts are currently underway at Westinghouse to further improve HBAR performance. These include development of tunable, shear mode devices exhibiting zero temperature coefficient at compatible, oven-control temperatures. Also included are studies (similar to those being conducted for conventional quartz resonators) aimed at reducing transducer self-noise levels and increasing device signal handling capability [Ref. 17].

Figure 16 shows the block diagram for a VHF, HBAR - stabilized oscillator recently fabricated at Westinghouse.

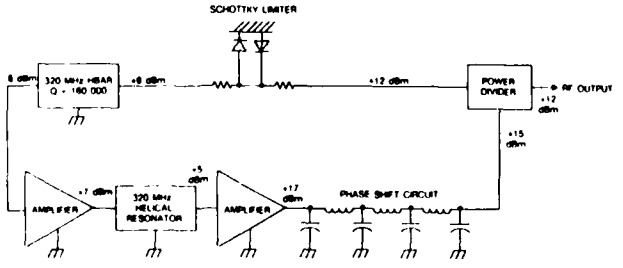
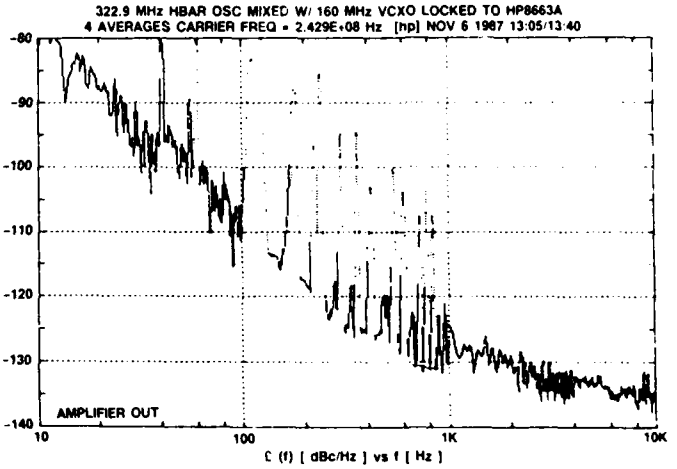



**Figure 16. Block Diagram for Prototype HBAR-Stabilized Oscillator**

The design incorporates use of low flicker noise modular amplifiers and a miniature helical resonator used to select the desired HBAR resonance. The resonator itself exhibits high Q (150,000 at 320 MHz) resonances every 5 MHz and a 50 ohm insertion loss of 13 dB. HBAR (phase bridge) short-term stability measurements for six prototype devices indicated a range of flicker noise levels of $S_y(100Hz) = 1x10^{-25}$ to $4x10^{-24}$. Installation of the lowest noise device in the oscillator however, resulted in $S_y(100\ Hz) = 2x10^{-24}$. A 13 dB increase in resonator flicker noise was confirmed and was likely a result of overheating during installation on the oscillator board.

Nevertheless, the flicker noise performance obtained on a $S_y(f)$ basis, compare favorably with those reported recently for ultra-low phase noise SAW resonator oscillators [Ref. 9], as shown in figure 17.




NOTE: NOISE SPECTRUM FOR $f_m \geq 1$ kHz DUE TO HP8663

**Figure 17. Measured Phase Noise Spectrum for HBAR-Stabilized Oscillator**

## Conclusions

In the range of carrier offset frequencies and for the types of devices of concern in this paper (i.e., fm $\geq$ several Hz), the signal spectral characteristics are divided into two distinct regions. In the near-carrier (flicker of frequency) region, the effects of flicker of phase noise in the oscillator sustaining stage can be made negligible by employing transistor linearization techniques such as emitter degeneration, use of low noise (Schottky diode) amplitude limiting, and through minimal resonator loading (Q degradation). The dominant contributor to signal instability is then due to short-term frequency instability in the resonator itself.

Resonator selection must then be based on known or measured levels of device self noise, using the spectral density of the fractional frequency fluctuations, $S_y(f)$, as a convenient comparative measure of resonator short term stability performance.

At high carrier offset frequencies, various methods are in-hand for improving attainable signal noise floor levels, and have been described in this paper. These include:

1) Highest possible oscillator operating frequency consistent with satisfying signal flicker-of-frequency spectral requirements;

2) Use of sustaining stage transistors having high $f_T$ in an oscillator configuration making use of the resonator out of band impedance to minimize output signal white noise content;

3) Use of (passive and phase lock loop) narrow bandwidth, acoustic resonator, spectral cleanup filters at UHF; and

4) Maintenance of high RF drive in oscillator; frequency multiplier, amplifier and spectral filter circuitry.

In connection with items (1) and (3), use of non-conventional, very high Q acoustic resonators is a promising method for achieving additional signal short term stability improvement.

## REFERENCES


(1) D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," Proc. IEEE, Vol. 54, no. 2, Feb., 1966, pp 329-330.

(2) F. L. Walls and A. E. Wainwright, "Measurement of The Short-Term Stability of Quartz Crystal Resonators and the Implications for Crystal Oscillator Design," IEEE Trans. on I. and M., Vol. IM-24, no. 1, March, 1975, pp 15-20.

(3) J. J. Gagnepain, "Fundamental Noise Studies of Quartz Crystal Resonators," Proc. 30th Frequency Control Symposium, June, 1976, pp 84-87.

(4) J. J. Gagnepain, "A Comparison of Frequency Noise of Quartz Resonators," Proc. 39th Frequency Control Symposium, May, 1985, pp 135-137.

(5) J. J. Gagnepain, et al., "Relation Between 1/f Noise and Q Factor in Quartz Resonators at Room and Low Temperatures, First Theoretical "Interpretation," Proc. 35th Frequency Control Symposium, May, 1981, pp 476-479.

(6) T. E. Parker, "1/f Frequency Fluctuations in Acoustic and Other Stable Oscillators," Proc. 39th Frequency Control Symposium, May, 1985, pp 97-106.

(7) S. Elliott and R. C. Bray, "Direct Phase Noise Measurements of SAW Resonators," Proc. 1984 IEEE Ultrasonics Symposium, Nov., 1984, pp 180-185.

(8) M. M. Driscoll and B. W. Kramer, "Spectral Degradation in VHF Crystal Controlled Oscillators Due to Short Term Instability in the Quartz Resonator," Proc. 1985 IEEE Ultrasonics Symposium, Oct., 1985, pp 340-344.

(9) M. J. Loboda, et al., "Reduction of close-to-carrier Noise in Surface Acoustic Wave Resonators", Proc. 1987 IEEE Ultrasonics Symposium, Oct., 1987, pp 43-46.

(10) D. Halford, et al., "Flicker Noise of Phase in RF Amplifiers and Frequency Multipliers: Characterization, Cause, and Cure," Proc. 22nd Frequency Control Symposium, April, 1968, pp 340-341.

(11) A. VanDerZiel, "Noise in Solid State Devices and Lasers," Proc. IEEE, Vol. 58, no. 8, Aug., 1970, pp 1178-1206.

(12) M.M. Driscoll, "Low Noise VHF Crystal-Controlled Oscillator Utilizing Dual. SC-Cut Resonators," IEEE Trans. on UFFC, Vol. UFFC-33, no. 6, Nov., 1986, pp 698-704.

(13) M. M. Driscoll, "Low Noise Crystal Oscillators Using 50 Ohm, Modular Amplifier Sustaining Stages," Proc. 40th Frequency Control Symposium, May, 1986, pp 329-335.

(14) M. M. Driscoll, "Surface Acoustic Wave Resonator Tracking Filter for UHF Signal Stability Improvement," Proc. 1983 IEEE Ultrasonics Symposium, Nov., 1983, pp 271-276.

(15) V. F. Kroupa, "Frequency Synthesis," J. Wiley and Sons, New York, 1973.

(16) J. T. Haynes, et al., "Stable Microwave Source Using High Overtone Bulk Resonators," 1985 IEEE MTT-S Digest, pp 243-246.

(17) H. L. Salvo, et al., "Shear Mode Transducers for High Q Bulk Microwave Resonators," Proc. 41st Frequency Control Symposium, May, 1987, pp 388-390.

42nd Annual Frequency Control Symposium - 1988

# LOW PHASE NOISE OSCILLATORS ABOVE 1 GHz UTILIZING HERMETICALLY SEALED ALL QUARTZ RESONATORS AND HARMONIC EXTRACTION

C. E. Chase, R. W. Laton, C. L. Yuen

Raytheon Company
Special Microwave Devices Operation
Northborough, MA 01532

## ABSTRACT

A compact hybrid SAW oscillator with an output frequency of 1.48 GHz, low aging rate, and low phase noise was realized by employing the various technologies developed for SAW devices and a unique second harmonic extraction technique.

## INTRODUCTION

SAW resonator stabilized oscillators are compact and capable of very low close-to-carrier phase noise, precise frequency settability, and low aging rates at frequencies considerably higher than bulk crystal oscillators. Recent work at Raytheon's Research Division[1] has demonstrated that a combination of wide-aperture design, use of copper-doped aluminum metallization, and burn-in procedures can improve phase noise performance as much as 15 dB beyond the previously-known state of the art. A few parts per million accuracy in absolute frequency can be achieved by reactive ion etching (RIE). Use of Raytheon's proprietary All-Quartz Package (AQP) technology makes aging rates of less than 1 ppm/year readily achievable[2].

The practical upper frequency limit for fundamental-mode SAW resonators determined by line-width limitations of conventional photolithography is about 1 GHz. Higher frequencies up to X-band and beyond are often needed for radar, ECM, and communications systems; these can be reached by means of frequency multiplication at some cost in phase noise performance and system complexity. The availability of SAW resonator oscillators operating at frequencies above 1 GHz would limit the amount of multiplication required, leading to better performance and smaller, less complex systems. Such oscillators can be conveniently and efficiently realized by means of second-and third-harmonic extraction techniques. Conventional SAW oscillators are based upon a delay line or resonator forming a feedback loop around an amplifier such that excess gain and an integral number of $2\pi$ radians of phase shift are provided at the desired frequency. White noise present in the circuit will then be amplified and reinforced at the desired operating frequency, leading to a buildup of coherent oscillation. As the oscillation amplitude grows to a large-signal condition, the gain will be compressed, and a steady state will be reached in which the net gain has compressed to unity. The resulting nonlinearity in the closed-loop transfer function in the operating region is responsible for relatively strong second, third, and higher-order harmonic content in the steady-state waveform. Measurements of these harmonics reveal levels within 13 - 15 dB of the fundamental, for 2 - 4 dB of gain compression which is typical for practical oscillators. A comparison of second-harmonic extraction with the performance of a typical broad-band frequency doubler is shown in Fig. 1.

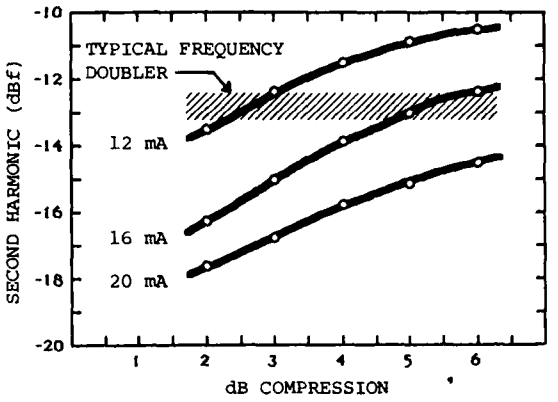



Fig. 1. Second-harmonic generation for typical amplifier compared with typical frequency doubler.

Bandpass filters may be employed to select the desired harmonic for subsequent amplification and delivery to a load. The usual 6 dB per octave degradation of phase noise with increasing center frequency is applicable to the higher harmonics, just as it would be if the fundamental were multiplied in external circuitry. This principle has been demonstrated at output frequencies as high as 2.25 GHz. The upper limit to the usefulness of this technique is under study.

## EXPERIMENTAL RESULTS

A fixed frequency hybrid circuit SAW resonator oscillator operating at the second harmonic of a 740 MHz two-port resonator has been constructed, and is described in this paper. Since the frequency of the resonator was at one-half the oscillator output frequency, the linewidth requirement was significantly relaxed and conventional optical photolithography was used. The SAW resonator was sealed in an AQP to provide good long-term stability[2] and had also been "burned-in" to improve the close-in phase noise performance before it was put into the hybrid oscillator circuit.

A simplified block schematic diagram of the 1.48 GHz hybrid SAW oscillator is shown in Fig. 2. The os-

CH2588-2/88/0000-378 $1.00 C 1988 IEEE

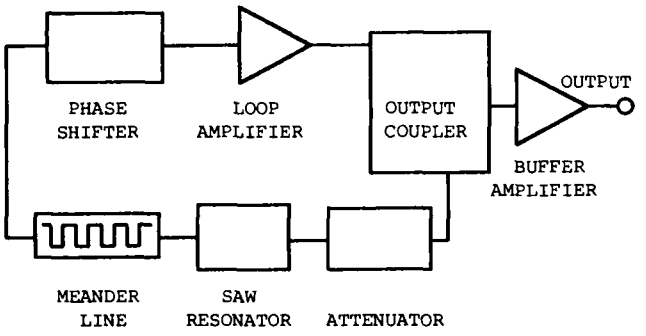



Fig. 2. Simplified block diagram of 1.48 GHz second-harmonic SAW resonator oscillator.

cillator loop contains a single loop amplifier, a stepwise-adjustable phase shifter using binary capacitors, a meander line attenuator, and power splitter. For VCO purposes a varactor-tuned phase shifter could be substituted for the one used here. The combination of phase shifter and meander line is used to adjust the phase shift around the loop to achieve oscillation; the attenuator is adjusted to set the excess small-signal loop gain to about 3 dB. The operating level was chosen on the basis of observed second-harmonic generation by the loop amplifier used. The power splitter is designed for efficient second-harmonic extraction, while coupling as little fundamental-frequency power as possible out of the loop. The level of second harmonic in the loop is typically only about 15 dB below the fundamental. Since this can be extracted with high efficiency, the approach compares favorably with frequency multiplication techniques. In addition, very little (<1 dB) fundamental power is withdrawn from the loop, so that the gain required of the loop amplifier is reduced. This has the effect of lowering the phase noise floor, which varies directly with loop amplifier gain.

The extracted second-harmonic power is amplified to achieve the desired output power, and to buffer the oscillator against frequency pulling by load impedance variations. Additional 740 MHz filtering is used to reduce the fundamental frequency content of the output below -40 dB with respect to the second harmonic; a low-pass filter at the output similarly reduces the third-harmonic component to an acceptable level. The output spectrum of the oscillator is shown in Fig. 3 and measured phase noise in Fig. 4.

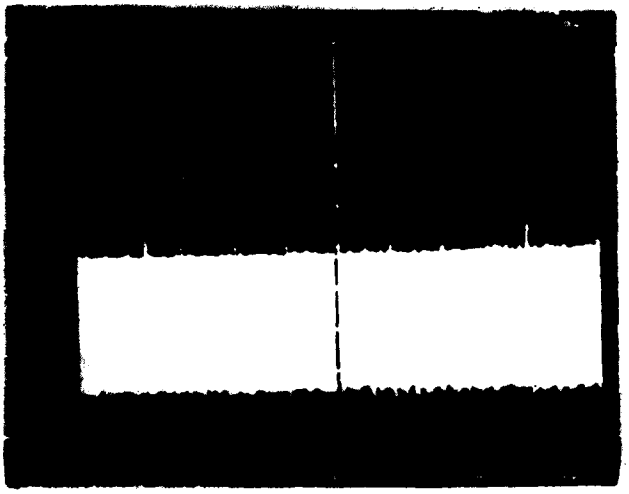


Fig. 3. Output spectrum of second-harmonic oscillator.

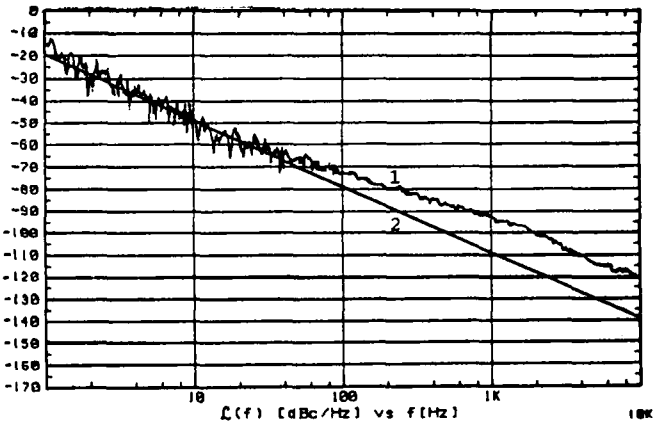



Fig. 4. Close-to-carrier phase noise spectrum.
(1) Noise contribution of measurement system synthesizer (VCO).
(2) Straight-line projection of data between 1 and 100 Hz, representative of unit under test to thermal noise floor at ~ -165 dBc/Hz.

In order to ensure an acceptable window of output power over a wide range of operating temperatures, care must be taken in biasing the loop amplifier correctly and adjusting the excess small-signal gain for harmonic production. The temperature dependence of other components must also be considered. In the present case, the output power is flat within ± 2 dB between -54 and +95°C.

Since the purpose of this investigation was to demonstrate the advantages of harmonic extraction techniques rather than to achieve the lowest possible phase noise, the SAW resonator was not optimized using the full arsenal of noise-reduction techniques described in the first paragraph. Likewise, commercially available MMIC amplifiers that were employed contribute somewhat to phase noise degradation. Nevertheless, the observed phase noise level of -80 dBc/Hz at 100 Hz offset is quite attractive for a 1.5 GHz oscillator. An ongoing optimization of the design is expected to provide eventual performance better than -95 dBc/Hz at 100 Hz offset.

## REFERENCES

1. M. J. Loboda, T. E. Parker, J. A. Greer, and G. K. Montress, "Reduction of Close-to-Carrier Phase Noise in Surface Acoustic Wave Resonators", to be published in the Proceedings of the 1987 Ultrasonics Symposium.

2. T. E. Parker, J. Callerame, and G. K. Montress, "A New All Quartz Package for SAW Resonators", Proceedings of the 39th Annual Symposium on Frequency Control, 1985, pp. 519-525.

42nd Annual Frequency Control Symposium - 1988

# MEASUREMENT and ANALYSIS of THERMAL HYSTERESIS in RESONATORS and TCXO's

Raymond L. Filler
US Army Electronics Technology & Devices Laboratory (LABCOM)
Fort Monmouth, NJ 07703-5000

## SUMMARY

Frequency vs. temperature (f-T) non-repeatability, i.e., thermal hysteresis, is an often encountered, but poorly understood phenomenon. Experience has shown thermal hysteresis to be worst at low temperatures. A group of TCXO's with anomalous thermal behavior prompted a closer look.

The resonators from these "anomalous" TCXO's were temperature cycled apart from the oscillators. Analysis of the resonator f-T data, to account for thermal lag between the resonator and the external thermometer, showed that the inherent hysteresis of the resonator is much lower than that of the TCXO and has a different temperature dependence.

A model was developed that accounts for both the normally-encountered and anomalous thermal hysteresis. This model can separate apparent hysteresis from "true hysteresis". Some hysteresis, which in the past has been attributed to stress and contamination transfer in resonators, may, in fact, be due to improper temperature measurement.

This analysis has the potential for allowing more accurate characterization and specification of TCXO's and resonators, and for prediction of behavior under dynamic temperature conditions.

## INTRODUCTION

The apparent thermal hysteresis of a temperature compensated crystal oscillator (TCXO) is defined, for the purpose of this report, as the absolute value of the difference between the frequency at a given temperature during the increasing temperature portion of a full-cycle temperature run and the frequency at that same temperature during the decreasing-temperature portion of the cycle. In most cases the apparent hysteresis is a function of temperature and of thermal history, including rate-of-change of temperature. Figure 1 is an example of a typical frequency vs. temperature cycle for a TCXO.

The curve consists of f-T data taken from -55°C, to +85°C and then, completing the cycle, to -55°C. The arrows on the curve indicate the direction of the temperature change. The apparent thermal hysteresis for this oscillator is plotted in Figure 2. It can be seen that the hysteresis is largest at the low end of the temperature range and is a minimum near 50°C.

In the course of evaluating the f-T data on a group of TCXO's, the frequency vs. temperature curve shown in Figure 3 was analyzed. (The example discussed in this paper is one of four TCXO's that have similar characteristics).

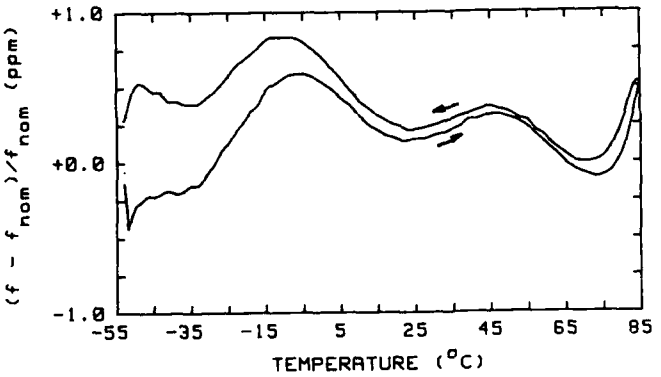



Figure 1 - Typical f-T data for a TCXO.

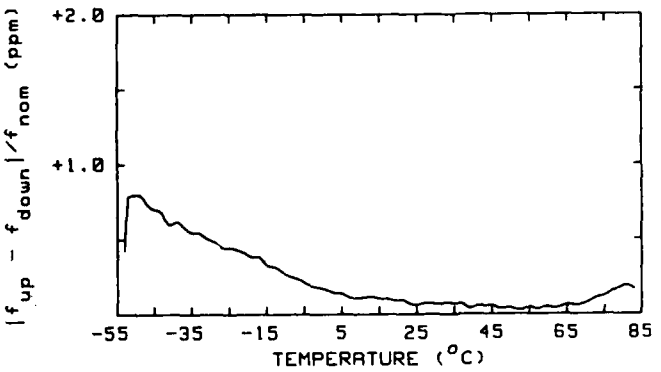



Figure 2 - Apparent thermal hysteresis for the TCXO from Fig. 1.

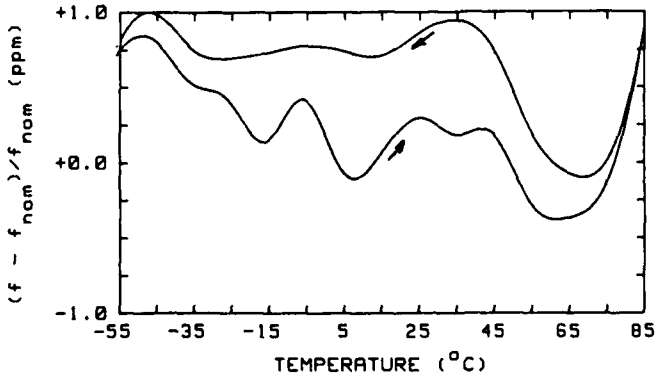



Figure 3 - "Anomalous" TCXO f-T data.

The apparent thermal hysteresis for this oscillator is plotted in Figure 4. The thermal hysteresis in this oscillator is maximum near the middle of the temperature range and is zero near the upper and lower temperature extremes. This is counter to the common experience that the greatest hysteresis occurs at low temperatures.

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT 

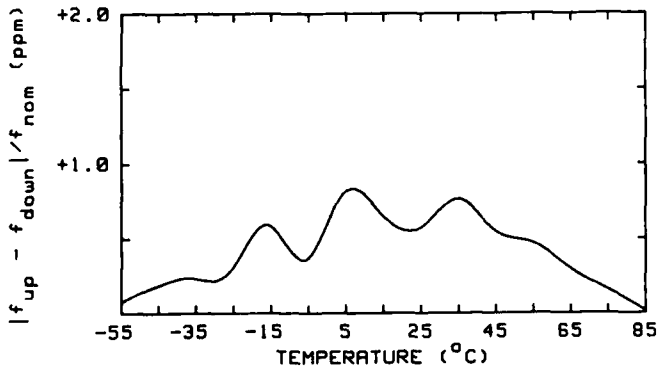



Figure 4 - "Apparent" hysteresis for the TCXO in Fig. 3.

The cause of the anomalous hysteresis was postulated to be thermally induced stress on the quartz by the chrome-gold electrodes that were deposited on the resonator. The resonator was removed from the oscillator to separate the hysteresis contribution of the resonator from the hysteresis contribution of the circuitry.

## RESONATOR f-T MEASUREMENT

Four f-T cycles were performed on the resonator using a Pi-network and vector voltmeter-synthesizer system to measure the *frequency*.[2] The resonator, without a thermal sink or baffling, and a quartz thermometer were in close proximity in a temperature chamber. For each temperature cycle, the temperature chamber was "cold soaked" at -55°C for 30 minutes, then stepped in 2°C increments. The differences among the four runs were either the rate of change of temperature or the upper limit of the temperature excursion. The temperature cycles are designated "fast", "slow", "slow-85", and "slow-50". During the "fast" run, the soak duration between when the temperature was set and when the measurement was made was 2 minutes. During the "slow" runs, the soak duration was 4 minutes. In the run designated "slow-50" the upper temperature was 50°C; it was 85°C for the others. Figures 5-8 are the raw f-T data for the four temperature cycles.

## RESIDUALS

Figures 9-12 are the residuals from a least squares fit of the composite of the up and down f-T data to a 6th order polynomial. Data points outside of the specified temperature range were not used in the fit and are denoted by "*". A residual curve for a resonator is similar to an f-T curve for a TCXO in the sense that a function is subtracted from the resonator f-T characteristic. Therefore Figures 9-12 can be directly compared to figures 1 and 3.

## APPARENT HYSTERESIS CURVE

Figures 13-16 are the apparent hysteresis curves for the four cycles. It is the absolute value of the difference between the up and down portions of the residual or f-T curves.

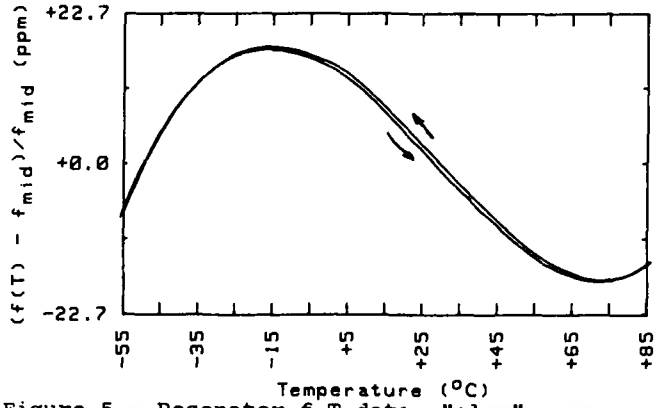



Figure 5 - Resonator f-T data, "slow" run, 2°C/4 minutes.

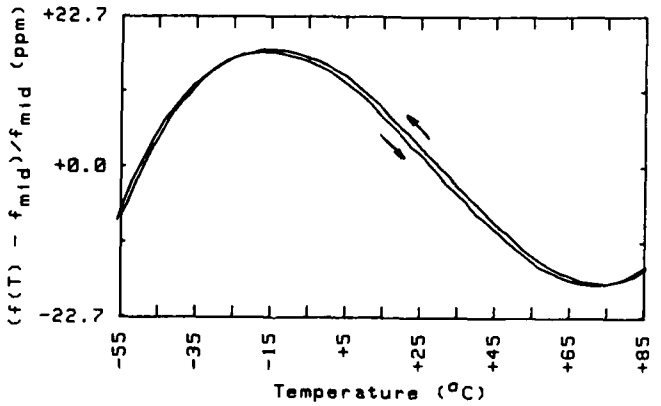



Figure 6 - Resonator f-T data, "fast" run, 2°C/2 minutes.

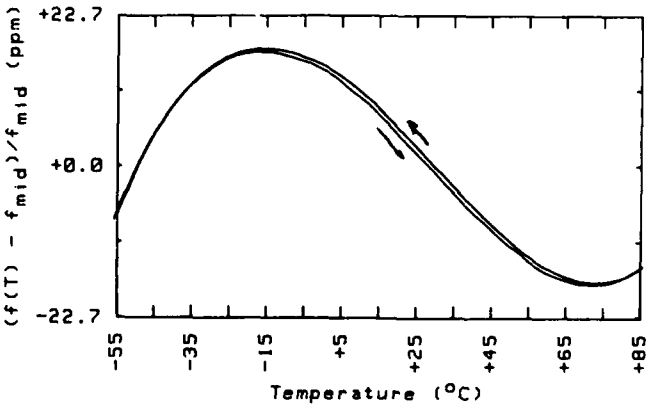



Figure 7 - Resonator f-T data, "slow-85"; 2°C/4 minutes.

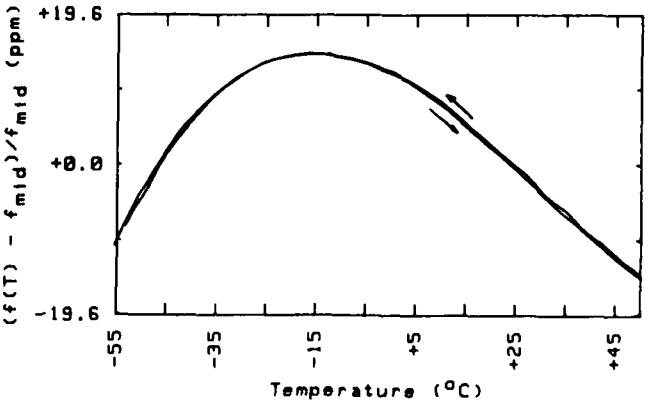



Figure 8 - Resonator f-T data, "slow-50", 2°C/4 minutes.

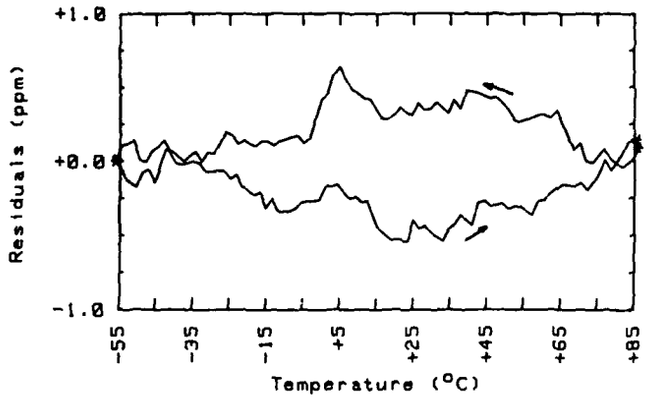



Figure 9 - Residuals from a 6th order polynomial for the "slow" run.

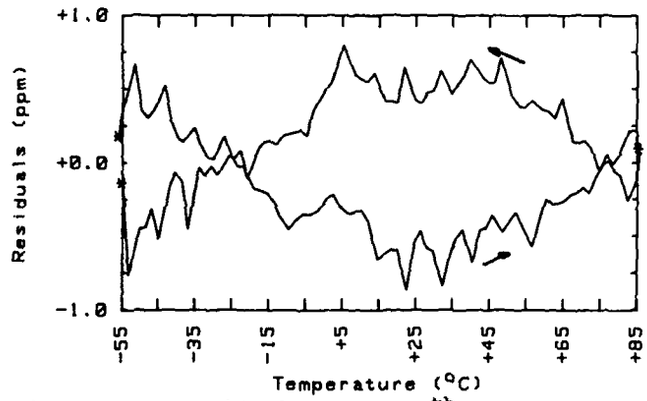



Figure 10 - Residuals from a 6th order polynomial for the "fast" run.

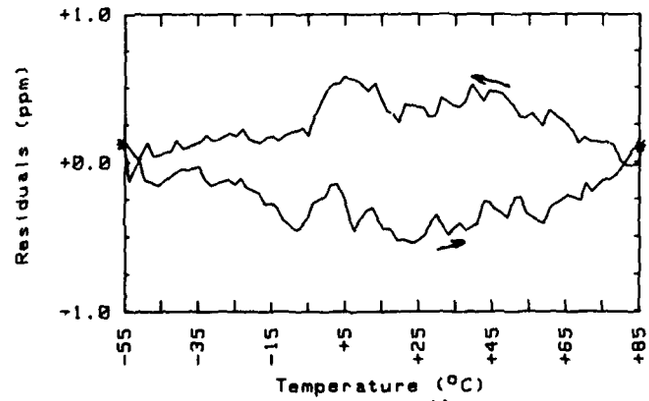



Figure 11 - Residuals from a 6th order polynomial for the "slow-85" run.

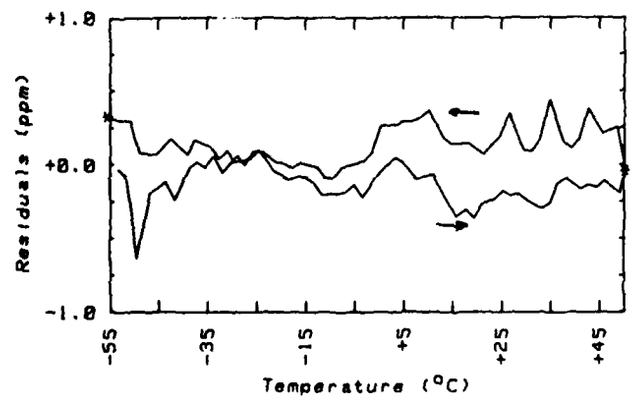



Figure 12 - Residuals from a 6th order polynomial for the "slow-50" run.

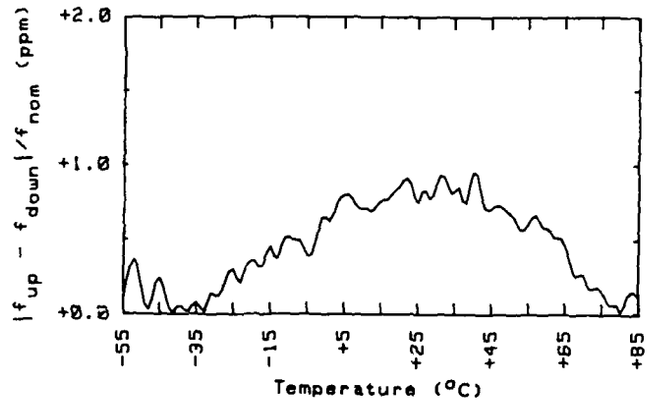



Figure 13 - Hysteresis for the "slow" run.

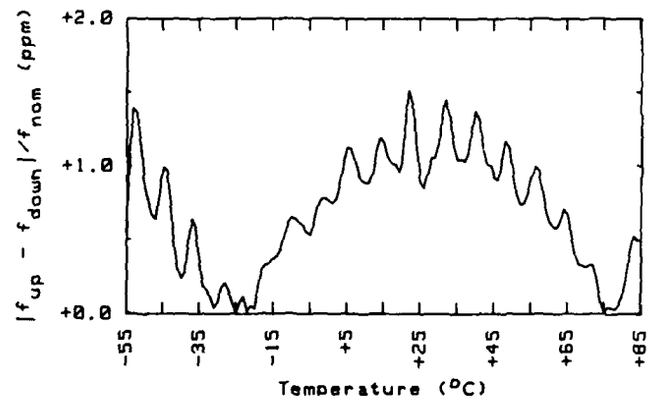



Figure 14 - Hysteresis for the "fast" run.

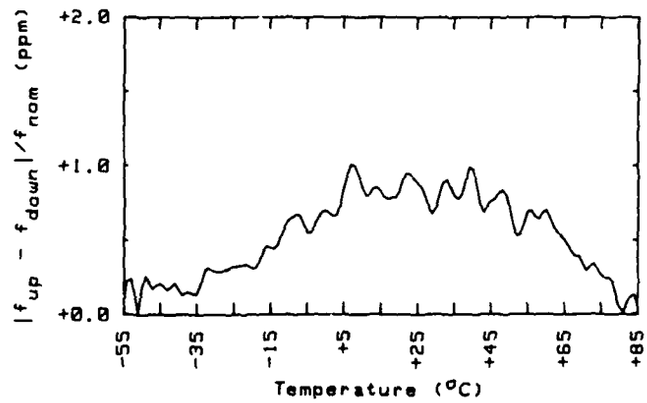



Figure 15 - Hysteresis for the "slow-85" run.

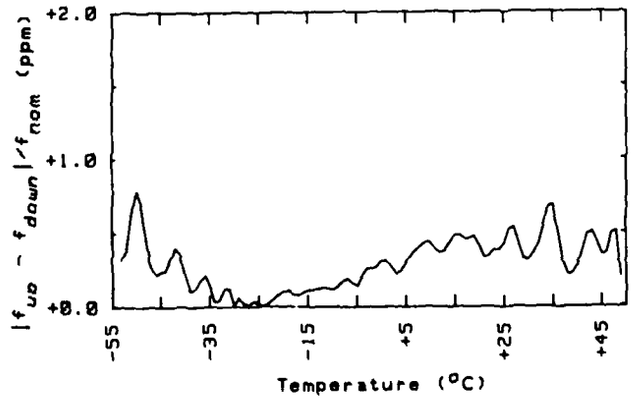



Figure 16 - Hysteresis for the "slow-50" run.

The salient feature of the resonator apparent hysteresis curves, especially the ones that cover the complete -55°C to + 85° cycle, is that they are similar in shape to the "anomalous" TCXO hysteresis curve. That is, the hysteresis is greater near the inflection temperature than near the ends of the temperature range. However, the details of the curves are different. Significantly, the maximum hysteresis (which is near the inflection temperature) is greatest for the "fast" run and least for the "slow-50" run. Furthermore, the temperature at which the hysteresis goes to zero varies with the rate of change of temperature. The low temperature "zero point" is near -35°C for the "slow" run and near -25°C for the "fast" run. The two runs, "slow" and "slow-85", which were performed under identical conditions albeit separated by several months and several temperature cycles are identical. This demonstrates that hysteresis is reproducible.

## TEMPERATURE SHIFT

In Tables I-IV are the inflection temperature ($T_i$), upper turnover temperature (UTP), and lower turnover temperature (LTP) for each set of data from figures 5-8 under the column labeled "composite". This is in some sense an average value for the temperatures. The other two columns are the differences between the characteristic temperatures of the composite and those of the temperature increasing portion of the temperature cycle (labeled "Δ Up"), and the decreasing portion of the cycle (labeled "Δ Down").

| | Composite | Δ Up | Δ Down |
|---|---|---|---|
| LTP = | -16.16 | -0.33 | +0.34 |
| $T_i$ = | +28.71 | -0.42 | +0.43 |
| UTP = | +73.59 | -0.51 | +0.52 |

Table 1 - Characteristic temperatures for the "slow" run.

| | Composite | Δ Up | Δ Down |
|---|---|---|---|
| LTP = | -16.16 | -0.65 | +0.69 |
| $T_i$ = | +28.69 | -0.75 | +0.77 |
| UTP = | +73.55 | -0.85 | +0.85 |

Table 2 - Characteristic temperatures for the "fast" run.

| | Composite | Δ Up | Δ Down |
|---|---|---|---|
| LTP = | -16.22 | -0.27 | +0.28 |
| $T_i$ = | +28.64 | -0.37 | +0.39 |
| UTP = | +73.50 | -0.48 | +0.50 |

Table 3 - Characteristic temperatures for the "slow-85" run.

| | Composite | Δ Up | Δ Down |
|---|---|---|---|
| LTP = | -15.45 | -0.29 | +0.32 |
| $T_i$ = | +31.27 | -0.37 | +0.47 |
| UTP = | +77.99 | -0.45 | +0.62 |

Table 4 - Characteristic temperatures for the "slow-50" run.

It can be seen that the measured $T_i$, UTP, and LTP for the up half-cycle data are lower than the corresponding temperatures for the down half-cycle for all four runs. This is indicative of a thermal lag existing between the thermometer and the resonator with the thermometer lagging behind the resonator.

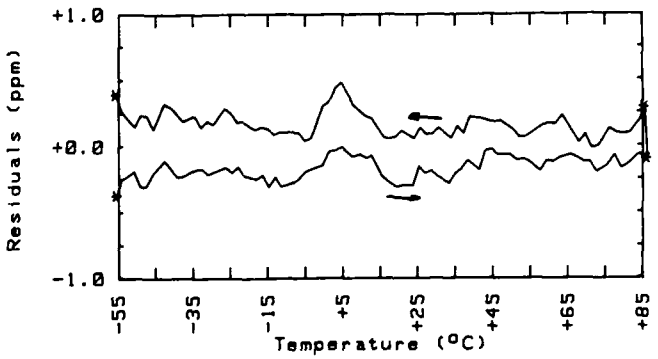



Figure 17 - Residuals from a 6th order polynomial after temperature shifting for the "slow" run.

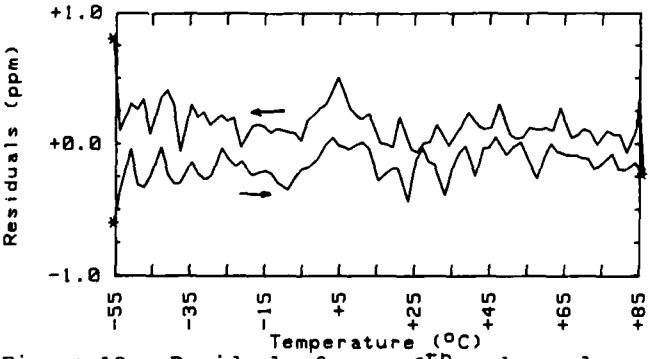



Figure 18 - Residuals from a 6th order polynomial after temperature shifting for the "fast" run.

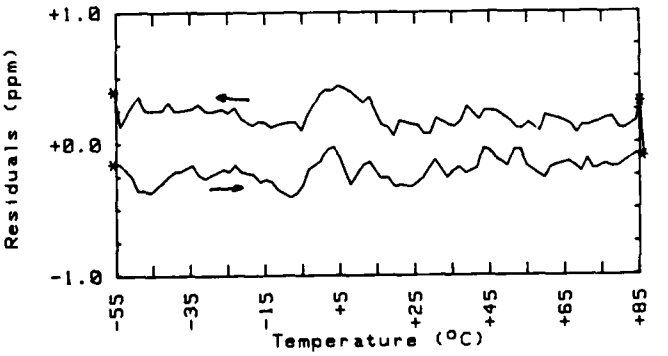



Figure 19 - Residuals from a 6th order polynomial after temperature shifting for the "slow-85" run.

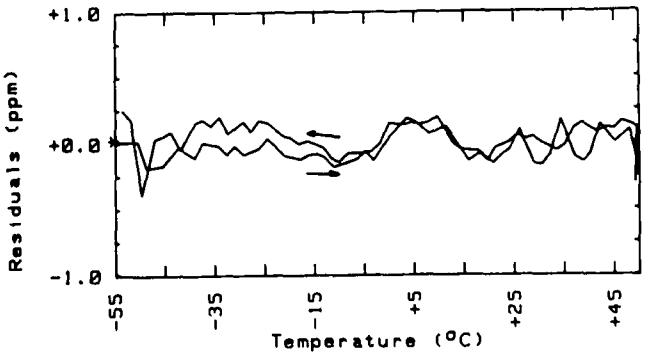



Figure 20 - Residuals from a 6th order polynomial after temperature shifting for the "slow-50" run.

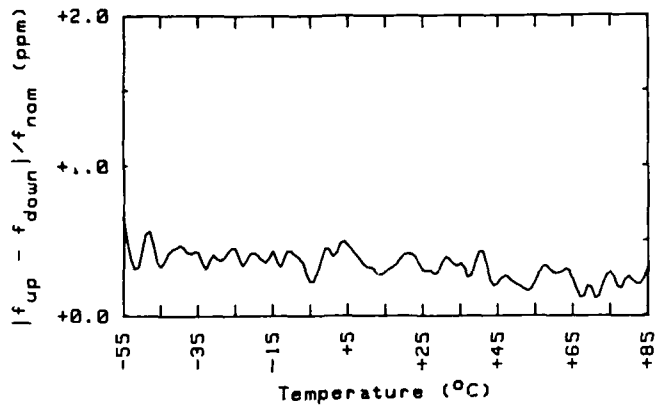



Figure 21 - Hysteresis after shifting for the "slow" run.

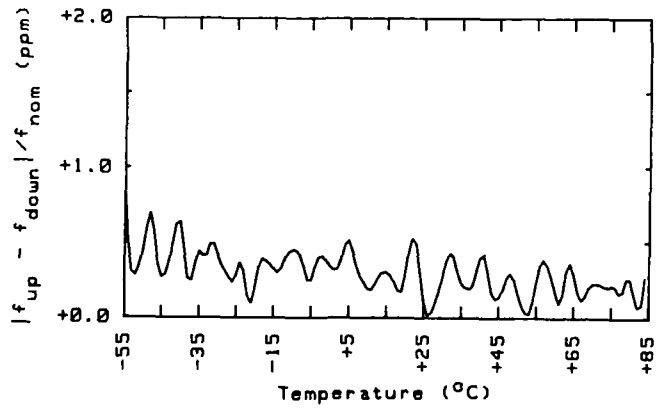



Figure 22 - Hysteresis after shifting for the "fast" run.

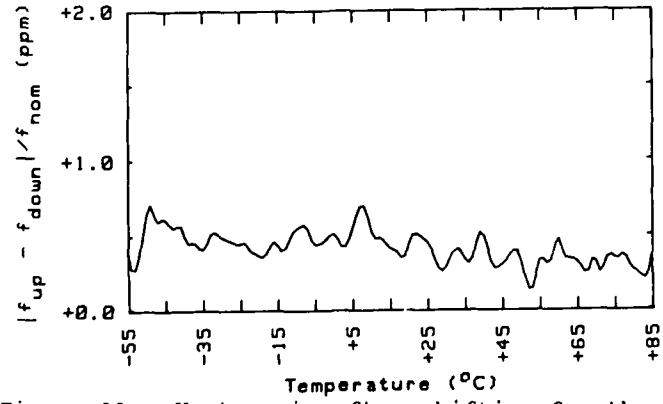



Figure 23 - Hysteresis after shifting for the "slow-85" run.

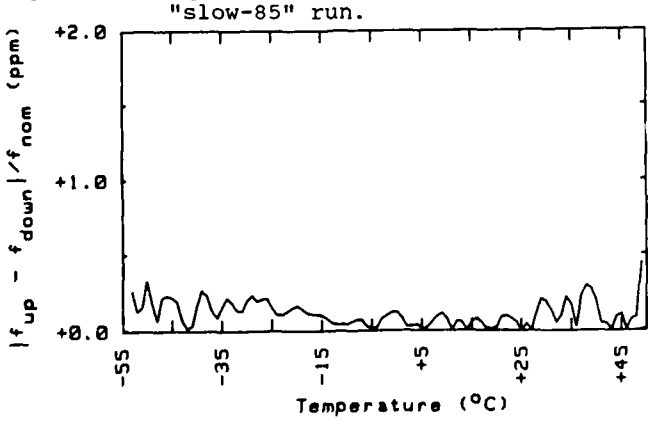



Figure 24 - Hysteresis after shifting for the "slow-50" run.

In order to "compensate" for this thermal lag, a temperature dependent $\Delta T_{up}$ was added to the up half-cycle data and a similar $\Delta T_{down}$ was subtracted from the down half-cycle data. The value of $\Delta T$ for each data point was selected by fitting the temperature differences in each column to a quadratic function. The end result is that the entire up-curve was shifted to the right and the entire down-curve was shifted to the left by the amount needed to make the three characteristic temperatures coincide.

The shifted residual curves for the four runs are shown in figures 17-21 and the shifted hysteresis curves are shown in figures 22-26.

The features of the shifted curves differ markedly from the unshifted curves. For the shifted data the hysteresis is virtually constant. It is important to note that the hysteresis for the temperature shifted data appears to be independent of rate of change of temperature. This can be seen in Figure 25 which is a superposition of the hysteresis curves from the shifted data for the "slow" and the "fast" run which are, to the limit of the resolution of the measurement, identical. The hysteresis obtained after compensating for thermal lag is the "true" hysteresis of the resonator.

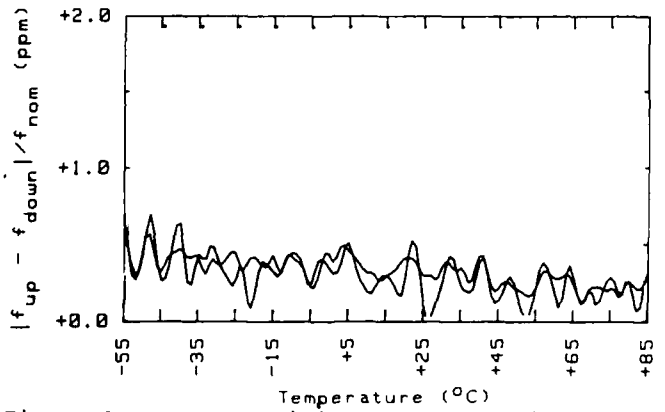



Figure 25 - Superposition of hysteresis curves for the shifted "slow" and "fast" runs.

The shifted and unshifted curves always cross at the turnover temperatures. This is due to the geometrical fact that the sign of the slope changes at a turnover and a horizontal shift causes the thermal lag contribution to the apparent hysteresis to change sign at each turnover point.

## SIMULATION

Since the uncompensated resonator f-T data is unavailable in TCXO's, in order to study the interaction between thermal hysteresis and thermal lag, it was necessary to model the f-T behavior and simulate the effects on the computer. The baseline resonator f-T curve is the cubic obtained from the resonator shown in figure 5.

### Thermal lag

Thermal lag was introduced by assuming one of two conditions for the relative values of the thermal time constant of the resonator,

and the thermal time constant of the thermometer, $\tau_r$ and $\tau_{th}$, respectively.

Condition 1 : $\tau_r >> \tau_{th}$; $\tau_{th} \sim 0$
Condition 2 : $\tau_{th} >> \tau_r$ ; $\tau_r \sim 0$

In condition 1 the resonator lags the thermometer while the thermometer follows the chamber. In condition 2 the thermometer lags the resonator while the resonator follows the chamber. The temperature at each measurement time $t_i$ for the resonator and the thermometer is therefore,[1]

Condition 1:

$$T_r(t_i) = T_i + (T_i - T(t_i-1)e^{-t_s/\tau_r}$$

$$T_{th}(t_i) = T_i$$

Condition 2:

$$T_r(t_i) = T_i$$

$$T_{th}(t_i) = T_i + (T_i - T(t_i-1)e^{-t_s/\tau_{th}}$$

where $T_i$ is the set temperature at each measurement time. For both cases the thermal lag, i.e., the difference between $T_r$ and $T_{th}$, reaches a steady value after about 2-3 soak durations. The soak duration and thermal time constant need to be expressed only in terms of their ratio.

Hysteresis

Hysteresis was included by adding (or subtracting) a constant frequency offset to the calculated frequencies for the down half-cycle.

Simulation Results

Figures 26 and 27 are the simulated "slow" f-T cycle residual and hysteresis curves. The ratio of the soak duration to the thermal time constant is 1.6 with the thermometer lagging the resonator. The frequency offset added to the down f-T is +0.3 ppm. Figures 28 and 29 are the simulation of the "fast" f-T. The soak-duration-to-time-constant ratio is 1.2.

The same shift operation as was performed on the real data was performed on the simulated data. The results are shown in Figures 30 and 31 for the simulated "slow" cycle and Figures 32 and 33 for the simulated "fast" cycle. The hysteresis is approximately the +0.3 ppm input into the simulation. The odd shape of the left part of the residual curves is due to the thermal model, which assumes an initial lag of zero which then grows to a steady state value until the reversal at the upper temperature end.

## SIMULATION vs. REAL

Figures 34 and 35 are superpositions of the simulated and actual data for the "slow" cycle, both unshifted and shifted. Figures 36 and 37 are for the "fast" cycle. It is remarkable how well the simulation matches the real data. The flare at the low temperature end, which is predicted from the simulation, suggests that the long-held belief that hysteresis is worse at the low temperature end may, in fact, be due to thermal lag.

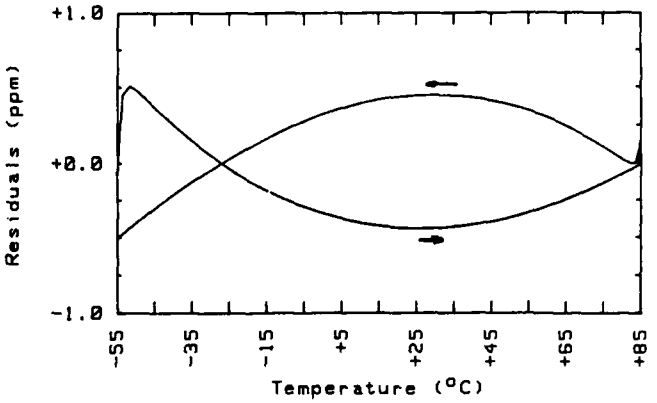



Figure 26 - Residuals from the simulation of the "slow" run.

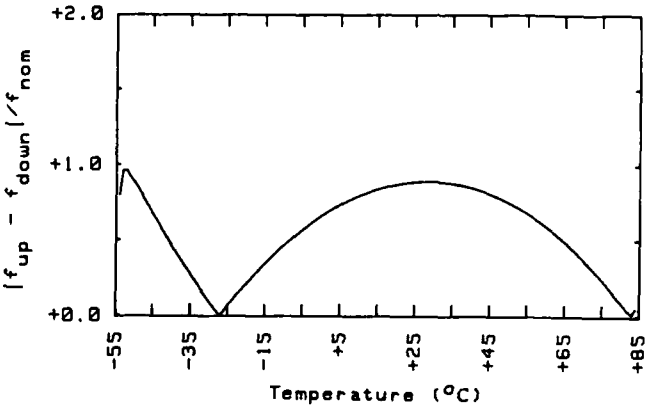



Figure 27 - Hysteresis from the simulation of the "slow" run.

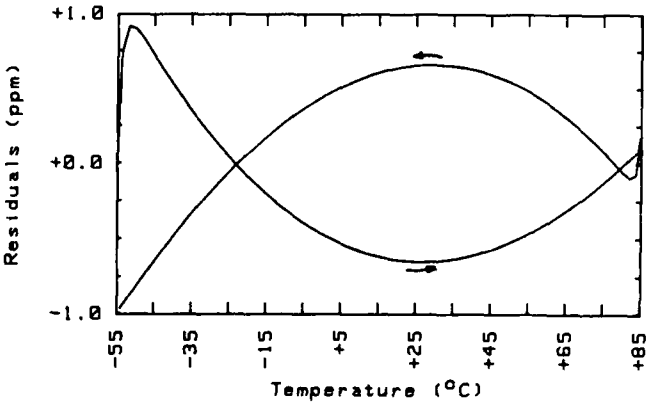



Figure 28 - Residuals from the simulation of the "fast" run.

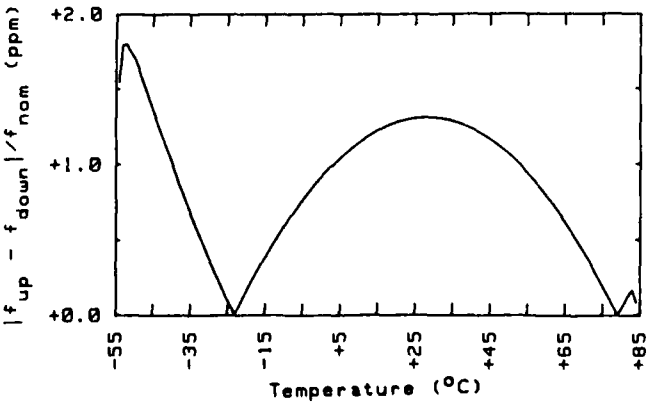



Figure 29 - Hysteresis from the simulation of the "fast" run.

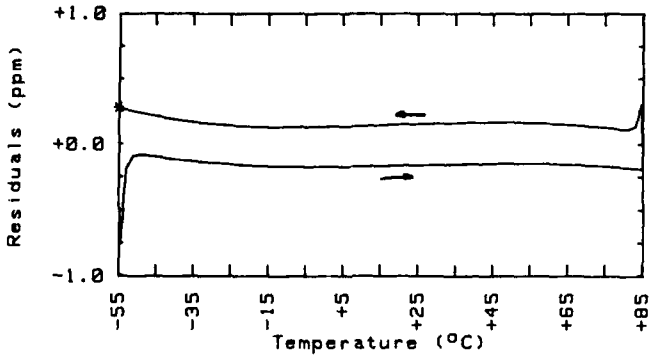



Figure 30 - Residuals from the simulation of the "slow" run after shifting.

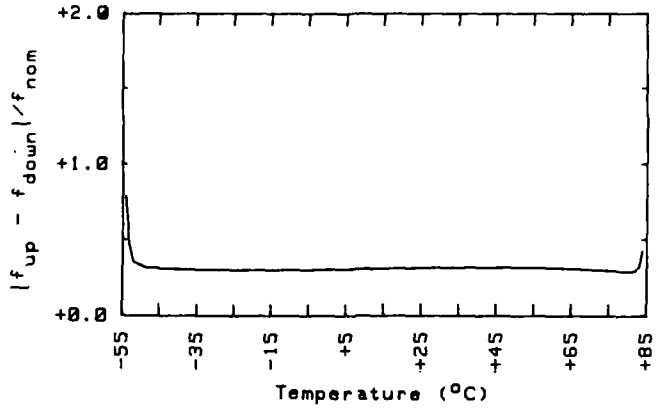



Figure 31 - Hysteresis from the simulation of the "slow" run after shifting.

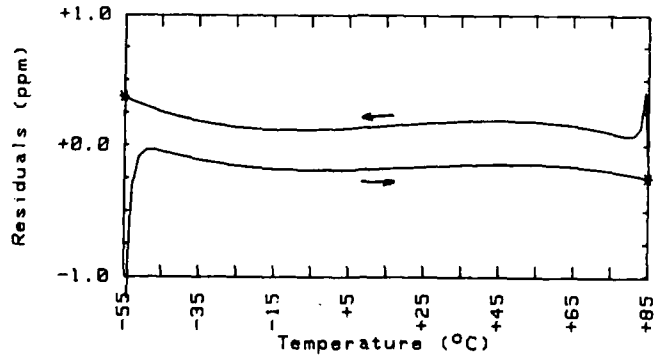



Figure 32 - Residuals from the simulation of the "fast" run after shifting.

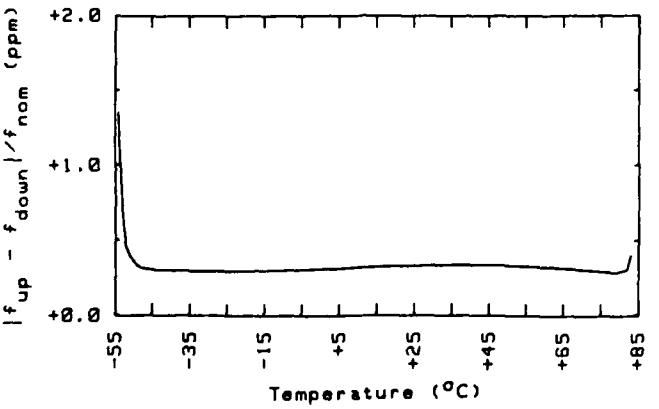



Figure 33 - Hysteresis from the simulation of the "fast" run after shifting.

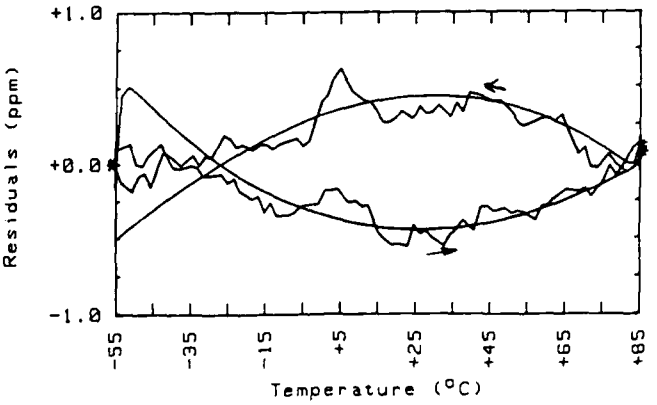



Figure 34 - Superposition of simulated and actual residuals for the "slow" run.

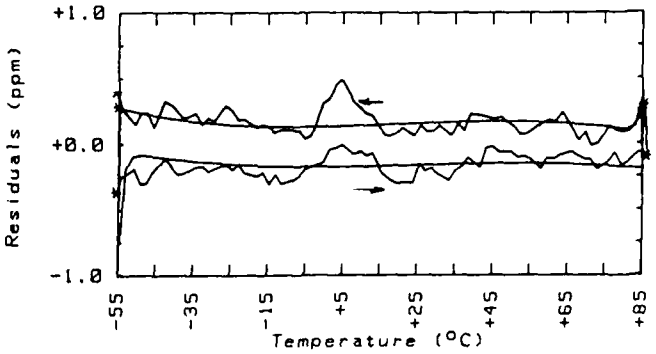



Figure 35 - Superposition of the simulated and actual residuals after shifting for the "slow" run.

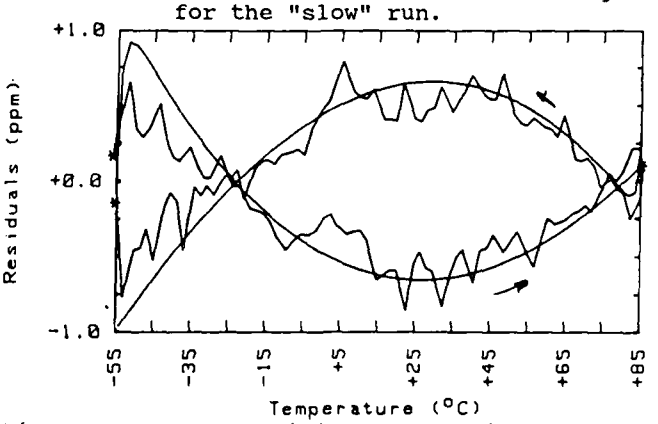



Figure 36 - Superposition of the simulated and actual residuals for the "fast" run.

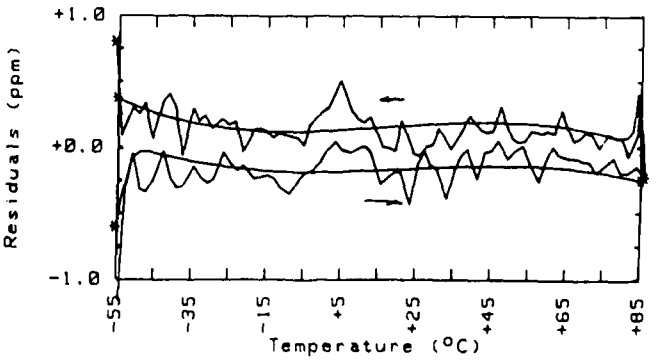



Figure 37 - Superposition of the simulated and actual residuals after shifting for the "fast" run.

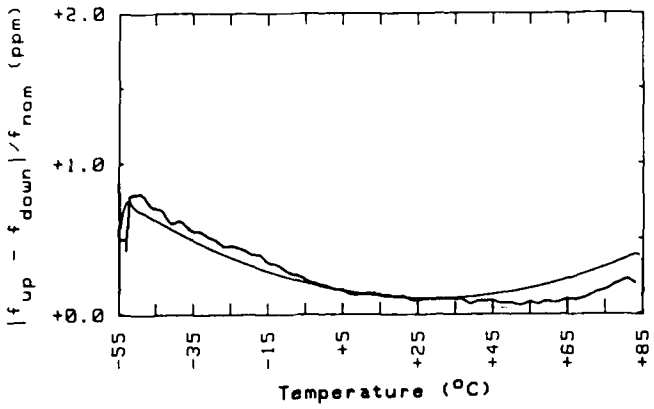



Figure 38 - Superposition of the simulated and actual hysteresis for the "typical" TCXO of Fig. 1.

"Typical" TCXO

The typical TCXO from Figure 1 was simulated with a hysteresis of +0.3 ppm and a thermal time constant ratio of 2.4 with the resonator lagging the thermometer. Figure 38 shows the actual hysteresis data from Fig.2 superimposed on the simulation.

OTHER THERMAL CONDITIONS

Various combinations of thermal lag and hysteresis can give similar looking residual and hysteresis curves. Figure 39 is a group of curves all with the same absolute value of hysteresis (0.3 ppm) and thermal time constant (2.4) and all four permutations of the sign of the hysteresis and lead-lag conditions.

Figure 40 show 3 curves with the same hysteresis (+0.3 ppm) but increasing rate of temperature change, i.e., increasing thermal lag. The crossover phenomena shown in Fig. 40 can occur in 2 ways. Figure 41 shows two thermal lag conditions. It can be seen that both conditions can give rise to crossover. In both cases the hysteresis is +0.3 ppm. The two conditions can be distinguished by the direction of the arrows and the fact that the crossover occurs outside of the TP's for the left case and inside of the TP's for the right case.

CONCLUSION

The similarity between simulated and actual data leads one to postulate that real (or apparent) thermal lag is a major contributor to thermal hysteresis in TCXO's. Apparent thermal lag is an error in the thermometer due to some effect other than temperature, such as stress in a thermistor. Thermal lag induced errors may have occurred during the compensation process due to a different thermal gradient or they may occur during use.

Thermal lag problems must be minimized in order to achieve better performance. The resonator and thermometer must be very well coupled thermally to each other and to any other thermally sensitive components in the oscillator. In addition, the mechanical configuration should be such as to ensure that those elements have the same thermal path to the ambient. Ideally they should be shielded from rapid temperature swings by placing them near the center of the device.

The result that thermal hysteresis is independent of rate of change of temperature, when thermal lag is accounted for, may give some new insight into the causes of the effect and a new technique for measuring and describing hysteresis.

REFERENCES

1. S. Schodowski, "Temperature Performance Measurement Methods for Temperature-Compensated Quartz Oscillators," Research and Development Technical Report ECOM-2896, National Technical Information Service, 1967.

2. International Electrotechnical Commission Document 444; "Basic Method for the Measurement of Resonance Frequency and Equivalent Resistance of Quartz Crystal Units by Zero Phase Technique in a Pi-network."

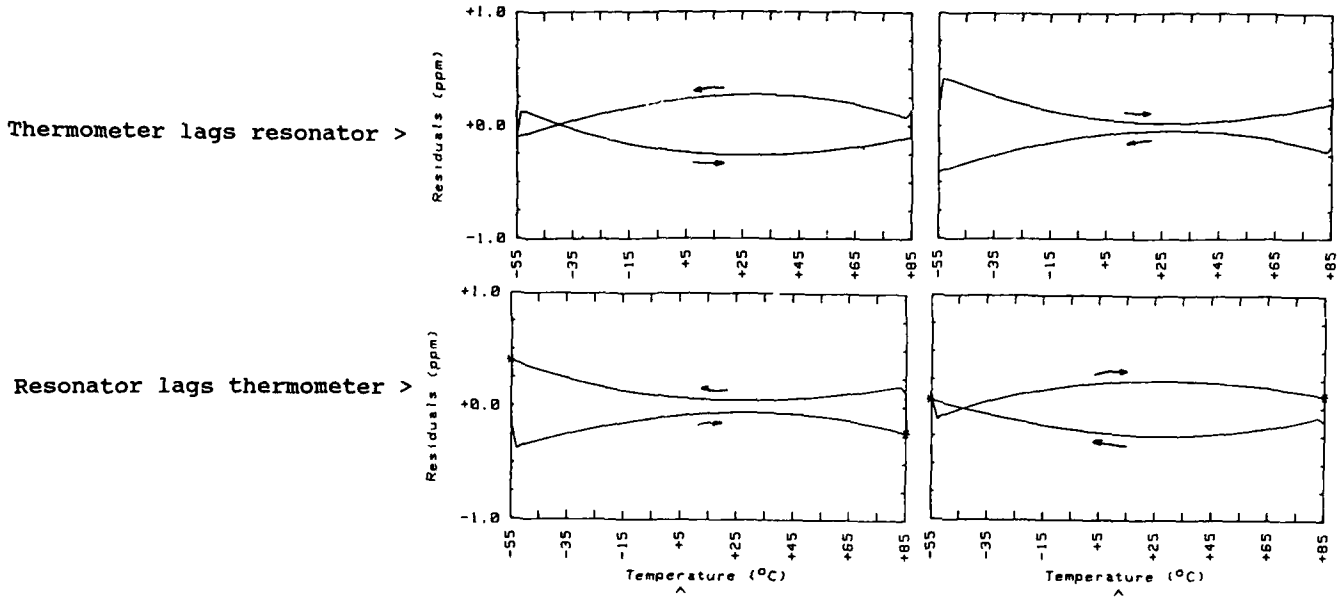



Figure 39 - Residuals from simulation with four conditions of hysteresis and thermal lag.

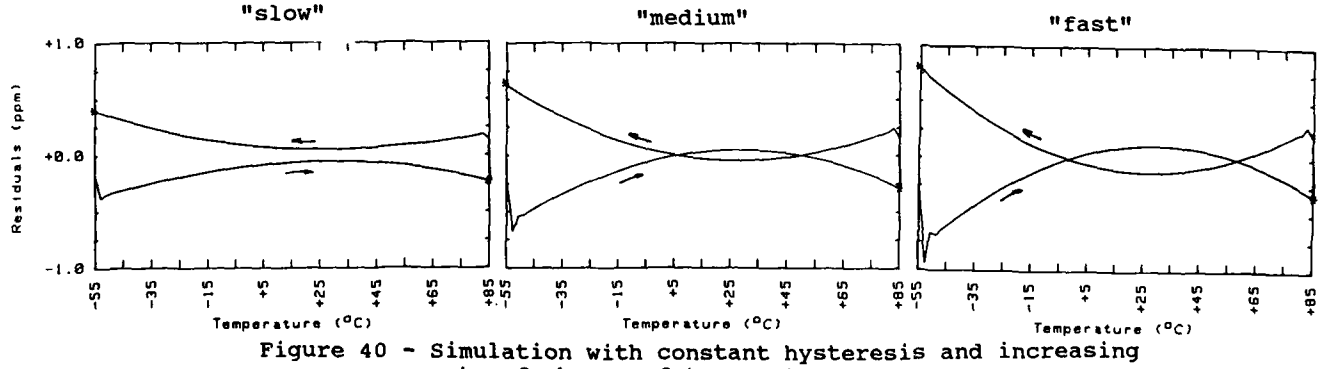



Figure 40 - Simulation with constant hysteresis and increasing rate of change of temperature.

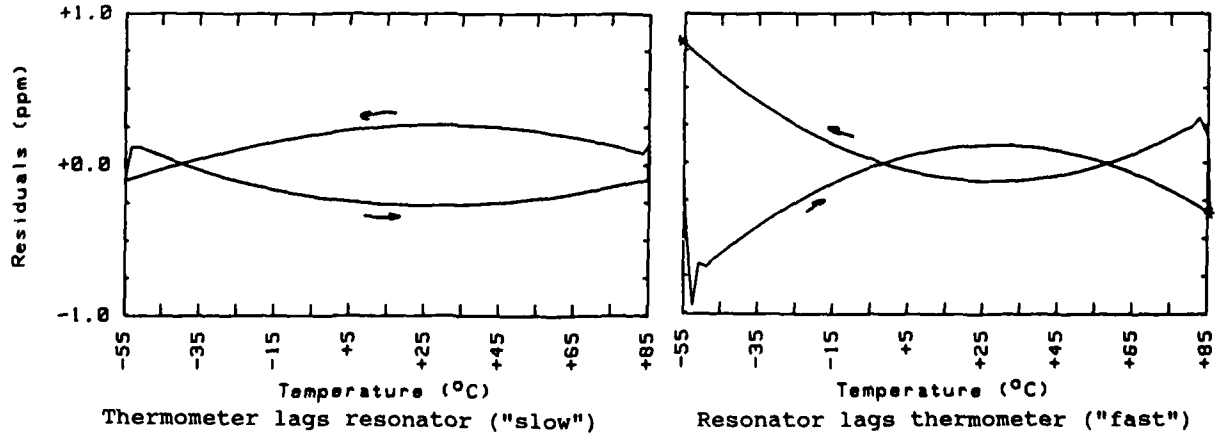



Figure 41 - Crossover caused by two opposite thermal lag conditions.

42nd Annual Frequency Control Symposium - 1988

# Burst Noise and 1/F Noise in Quartz Crystals and Oscillators

by: Grant E. Moulton, Hewlett-Packard Company
Rohnert Park, CA

## Summary

Short term frequency changes similar to the chaotic burst noise or popcorn noise observed in operational amplifier integrated circuits have also been seen in both quartz crystals and quartz crystal oscillators. This noise is distinct in character from the 1/f noise also observed in crystals. Two methods of measuring the burst frequency noise and phase noise of crystals are described and their relative advantages compared. Observations of the frequency changes and phase noise of crystals and crystal oscillators are reported, including comparison of burst frequency noise at both the fundamental and overtones simultaneously. Efforts to find correlation between burst frequency noise and other crystal parameters are reported. The previously reported relationship of S f(1 Hz), the spectral density of frequency fluctuations at 1 Hz offset, and $1/Q^4$ is examined in light of the observed burst frequency noise.

## Background

During the design and production of the Hewlett-Packard 70900A Local Oscillator Module and 70310A Precision Frequency Reference Module, several 100 MHz crystal oscillators had significantly higher phase noise than normally seen, particularly at less than several hundred Hz offset frequency. (See figure 1)

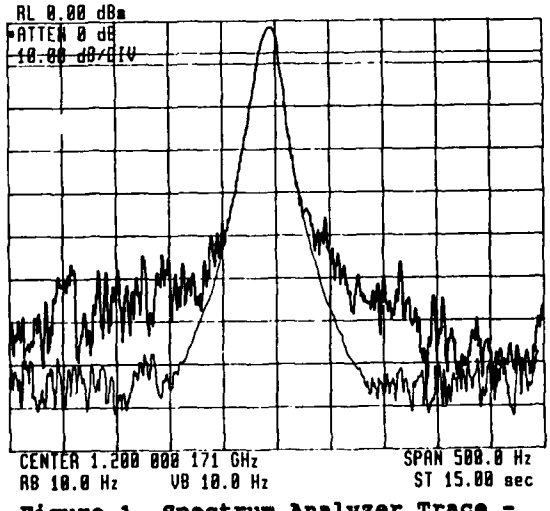



**Figure 1 Spectrum Analyzer Trace - Quiet & Noisy Crystals**

Other observations of crystal oscillators indicated problems with repeatability of phase noise results from one measurement to the next. Some oscillators have shown as much as 20 dB variation between successive tests of phase noise. Similar observations led Michael C. Fischer to investigate the fluctuations of the phase noise of a 100 MHz crystal oscillator over time. He noticed the same variations even when averaging 64 traces over a two minute interval and plotting the average noise over long time periods[1].

While measuring the phase noise of crystal oscillators with the Hewlett-Packard 3047 Phase Noise Measurement System, the time domain phase signal at the input of the Hewlett-Packard 3582A FFT spectrum analyzer bursts were seen by Christian K. Seiracki[2] while investigating the phase noise caused by a 20 MHz crystal filter, and also reported by R. Baer et al[3]. while observing SAW filters. Based upon these observations, a burst noise test fixture was developed by Val D. Mc Omber and Thomas B. Fetter to further investigate and screen the crystals for use in quiet reference oscillators.

## Measurement Methods

Several methods[4,5,6,7,8,9,10] have been proposed to measure the phase noise of crystals both as crystal resonators alone and in oscillators. These methods include two basic configurations: phase locking two crystal oscillators while observing the phase output of the phase detector, or using the crystal resonator as one or both arms of a frequency discriminator. The details of implementing these measurements have been previously reported, but will be reviewed here.

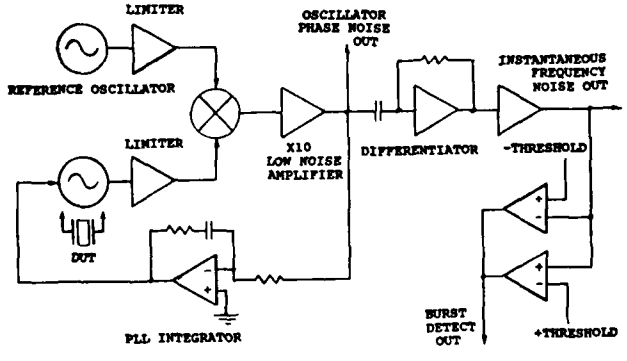



**Figure 2 Phase Locked Loop Method Block Diagram**

The two oscillator method of figure 2 uses a mixer as a phase detector. Two oscillators are phase locked in a narrow lock bandwidth. The output of the mixer accurately represents phase noise for frequencies far beyond the lock bandwidth. Corrections for the attenuation of the phase locked loop can be made as described in the operating manual of the Hewlett-Packard 11729C Carrier Noise Test Set[6]. Frequency deviations can also be measured by differentiating the phase noise output. With no correction for the phase locked loop attenuation, the frequency output will be valid between the loop bandwidth and the upper frequency limit of the differentiator. The noise observed will be a power sum of the noise in both oscillators at any particular offset frequency, so accurate measurement requires a very quiet crystal in the reference oscillator.

CH25.. -.., 8/0000-389 $1.00 C 1988 IEEE

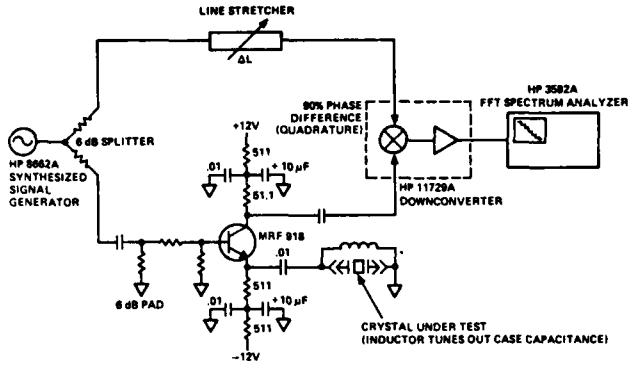



**Figure 3 Crystal Filter Freq. Discriminator Block Diagram**

The second method available to measure crystal noise measures the noise of the crystal as a filter, as shown in figure 3. A signal from a reference source is split into two paths. One signal is sent through the resonator under test. The second signal is sent through a second (reference) crystal filter or directly to an adjustable delay line to achieve quadrature at the mixer (phase detector).

Care must be taken to present as low an impedance as possible to the crystal resonator. If the crystal sees a real impedance higher than a small fraction of the series resistance of the resonator, the phase changes of the crystal will be attenuated. Too large an imaginary impedance will pull the crystal away from nominal series resonance.

The local negative feedback of the crystal impedance at resonance will reduce the driving transistor noise[11] sufficiently such that the crystal noise dominates, at least for crystals with series resistance larger than 20 to 30 ohms.

The proposed model of the crystal at resonance is a finite real series resistance in series with a small time varying imaginary impedance. If an external resistance equal to the series resistance is added in series with the crystal by the test fixture, the phase fluctuations of the network will be reduced by one half[8,9], or 6 dB. The reduction can be compensated for by calibrating the system for frequency noise.Calibrate the system using a single resonator with a through line in the other arm by measuring the output voltage change for a source frequency change. However, the extra resistance will degrade the noise floor of the measurement.

All three of the noise measurement systems have their advantages. The phase locked loop method can be calibrated once, and retain that calibration from crystal to crystal. To maintain calibration, select the reference oscillator as the locked oscillator for repeatable lock loop bandwidth, and ensure that the mixer input power is leveled to retain constant phase detector sensitivity. The phase locked loop method does require a known good oscillator that is both tunable and has lower phase noise than the noise to be measured.

The single crystal discriminator also requires a source oscillator with very low phase noise, since the source noise will be detected by the FM demodulation of the discriminator[8,9]. However, there is only one crystal to cause noise, if the source is known to be good. A VHF crystal will most likely be noisier than a synthesizer that is phase locked to a precision ovenized oscillator at 5 or 10 MHz, at least at offset frequencies below 100 Hz.

The two crystal discriminator reduces the sensitivity to source noise while it adds the requirement of a standard quiet reference crystal. One further advantage of the discriminator methods is the lack of a requirement for a low noise crystal oscillator for each crystal frequency, assuming a quiet synthesized source such as the Hewlett-Packard 8662A is available. The variable frequency source used with a simple test fixture can test crystals of many operating frequencies.

## Measurement Results

The measurements done using a phase locked loop allowed simultaneous measurements of phase noise and frequency fluctuations versus time. A comparator was set to thresholds slightly above and below the excursions due to residual FM that were expected for quiet crystals. Each time that the frequency noise exceeded either one of the thresholds (positive or negative) of the comparator, a "pop" or burst was recorded on an event counter. The number of bursts counted during a two minute averaging period was plotted along with the average phase noise for that same time period. The power dissipated in the crystal during the measurement was approximately 0 dBm.

Many crystals showed no burst noise and also had relatively stable, low levels of phase noise as shown in figure 4. Other

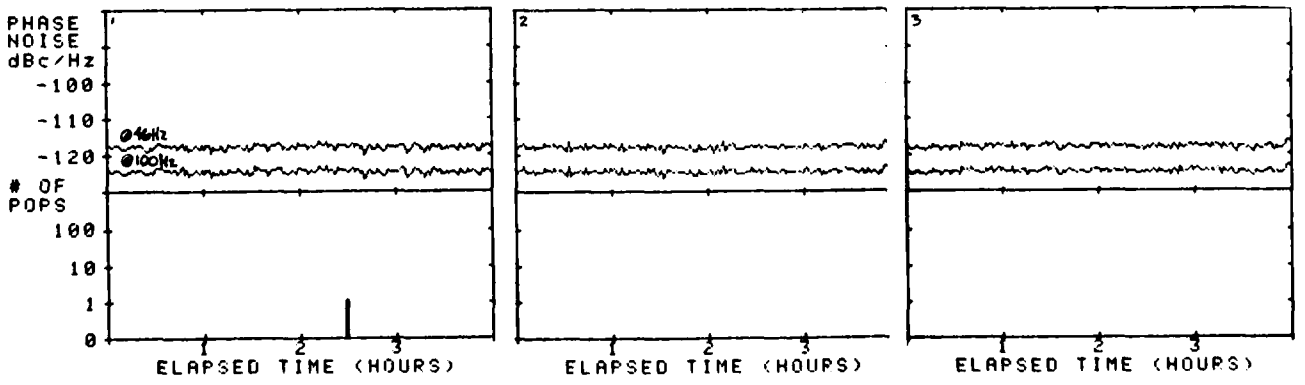



**Figure 4 Phase Noise and Burst Count Vs. Time Quiet Crystal**

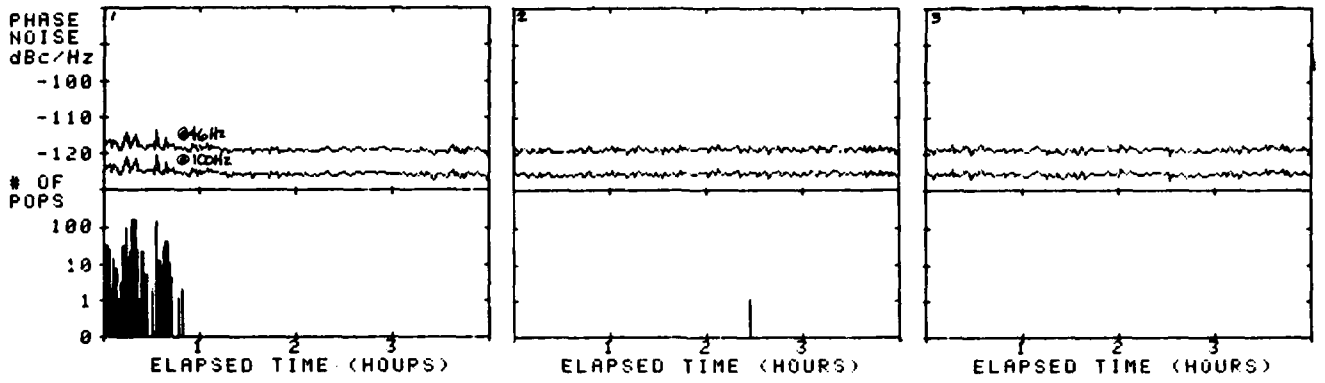



**Figure 5 Reduction of Noise With Time**

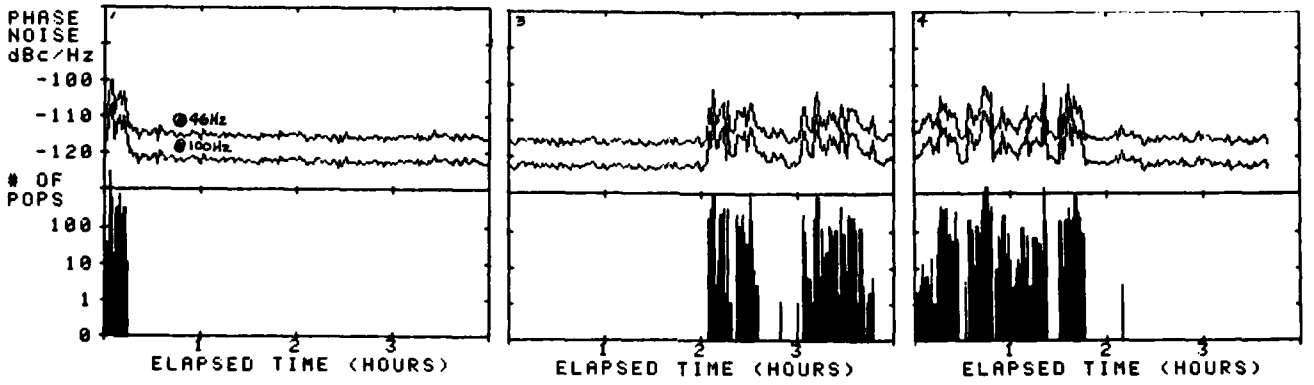



**Figure 6 Intermittently Noisy Crystal**

crystals had high levels of burst noise for approximately the first half hour, after which they showed little or no bursts (figure 5). Still other crystals showed intermittent bursts and very high phase noise (figure 6), while a few showed constant bursts and very high noise.

The first few minutes of power to the crystal appear to be the worst, and experimental results suggest that those crystals that have fewer than 10 to 20 bursts counted in their first two minutes under drive will settle into low noise behavior. This observation has been reported as annealing under power in both crystals[1] and SAW resonators[12,13,14] and may have similarities to the change of series resistance with increasing drive power also observed on filter crystals.

The time domain frequency fluctuations show complex behavior. The frequency measured is the instantaneous frequency, the derivative of phase noise, bandlimited from 20 Hz to 1 kHz. Compare the signals shown in figure 7 (quiet crystal) and figure 8 (very noisy crystal). Figures 9 through 12 show increasingly finer detail. None of the long term thermally[15] driven changes show up due to the lower frequency response limit at the lock loop bandwidth. The rise time of about 150 uS was still seen using a 10 kHz bandwidth, suggesting that the crystal filtering action limits the time response to changes in crystal frequency.

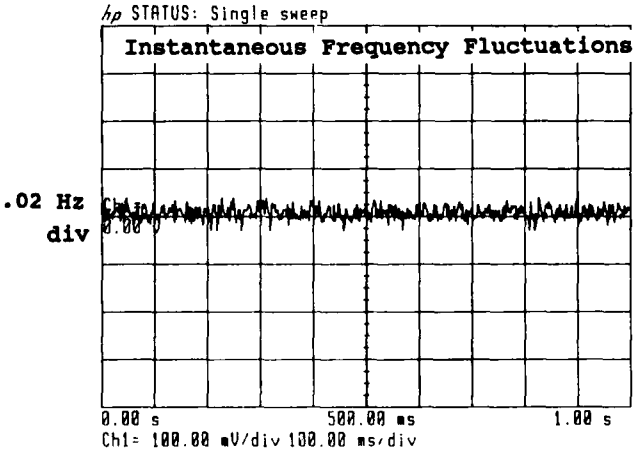



**Figure 7 Typical Good Crystal**

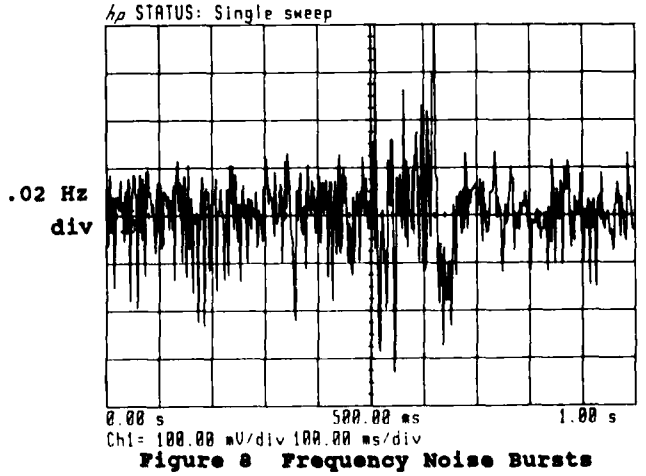



**Figure 8 Frequency Noise Bursts**

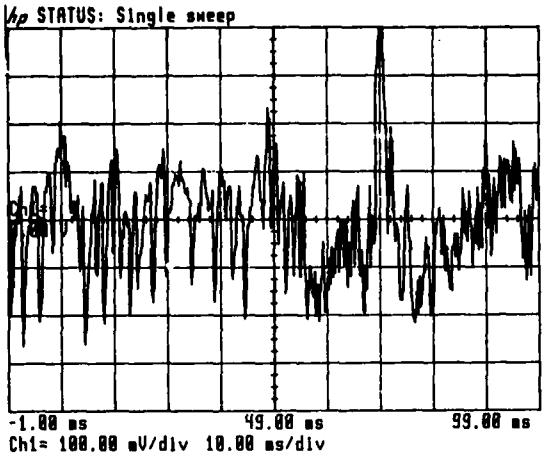



**Figure 9 Burst Train Followed by Large Burst**

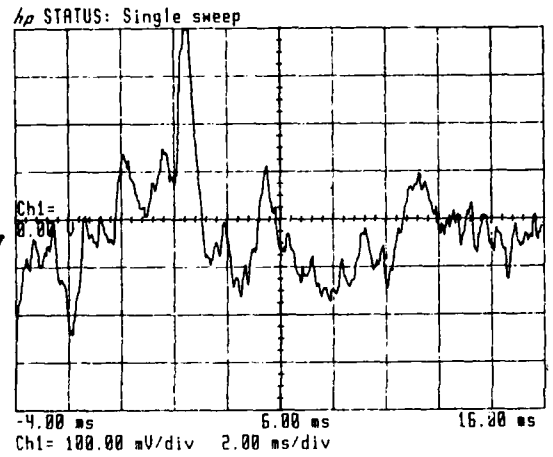



**Figure 10 Closer Detail of Bursts**

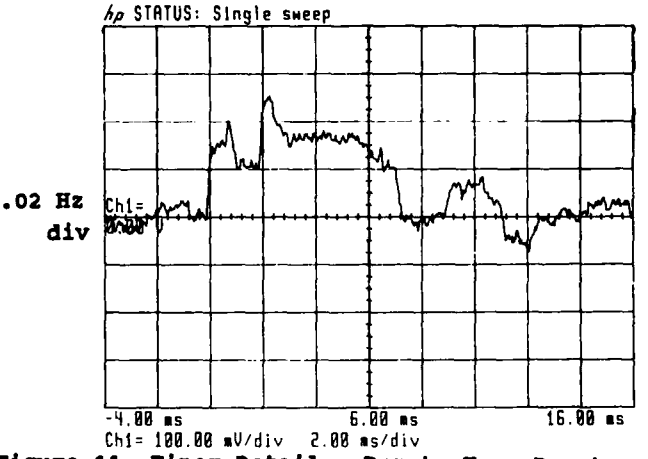



**Figure 11 Finer Detail - Bursts Upon Bursts**

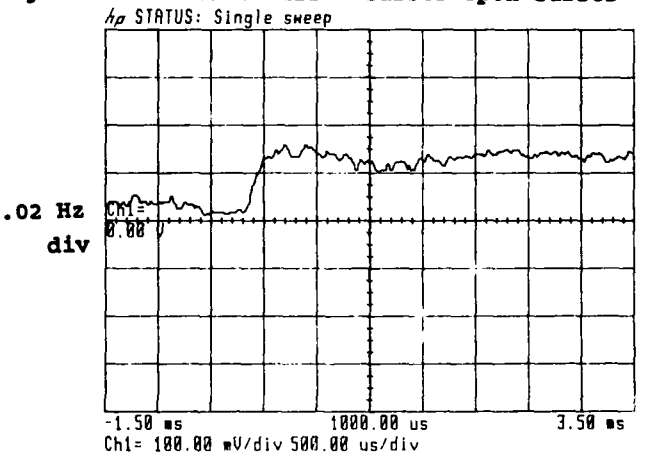



**Figure 12 Rise Time Detail (Approx. 150 uS)**

The detail visible shows steps, pulses, pulses upon pulses, and a base of noise that is similar to the level of the "quiet" crystal. One of the most interesting patterns seen was a series of very narrow pulses followed by a single very large burst that strongly suggests a nonlinear chaotic system[25,26,28] driven to the breaking point. After the large burst, the crystals were often very quiet for a time. Similar chaotic behavior was reported earlier by J. J. Gagnepain, M. Olivier and F. L. Walls[16,17], but showed up at very high drive levels (7 Volts). Some crystals may be more susceptable to the nonlinear chaotic oscillations at lower drive levels.

## Fundamental and Overtone Measurements

Another type of test system was built to observe if the bursts occurred at the fundamental and the fifth overtone simultaneously. Figure 13 shows the block diagram of a system built to drive the crystal at 100 MHz and 19.978 MHz while observing the phase (and frequency) noise independently at each frequency. The system is essentially the same as the single crystal discriminator, except that the crystal receives the combined input signal from both sources, while the output is split, filtered, and sent to each mixer. The fact that the fundamental mode is not exactly one fifth the frequency of the fifth overtone allows the observation to be made without interference. Even so, the mixer ports have to be heavily filtered to reject the beat frequency of 100 MHz mixed with 5 X 19.978 MHz, or 110 kHz.

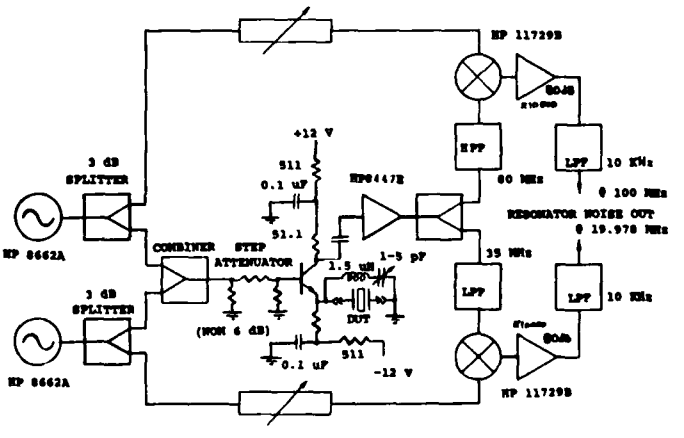



**Figure 13 Simultaneous Measurement of Frequency Noise at Fundamental and Fifth Overtone**

The system is calibrated by varying the frequency of one source and recording the dc voltage out of the mixer amplifier, while driving the second source at the other series resonant frequency. The inverse of the slope of the voltage versus frequency at quadrature gives the calibration multiplier to be used.

**Instantaneous Frequency Fluctuations**

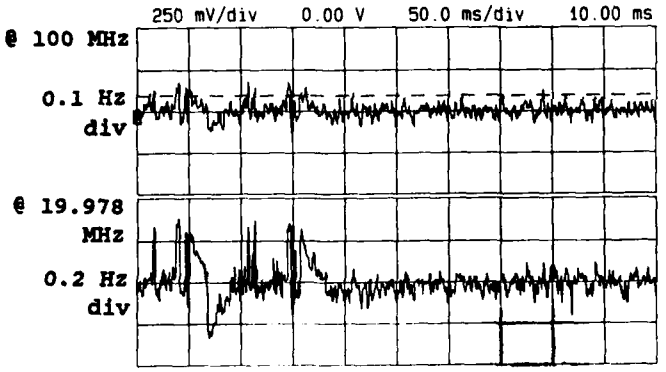



**Figure 14 Independent Noise Burst @ 19.978 MHz**

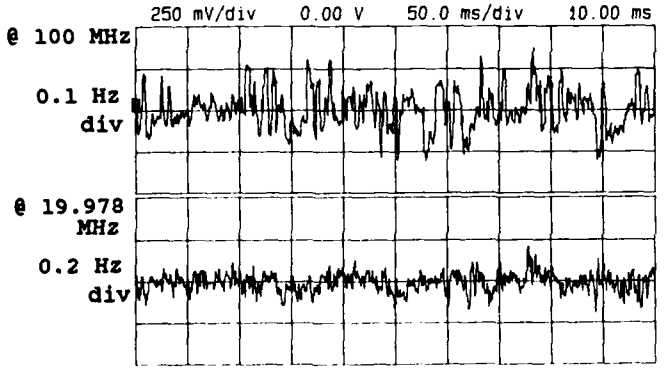



**Figure 15 Independent Noise Burst @ 100 MHz**

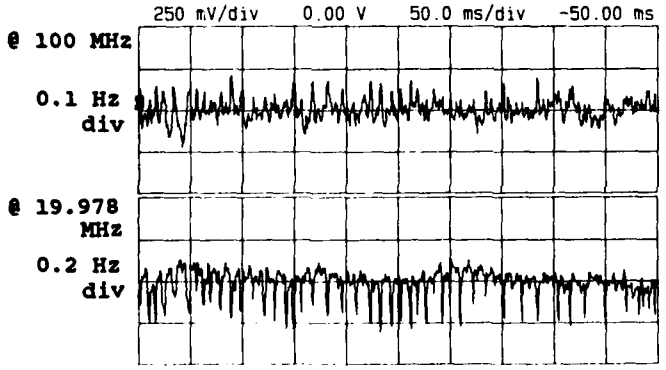



**Figure 16 Related Simultaneous Noise at Fundamental and Overtone**

Observations of the two frequency noise outputs show both independent (figures 14 and 15) and correlated noise bursts (figure 16). The narrow pulse train of figure 16 shows up at opposite polarities on the fundamental and the fifth overtone.

The independent nature of the bursts seen at each frequency indicates that the source of the mechanism may be confined to a very small area of the crystal. Tests done by Christian K. Sieracki indicated that the series resistance also changed with time on those crystals that showed frequency bursts with time. This suggests that the conclusions of J. E. Knowles [24] that fluctuations of series resistance are connected to surface defects (deep pits and scratches) may also apply to the burst sources.

**Correlation of Bursts and Crystal Parameters**

Several of the crystal parameters were measured in addition to the phase noise and burst count and an attempt made to correlate the crystal parameters. The data was not taken on a complete set of crystals, but on a population of crystals chosen to have the complete range of burst count seen, i.e. there were fewer of the quiet crystals in the sample than were typically received. Neither series resistance (figure 17) nor Q (figure 18) showed a strong correlation with observed phase noise at 100 Hz offset from 100 MHz (dBc/Hz). There may be a slight correlation of phase noise and higher motional capacitance indicated, but it isn't a strong link (figure 19).

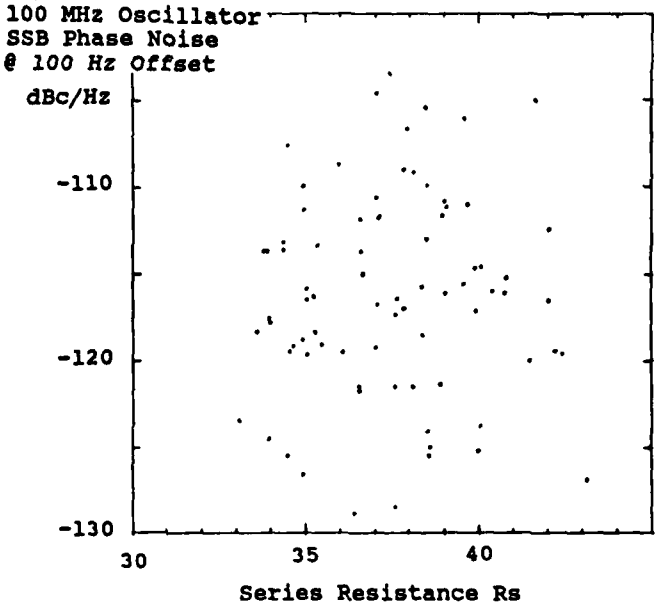



**Figure 17 Lack of Correlation Between Noise and Rs**

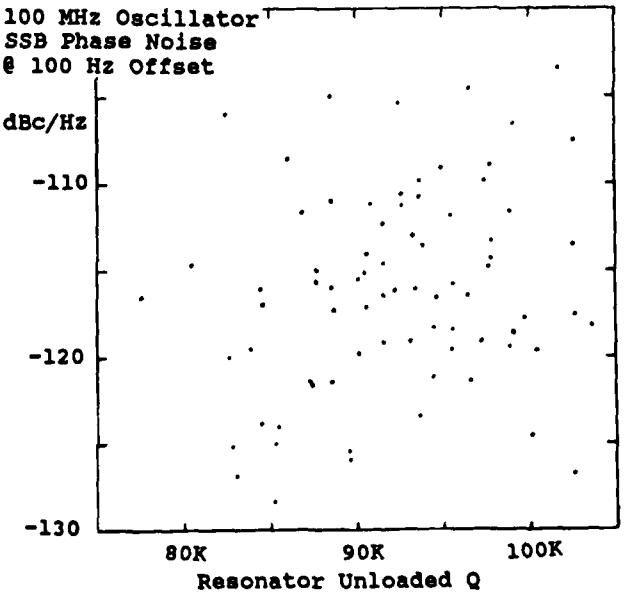



**Figure 18 Lack of Correlation Between Noise and Q**

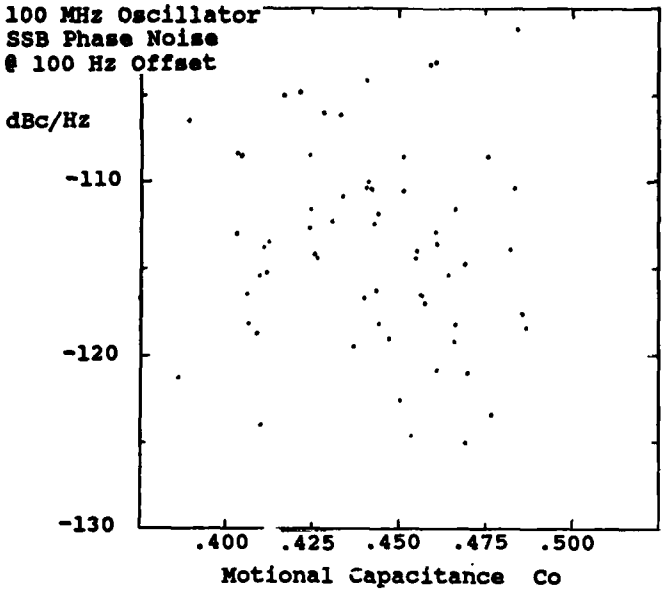



**Figure 19 Possible Link Between Noise and Co**

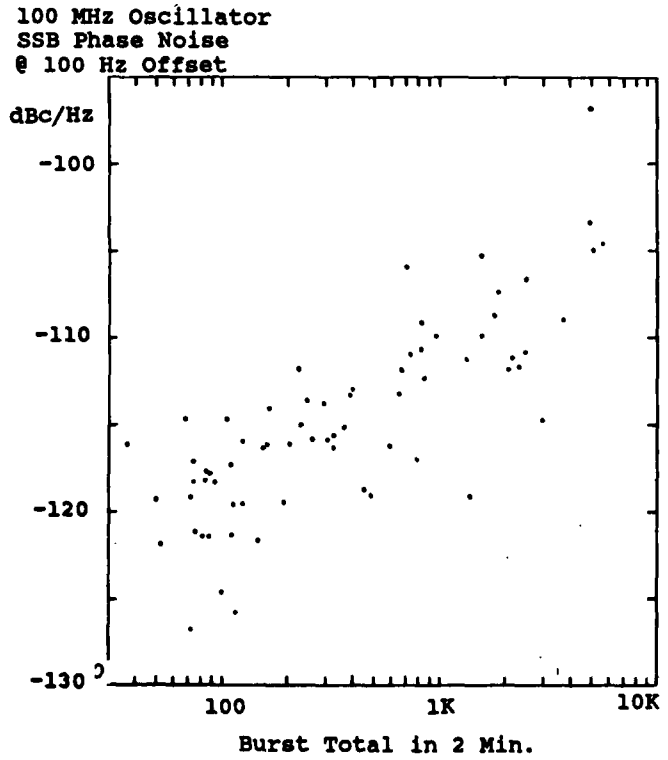



**Figure 20 Relationship Between Noise and Bursts**

There did appear to be a high correlation between average phase noise and burst count during that averaging period (figure 20). Note that an "eye" average shows low noise levels since the noise is usually low, but an FFT spectrum analyser using an RMS average remembers the high noise during bursts, and raises the noise level. The same correlation was visible in the plots of noise and bursts over time (figure 10). Based upon these observations, the bursts appear to account for most of the high noise and variability seen in phase noise measurements of crystal oscillators.

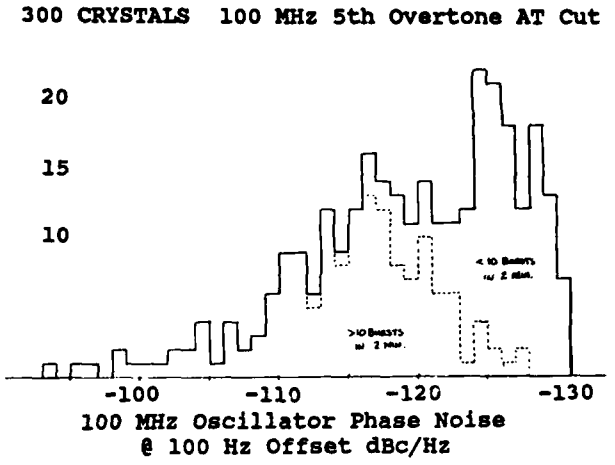



**Figure 21 Histogram of Oscillator Noise 100 Hz Offset from 100 MHz**

Figure 21 shows a histogram of observed oscillator phase noise at 100 Hz offset for three hundred 100 MHz crystals previously screened (by eye average) to -110 dBc/Hz. Note the concentration of samples near -125 dBc/Hz. The distribution also shows the limit of -130 dBc/Hz due to noise floor of the reference oscillator. There may be a combination of distributions, one with, and one without bursts. The high noise crystals all show high burst counts, as indicated by the shaded areas which indicate crystals with higher than ten bursts during the two minute averaging time.

The clustering of 1/f phase noise levels near -125 dBc/Hz at 100 Hz offset can be extrapolated to a level of near -120 dBc/Hz at 1 Hz for the phase noise of the resonator alone. These oscillator measurements agree with separate measurements of resonator phase noise of many 100 MHz crystals[9,10,21]. The phase noise at 1 Hz clustered about -120 to -125 dBc/Hz, based upon data taken at offsets in the 1/f region from 10Hz to 200 Hz. This agrees quite well with the measurements on SAW devices reported by B. Bray and S. Elliot[6], and by T.E. Parker[18]. These measurements also fall on the plot of the spectral density of frequency stability versus unloaded Q proposed by J.J. Gagnepain[19,20].

Consider only those crystals that have Q near the material limit, and only those that show no burst noise. If all crystals have resonator phase noise near -120 dBc/Hz at 1 Hz offset frequency, that fact may explain the $1/Q^4$ dependence of frequency fluctuations. If phase noise at 1 Hz is observed with a resonator bandwidth wider than a few Hz, then by Leeson's model[22] the 1 Hz phase noise will rise as the square of the bandwidth (-20 dB/decade noise slope added for offsets less than half the bandwidth). The bandwidth of the resonator is given by:

$$BW = fo/Q$$

The material limit of Q restricts the achievable Q to:

$$Q = k/fo$$

(where k is some fraction of the material limit)

then: $BW = k/Q^2$

since: $\mathcal{L}(1\ Hz) = 10 \log(BW^2) + N$

N is the resonator and amplifier phase noise (in dBc/Hz at 1 Hz offset frequency) of the oscillator with the feedback path removed. The oscillator phase noise falls -20 dB per decade inside the crystal bandwidth in addition to approx. -10 dB/decade for resonator noise.

thus: $\mathcal{L}(1\ Hz) = 10 \log(k^2/Q^4) + N$

This connection to constant resonator phase noise (not oscillator noise) was also mentioned by T. E. Parker[18]. Recently the relation of stability versus Q for many types of oscillators from pendulum clocks to atomic standards was discussed by D.A. Bateman[23], and other types of resonators were shown to have correlation between stability and Q.

## Conclusions

The intermittent nature of phase noise has been traced to burst noise. The bursts have been observed to quiet with operation under power in as low as half an hour,but many crystals do not quiet down, or if they do, often resume their earlier noise levels intermittently. The first few minutes under power are usually the worst, and can be used as a preview of future behavior. Similar observations made on SAW devices show reduction of phase noise under drive level.

None of the electrical parameters of the crystal unit show a strong correlation to the bursts. However, there has been speculation that it is due to the same sources as the Rs changes with drive level linked to surface defects from imperfect polishing by J. E. Knowles[24]. A second possibility, possibly from similar causes, is that of nonlinear chaotic oscilations due to heating of the crystal under drive causing a change in the resonant frequency reported by J. J. Gagnepain et al[19,20], or due to surface contamination (defects or particles) causing the nonlinear behavior.

The author would like to thank Val D. Mc Omber and Thomas B. Fetter for their work designing the burst noise test set, especially for the idea of observing frequency noise to distinguish bursts. The assistance and encouragement of Robert C. Bray, G. Stephen Curtis, Michael C. Fischer, and Stephen Warwick is greatly appreciated.

## References

1. Michael C. Fischer, personal communication.
2. Christian K. Sieracki, personal communication.
3. R. L. Baer et al., "Phase Noise in SAW Filters", 1984 IEEE Ultrasonics Symposium pp. 30-35. Cat. # 84 CH 2112-1
4. A. E. Wainwright, F. L. Walls, and W. D. Mc Caa, "Direct Measurements of the Inherent Frequency Stability of Quartz Crystal Resonators", Proc. of the 28th Annual Symposium on Frequency Control, 1974, pp.177-180.
5. J. J. Gagnepain, "Fundamental Noise Studies of Quartz Crystal Resonators", Proc. of the 30th Annual Symposium on Frequency Control, 1976, pp.84-91.
6. Scott S. Elliot and Robert C. Bray, "Direct Phase Noise Measurements of SAW Resonators", 1984 IEEE Ultrasonics Symposium, pp. 180-186, Cat. # 84 CH 2112-1.
7. Hewlett-Packard Co. Product Note 11729B-1, "Phase Noise Characterization of Microwave Oscillators, Phase Detector Method", Part # 5952-8286
8. G. Stephen Curtis, "The Relationship Between Resonator and Oscillator Noise, and Resonator Noise Measurement Techniques", Proc. 41st Annual Frequency Control Symposium, 1987, Cat. # CH2427-3/87/0000-0470 $1.00 c 1987 IEEE.
9. Grant E. Moulton, "Analysis and Prediction of Phase Noise in Resonators and Oscillators", Hewlett-Packard Co. RF and Microwave Symposium, 1985.
10. Grant E. Moulton, "Dig for the Roots of Oscillator Noise", Microwaves and RF, Vol. 25, No. 4, April 1986, pp.
11. Bokuji Komiyama and Yoshiyuki Yasuda, "Precision Frequency Sources and Flicker Noise of Phase in RF Amplifiers", The Transactions of the IECE of Japan, Vol. J 60-B No. 3, pp. 197-204.
12. Tim L. Bagwell and Robert C. Bray, "Novel Surface Transverse Wave Resonators With Low Loss and High Q", 1987 IEEE Ultrasonics Symposium, pp. 319-325, Cat. # 0090-5607/0000-0319 $1.00 c 1987 IEEE.
13. R. C. Bray, L. L. Pendergrass, C. A. *Johnsen*, T. L. *Bagwell*, J. L. Henderson, "Annealing Behavior and Phase Noise Performance of SAW Resonators", 1985 IEEE Ultrasonics Symposium, pp. 247-252, Cat. # 0090-5607/85/0000-0247 $1.00 c 1985 IEEE.
14. M. J. Loboda, T. E. Parker, J. A. Greer and G. K. Montress, "Reduction of Close-to-Carrier Phase Noise in Surface Acoustic Wave Resonators", IEEE 1987 Ultrasonics Symposium, pp. 43-46, Cat. # 0090-5607/87/0000-0043 $1.00 c 1987 IEEE.
15. Y. Noguchi, Y. Teramachi, and T. Musha, "1/f Frequency Fluctuation of a Quartz Crystal Oscillator and Temperature Fluctuation", Proc. of the 35th Annual Frequency Control Symposium, 1981, pp. 484-491
16. J. J. Gagnepain, M. Olivier, and F. L. Walls, "Excess Noise in Quartz Crystal Resonators", Proc. of the 35th Annual Symposium on Frequency Control, 1983, pp. 218-225, Cat. # CH1957-0/83/0000-0218 $1.00 c 1983 IEEE.
17. M. Olivier and J. J. Gagnepain, "Chaotic States and Anormalous Noise in Resonators", pp. 319-323, Noise in Physical Systems and 1/f Noise, Elsevier Science Publishers, The Netherlands, 1983, ISBN 0 444 86800 3.
18. T. E. Parker, "1/f Frequency Fluctuations in Acoustic and Other Stable Oscillators", Proc. of the 39th Annual Symposium on Frequency Control, 1985, pp. 97-106, Cat.# CH2186-0/85/0000-0097$1.00 c 1985 IEEE.
19. J. J. Gagnepain, J. Uebersfeld, G. Goujon and P. Handel, "Relation Between 1/f Noise and Q-Factor in Quartz Resonators at Room and Low Temperatures, First Theoretical Interpretation", Proc. 35th Annual Frequency Control Symposium, 1981, pp. 476-483.

20. J. J. Gagnepain, "A Comparison of Frequency Noise of Quartz Resonators", Proc. 39th Annual Symposium on Frequency Control, 1985, pp. 135-137, Cat. # CH2186-0/85/0000-0135$1.00 c 1985 IEEE.
21. R. C. Smythe, "Phase Noise in Crystal Filters", Proc. 39th Annual Symposium on Frequency Control, 1985, pp. 138-139, Cat. # CH2186-0/85/0000-0138$1.00 c 1985 IEEE.
22. D. B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum", Proc. IEEE Feb. 1966 pp.329-330.
23. D. A. Bateman, "Q and Oscillator Stability", Electronics and Wireless World, 1987, pp. 843-844.
24. J. E. Knowles, "On the Origin of the 'Second Level of Drive' Effect in Quartz Oscillators", Proc. Annual Frequency Control Symposium, pp. 230-236.
25. J. M. T. Thompson and H. B. Stewart, "Nonlinear Dynamics and Chaos, Geometrical Methods for Engineers and Scientists", John Wiley and Sons, 1984, ISBN 0-471-90960-2.
26. Pierre Berge, Yves Pomeau, Christian Vidal, "Order Within Chaos, Towards a Deterministic Approach to Turbulence", John Wiley and Sons, 1984, ISBN 0-471-84967-7.
27. James Gleick, "Chaos, Making a New Science", R.R. Donnelly and Sons Co., 1987. ISBN 0-670-81178-5
28. Arun V. Holden, "Chaos", Princeton University Press, 1987. ISBN 0-691-08424-6

42nd Annual Frequency Control Symposium - 1988

# RESONATOR SURFACE CONTAMINATION - A CAUSE OF FREQUENCY FLUCTUATIONS?

Y.K Yong
Department of Civil and Enviromental Engineering
Rutgers University
P.O.Box 909
Piscataway, New Jersey 08855-0909

John R. Vig
Frequency Control and Timing Branch
U.S. Army Electronics Technology and Devices Laboratory
ATTN: SLCET-EQ
Fort Monmouth, New Jersey 07703-5000

## SUMMARY

Studied are the mass loading effects of adsorbing and desorbing contaminant molecules on the magnitude and characteristics of frequency fluctuations in a thickness shear resonator. The study is motivated by the observation that the frequency of a thickness shear resonator is determined predominantly by such mechanical parameters as the thickness of the resonator, elastic stiffnesses, mass loading of the electrodes and energy trapping. An equation was derived relating the spectral density of frequency fluctuations to: (1) rates of adsorption and desorption of one species of contaminant molecules, (2) mass per unit area of a monolayer of molecules, (3) frequency constant, (4) thickness of resonator and (5) number of molecular sites on one resonator surface. The induced phase noises were found to be significant in very high frequency resonators and are not simple functions of the percent area contaminated. The spectral density of frequency fluctuations was inversely proportional to the fourth power of the thickness if other parameters were held constant. Since the resonator frequency is inversely proportional to the thickness, the spectral density is, in effect, proportional the fourth power of resonator frequency.

## I. INTRODUCTION

Short-term stability (i.e., noise) is among the least well-understood resonator phenomena. Factors that have been recognized as sources of short-term instabilities include: Johnson noise, temperature fluctuations, random vibrations, acoustic losses, noise originating from interfaces between the electrodes and the quartz plate and between the mounting structure and the quartz plate, and noise due to the oscillator circuitry. Vig[1] proposed a preliminary model of frequency fluctuations induced by surface contaminations and described qualitatively the effects of such contaminations.

This paper explores and estimates the frequency fluctuations caused by adsorption and desorption of one species of molecules on the surface of a thickness shear quartz resonator. The characteristics of such frequency fluctuations are studied. Our study is motivated by the observation that the resonant frequency of a thickness shear resonator is determined predominantly by such mechanical parameters as the thickness of the resonator, elastic stiffnesses, mass loading of the electrodes and energy trapping. Hence, the effects on the mass loading of the stochastic characteristics of mass loading due to adsorption and desorption of contaminant molecules are considered. These effects are confined to the electrode area, since the vibrational energy is predominantly trapped within the electrodes.

## II. ADSORPTION/DESORPTION OF CONTAMINANT MOLECULES.

A dynamic equilibrium exists between the resonator surface and the gas above it. This equilibrium is maintained by the rate of adsorption and desorption of contaminant molecules on the resonator surface. The area of this surface is taken as the electrode area, since most of the resonator energy is confined within the electrodes (energy trapping).

The rate of arrival of molecules at the surface can be evaluated readily from simple kinetic theory of gases. A simple form of the relationship between the the rate of arrival, pressure, temperature and molecular weight of the molecule is [2]:

$$r = 3.51\times10^{22}\ P/(T\ M)^{1/2}\ \text{molecules/cm}^2\text{/s} \qquad (1)$$

where P, T and M are, respectively, the pressure in torrs, temperature in kelvins and molecular weight. In the enclosed space of a crystal resonator, the pressure will vary slightly with the fluctuations in the number of adsorbed molecules. For the present linear model, we assume a constant pressure. If we consider the following example:

P, pressure = 1 torr
T, temperature = 300 K
M, molecular weight = 28 (nitrogen),

the rate of arrival, $r_o$, of nitrogen molecules at a contaminant site with a spacing of 0.5 nm is

$$r_o = 9.6\times10^{5}\ \text{molecules/site/s}.$$

For a given species of molecule at a fixed temperature, the rate would be linearly dependent on the pressure. The reciprocal of this rate is the mean interarrival time $\tau_o$.

The contaminant molecules reside on the surface for a finite time. In a simple physical adsorption process, the mean residence time is [3]:

$$\tau_1 = \alpha\ \exp(E_d/(RT)) \qquad (2)$$

where R is the gas constant = $1.987\times10^{-3}$ kcal/mol K, T is temperature in kelvins and $E_d$ is the desorption energy in kcal/mol. The exact value of the parameter $\alpha$

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

is conceptually not essential to our derivations. For subsequent calculation purposes we assign to $\alpha$ a value of $1x10^{-13}$ seconds. If we use a typical range of values for $E_d$ of 10 to 20 kcal/mol., the mean residence time is:

$$\tau_1 = 2x10^{-6} \text{ to } 50 \text{ seconds.}$$

The mean desorption rate of molecules from the surface, $r_1$, is the reciprocal of this mean residence time.

A monolayer or less of adsorbed contaminant molecules is assumed. When the system of contaminant gas and resonator surface settles down to a steady-state dynamic equilibrium, the states of the contaminant sites, which are either uncontaminated (state 0) or contaminated (state 1), can be described by steady-state probabilities. A transistion diagram in Fig.1 shows the interaction of the two states in a site. The rate in which the site enters a particular state must be equal to the rate at which it leaves. Hence, balance equations for the steady-state probabilities can be written:

| STATE | RATE ENTERING | | RATE LEAVING |
|---|---|---|---|
| 0 | $p_1 r_1$ | = | $p_o r_o$ |
| 1 | $p_o r_o$ | = | $p_1 r_1$ |

(3)

where $p_o$ and $p_1$ are the steady-state probabilities of state 0 and state 1, respectively. In writing these balance equations, the assumption was made that the molecules arriving at a site would be adsorbed immediately if the site is uncontaminated but would otherwise bounce off the surface and return to the gas, that is, we assume a sticking probability of one if the site is uncontaminated and a sticking probability of zero if the site is contaminated. In reality, both probabilities generally have a value between zero and one. The exact value of the sticking probability will, depending on the desorption rate, lower or increase the magnitude of frequency fluctuations, but will otherwise not affect the logic of our derivations.

Since the site must at any given time be in either state 0 or 1, $p_o$ and $p_1$ must sum to one:

$$p_o + p_1 = 1. \tag{4}$$

Therefore, from Eq.(3) and Eq.(4) we obtain

$$p_o = \frac{r_1}{r_o + r_1} \text{ and } p_1 = \frac{r_o}{r_o + r_1}. \tag{5}$$

$p_1$ can be interpreted as the ratio of contaminated area to total surface area, that is, the fraction of area contaminated. A highly contaminated or 'dirty' surface would have a value of $p_1$ close to unity.

## III. FREQUENCY CHANGE DUE TO A VERY SMALL MASS LOADING.

### (a) Frequency Influence Curve of a SC-Cut Strip.

The effect of a very small mass on the surface of a thickness shear resonator is investigated, namely, the effect on the resonant thickness shear frequency. For this purpose a one-dimensional finite element model is employed[4]. Since the ratio of the mass loading per unit area to the resonator mass per unit area is expected to be very small ($< 1x10^{-4}$), the magnitude of frequency change is linearly proportional to the mass loading. Consequently, the finite element calculations for frequency change at a mass loading ratio of, say, $1x10^{-4}$ can be scaled linearly to yield frequency change due to mass loading of contaminant molecules.

Figure 2 shows a unit width SC-cut strip subjected to a small mass loading (length 99 μm, thickness 5 nm and density equal to that of quartz). The position of the mass is given by the coordinate along the length of the strip. If the frequency change can be plotted against the position of the mass on the strip, a curve known as the frequency influence curve is obtained. Figure 3 gives the frequency influence curve for a mass loading (thickness 0.5 nm and length 0.5 nm) obtained from scaling the finite element calculations for the strip of Fig.2. We observe that the mass loading has predominant frequency influence over the middle third portion of the strip. The area under this frequency influence curve represents the frequency change due a 0.5 nm monolayer of molecules.

### (b) Frequency Change per Site over a 2-D Surface.

We assume for simplicity a flat frequency influence surface over the entire electrode area, although the actual frequency influence of contaminant molecules is predominantly over the middle third area of electrodes. It will be shown subsequently that the noise levels are inversely proportional to the surface area if other parameters are kept constant. Therefore, the assumption of a flat frequency influence surface will lower the noise levels. Based on the one-dimensional frequency influence curve of section (a), the assumption was estimated to create noise levels which were at least 3 dBc lower than when the actual frequency influence surface was used.

For a flat frequency influence surface, the change in frequency due to a monolayer of contaminant molecules is

$$\Delta f = f_o \frac{m'}{m} \tag{6}$$

where $f_o$, $m'$ and $m$ are, respectively, the resonator frequency, mass per unit area of a monolayer of molecules and mass per unit area of resonator. The frequency change per site is constant

$$\xi_i = \frac{\Delta f}{N}, \quad i = 1,2,\ldots,N. \tag{7}$$

where N is the number of sites in the active area of the resonator, that is, the area of electrode. The frequency change per site is calculated for two resonators

Example 1. 10 MHz Thickness Shear Resonator.

Electrode area = 10 $mm^2$
Thickness = 0.165 mm
0.5 nm monolayer of molecules (density = quartz)
Number of sites = $4x10^{13}$ (0.5 nm spacing)

$\Delta f$ = 30 Hz or 3 ppm
$\xi_i = 7.5x10^{-13}$ Hz

Example 2. 525 MHz AT-Cut Resonator.

Electrode area = 0.003 $mm^2$
Thickness = 3.2 μm
0.5 nm monolayer of molecules (density = quartz)
Number of sites = $12 \times 10^9$ (0.5 nm spacing)

$\Delta f$ = 82 KHz or 156 ppm
$\xi_i = 6.83 \times 10^{-6}$ Hz

We observe that the frequency change per site can increase by orders of magnitude if the thickness and electrode area of resonator are reduced.

## IV. STOCHASTIC PROCESSES.

Each contaminant site yields a stochastic process which is assumed stationary, ergodic and mutually independent of other sites. The stochastic process of frequency fluctuations per site i can be written as

$$f_i(t) = \xi_i b_i(t) \qquad i = 1,2,\ldots,N, \tag{8}$$

where $b_i(t)$ is a Bernoulli random variable with a value of one if the site is contaminated and zero otherwise. The mean or expected value of $b_i(t)$ is simply [5]

$$E[b_i(t)] = p_1. \tag{9}$$

where $p_1$ is the probability of the site being contaminated which, from Eq.(5), is the same as the fraction of area contaminated. The variance of $b_i(t)$ [5] is

$$\begin{aligned} \mathrm{Var}[b_i(t)] &= E[b_i(t)^2] - E[b_i(t)]^2 \\ &= p_1(1-p_1) \end{aligned} \tag{10}$$

Equation (10) could be expressed in terms of the adsorption and desorption rates if $p_1$ was substituted using Eq.(5)

$$\mathrm{Var}[b_i(t)] = \frac{r_o r_1}{(r_o + r_1)^2}. \tag{11}$$

(a) Frequency Fluctuations.

The resonator frequency fluctuations induced by adsorption and desorption of contaminant molecules are obtained by summing over N sites the stochastic processes $f_i(t)$

$$f(t) = 2 \sum_{i=1}^{N} f_i(t) \tag{12}$$

where the factor 2 is included because the thickness shear resonator has two surfaces. The term on the right-hand side is substituted by Eq.(8) to yield

$$f(t) = 2 \sum_{i=1}^{N} \xi_i b_i(t). \tag{13}$$

Since $b_i(t)$ is assumed stationary, ergodic and mutually independent, the mean and variance of f(t) can be derived easily, using elementary principles of statistics. The mean of f(t) is, employing Eq.(9),

$$\begin{aligned} E[f(t)] &= 2 \sum_{i=1}^{N} \xi_i E[b_i(t)] \\ &= 2p_1 \sum_{i=1}^{N} \xi_i. \end{aligned} \tag{14}$$

Equation (14) can be written in a simpler form, since the sum of frequency change per site over N sites is equal to $\Delta f$

$$E[f(t)] = 2p_1 \, \Delta f. \tag{15}$$

Hence, the mean frequency fluctuations is twice the product of the fraction of area contaminated and frequency change due to a monolayer of molecules.

The variance of f(t), $\sigma^2$, can be similarly derived

$$\begin{aligned} \mathrm{Var}[f(t)] &= 4 \sum_{i=1}^{N} \xi_i^2 \, \mathrm{Var}[b_i(t)] \\ &= 4p_1(1-p_1) \sum_{i=1}^{N} \xi_i^2. \end{aligned} \tag{16}$$

where Eq.(10) was employed. An expression in terms of the rates of adsorption and desorption, frequency change of a monolayer of molecules and total number of sites is obtained by substituting Eq.(11) and Eq.(7) into the right hand side of Eq.(16)

$$\sigma^2 = \frac{4 r_o r_1}{(r_o + r_1)^2} (\Delta f)^2 / N. \tag{17}$$

Equation (17) is valid only if a flat frequency influence surface is assumed. The variance is inversely proportional to N if other parameters were kept constant.

(b) Autocorrelation and Spectral Density Functions.

The autocorrelation function $R(\tau)$ is defined as [6],

$$R(\tau) = E[f(t+\tau)f(t)] \tag{18}$$

where the mean of f(t) has been removed and $\tau$ is the time lag. Using Eq.(13) and the assumption that the sites are mutually independent, Eq.(18) becomes

$$R(\tau) = 4 \sum_{i=1}^{N} \xi_i^2 \, E[b_i(t+\tau) b_i(t)]. \tag{19}$$

In the steady state dynamic equilibrium of one species of contaminant molecule and one type of site on the resonator surface, there is one adsorption rate $r_o$ and one desorption rate $r_1$. Hence, the stochastic

process has only one correlation time $\tau_c$

$$\tau_c = (r_o + r_1)^{-1}. \tag{20}$$

The autocorrelation function in Eq.(19) yields an exponential autocorrelation function noise[6,7]

$$R(\tau) = 4 \sum_{i=1}^{N} \xi_i^2 \, \mathrm{Var}[b_i(t)] \exp(-|\tau|/\tau_c). \tag{21}$$

From Eq.(16) and Eq.(21), we obtain

$$R(\tau) = \sigma^2 \exp(-|\tau|/\tau_c). \tag{22}$$

The one sided spectral density function, $S_{\Delta f}(f)$, is derived by taking the Fourier transform [6] (Wiener-Khintchine relation) of Eq.(22)

$$S_{\Delta f}(f) = \frac{4\sigma^2 \tau_c}{1 + (2\pi f \tau_c)^2} \tag{23}$$

where f is the Fourier frequency. $\sigma^2$ on the right-hand side is substituted for Eq.(17) to yield the expression

$$S_{\Delta f}(f) = \frac{16 r_o r_1 (\Delta f)^2 / N}{(r_o + r_1)^3 + 4\pi^2 f^2 (r_o + r_1)}. \tag{24}$$

It is worthwhile to describe in words the meaning of Eq.(24). The equation states that the spectral density of frequency fluctuations caused by one species of contaminant molecule and site is (1) a function of some rational fraction of the rates of adsorption and desorption, (2) proportional to the square of the frequency change induced by a monolayer of molecules and (3) inversely proportional to the number of sites. The spectral density function exhibits a white frequency noise process when $f << (r_o + r_1)/(2\pi)$ and random walk in frequency when $f >> (r_o + r_1)$. There is a simple relation between the spectral densities of frequency fluctuations and phase fluctuations [8], namely

$$S_\phi(f) = S_{\Delta f}(f)/f^2,$$

$$= \frac{16 r_o r_1 (\Delta f)^2 / N}{(r_o + r_1)^3 + 4\pi^2 f^2 (r_o + r_1)} \cdot \frac{1}{f^2}. \tag{25}$$

The spectral density of phase fluctuations is more commonly used in the field.

## V. RESULTS.

The spectral densities of phase fluctuations for the two resonators in examples 1 and 2 of section III(b) are studied over a wide range of $r_o$ and $r_1$ values. Table 1 shows the calculations for the phase fluctuations of the two resonators at 300K. The second column of the table was calculated using Eq.(1) and the pressure values of column 1. A molecular weight of 28 and site spacing of 0.5 nm were assumed. Column 4 was computed by employing the desorption energies in column 3 and the reciprocal of Eq.(2). Columns 5 and 6 are obtained from Eq.(5) and Eq.(20), respectively. The last two columns present the phase fluctuations calculated at 1 Hz, using Eq.(25), the parameters from examples 1 and 2 of section III(b) and rates from columns 2 and 4.

We observe from the second to last column of Table 1 that the 10 MHz thickness shear resonator yields phase noise levels which are barely measurable by today's existing noise measurement technology. The last column, on the other hand, shows that the 525 MHz resonator can produce significant levels of phase noise. We note that different noise levels are obtained for the same fraction of area contaminated.

In order to study the effects of percent area contaminated on phase noise levels, we employ Eq.(5) to rewrite Eq.(25) in the following form

$$S_\phi(1) = \frac{16 r_1 p_1 (1-p_1)^2 \Delta f^2 / N}{r_1^2 + [2\pi(1-p_1)]^2}; \qquad 0 < p_1 < 1. \tag{26}$$

where the spectral density is evaluated at 1 Hz. If we fix the parameters $\Delta f$ and N, we can plot the spectral density against the fraction of area contaminated for a given value of $r_1$, i.e., surface desorption energy. Figure 4 shows such a plot for different values of desorption energy. The parameters for the 525 MHz resonator from example 2 of section III(b) were used. We observe that for $p_1$ less than 0.1, the phase noise decreases linearly with decreasing $p_1$. In general, the noise level decreases rapidly when $p_1$ is greater than 0.9. Hence the resonator is noisest when the fraction of area contaminanted is greater than 0.1 but less than 0.9. The magnitude of the peak noise level depends on the value of the desorption energy, the highest of which occurs in the range of $E_d$ values from 17 to 18 kcal/mol (the parameter $\alpha$ of Eq.(2) being kept constant at $1 \times 10^{-13}$ seconds). The value of $p_1$ at the peak noise level approaches unity when $E_d$ is greater than 20 kcal/mol.

The effects of the two state variables, namely, the pressure and temperature, on phase noise levels can be studied. Figure 5 exhibits the phase noise-temperature curves of the 525 MHz resonator at a pressure of 0.1 torr. We observe that when the temperature is increased from a reference temperature of, say, 25 $^{o}$C, the phase noise levels can increase or decrease depending on the values of $E_d$. The turnover temperatures of the curves increase with increasing values of $E_d$.

Figure 6 shows the phase noise-pressure curves of the 525 MHz resonator at a temperature of 300 K. The overall shapes of the curves look similar to the curves in fig. 5 but the turnover pressures decrease with increasing values of $E_d$. We observe that a decrease in pressure does not necessarily imply a decrease in noise levels.

Figure 7 shows the spectral density of phase fluctuations for the 525 MHz resonator using Eq.(25) and $r_o$ and $r_1$ equal to 100 and 0.1 $s^{-1}$, respectively.

For Fourier frequencies of less than 10 Hz, the white frequency fluctuations predominates. The random walk frequency noise is dominant at Fourier frequencies greater than 50 Hz.

We can make another important observation based on the parameters in Eqs.(24) or (25). For this purpose, we replace the term $\Delta f$ in Eq.(24) using Eq.(6)

$$S_{\Delta f}(f) = \frac{16 r_o r_1 (f_o m'/m)^2/N}{(r_o + r_1)^3 + 4\pi^2 f^2 (r_o + r_1)} . \qquad (27)$$

Upon a first glance, it would appear that the spectral density is proportional to the square of the resonator frequency $f_o$, if other parameters were held constant. But the frequency of a thickness shear resonator is inversely proportional to the thickness of the plate, namely,

$$f_o = k/h \qquad (28)$$

where h is the resonator thickness and k is the frequency constant determined by the elastic constants and overtone of thickness shear vibration. The mass per unit area is also related to the thickness

$$m = \rho h \qquad (29)$$

where $\rho$ is the resonator density. Substituting Eqs.(28) and (29) into Eq.(27), we obtain

$$S_{\Delta f}(f) = \frac{16 r_o r_1 (k\, m')^2/(N h^4)}{(r_o + r_1)^3 + 4\pi^2 f^2 (r_o + r_1)} . \qquad (30)$$

Hence, the spectral density is inversely proportional to the fourth power of h. Since the resonator frequency is inversely proportional to h, Eq.(30) shows, in effect, that the spectral density is proportional to the fourth power of resonator frequency. In an actual thickness shear resonator, the electrode area usually decreases with increasing resonator frequency, which leads to decreasing values of N and increasing magnitudes of phase noise.

Experimental results relating the flicker frequency fluctuations to the fourth power of quartz resonator frequency were reported by Parker[9] as well as others. Gagnepain, et.al[10] were the first to put forward a phenomenological law relating the flicker frequency noise levels to the reciprocal of the fourth power of resonator Q. Weissman[11] reported in a recent review paper that a superposition of a number of Lorentzian spectra similar to Eq.(30) will yield flicker frequency fluctuations. When more than one species of contaminant molecules and sites are present in the contaminant gas-resonator surface system, the stochastic process has more than one correlation time. Hence, depending on the distribution of the correlation times, the basic model in Eq.(30) can produce flicker frequency fluctuations over some range of Fourier frequencies.

## VI. CONCLUSIONS.

The following conclusions are applicable to phase noise in thickness shear resonators induced by one species of contaminant molecule and site:

(1) The phase noise is not a simple function of the percent area contaminated, as shown in Eq.(26). A very clean and a 'dirty', that is, a highly contaminated resonator can both have low noise levels.

(2) The phase noise due to surface contaminations can be significant in VHF, UHF and membrane resonators.

(3) Surface contamination caused noise levels can be altered by:
   a. changing the pressure of the contaminant gas above the surface.
   b. changing the temperature and treating the surface to obtain different rates of adsorption and desorption
   c. changing the area of the electrode. The smaller the area, the higher the noise levels.

4) If all parameters except the thickness h are held constant in Eq.(30), the spectral density of frequency fluctuations is inversely proportional to the fourth power of h. Since the resonator frequency is inversely proportional to h, the spectral density is also proportional to the fourth power of the resonator frequency.

One possible experiment to verify the findings in this paper may involve measuring the phase noise of a very high frequency resonator while in a controlled atmosphere in a vacuum chamber and gradually decreasing the pressure. Extra care must be taken to keep constant all other parameters in Eq.(30) and to reduce other noise sources so as not to interfere with the measurements. The gas above the resonator must consist of essentially one species of contaminant molecules. A concomitant experiment would be to keep the pressure constant and vary the temperature instead. The effects of decreasing thickness on phase noise levels in a given high frequency resonator should also be studied.


## ACKNOWLEDGEMENTS

This work was supported by the U.S. Army Research Office, contract no. DAAL03-87-K-0107. The authors thank Charles A. Greenhall, Jet Propulsion Laboratory and Thomas E. Parker, Raytheon Company, for their helpful comments. The authors would also like to thank James T. Stewart, Rutgers University, for computing the one-dimensional frequency influence curve.


## REFERENCES.


1) Vig, J.R., "The Effects of Surface Contamination on the Noise and Drive-Level Sensitivity of Piezoelectric Resonators," Research and Development Technical Report SLCET-TR-87-5, ADA181299,May 1987.

2) Prutton, M., Surface Physics, 2nd Edition, Clarendon Press, Oxford, 1983, pp 5-6.

3) Prutton, M., Surface Physics, 2nd Edition, Clarendon Press, Oxford, 1983, pp 103-104 and pp 110-112.

4) Yong, Y.K., "On the use of 1-D finite elements for the temperature behavior of a contoured and partially plated SC-cut resonator," Proceedings of the IEEE 1987 Ultrasonics Symposium, pp 353-358.

5) Ross, Sheldon M., Introduction to Probability Models, 3rd Edition, Academic Press, Orlando, Florida, 1985, Chapter 2, pp 21-82.

6) Bendat, J.S. and Piersol, A.G., Random Data, Analysis and Measurement Procedures, 2nd Edition, Wiley-Interscience Publication, New York, 1986, Chapter 5, pp 109-163.

7) Voss, R.F., "1/f (Flicker) Noise: A Brief Review," Proceedings of the 33rd Annual Symposium on Frequency Control, 1979, pp 40-46.

8) Howe, D.A., Allan, D.W. and Barnes, J.A., "Properties of Signal Sources and Measurement Methods," Proceedings of the 35th Annual Symposium on Frequency Control, 1981, pp A1-A47.

9) Parker, T.E., "1/f Frequency Fluctuations in Quartz Acoustic Resonators," Applied Physics Letters, Vol. 46, No. 3, 1985, pp 246-248.

10) Gagnepain, J.J., Uebersfeld, J., Goujon, G. and Handel, P., "Relation between 1/f noise and Q-factor in quartz resonators at room and low temperatures, first theoretical interpretation," Proceedings of the 35th Annual Frequency Control Symposium, 1981, pp 476-483.

11) Weissman, M.B., "1/f noise and other slow, nonexponential kinetics in condensed matter," Reviews of Modern Physics, Vol. 60, No. 2, April 1988, pp 537-571.


TABLE 1: PHASE FLUCTUATIONS DUE TO SURFACE CONTAMINATIONS IN 10 MHZ AND 525 MHZ THICKNESS SHEAR RESONATORS

| Pressure torr | $r_o$ 1/s | $E_d$ kcal/mol | $r_1$ 1/s | Fraction of Area Cont. | $\tau_c$ s | 10 MHz Resonator* $10\text{Log}S_\phi(1)$ $rad^2/Hz$ | 525 MHz Resonator* $10\text{Log}S_\phi(1)$ $rad^2/Hz$ |
|---|---|---|---|---|---|---|---|
| $10^3$ | $10^9$ | 5.49 | $10^9$ | .500 | $5\times10^{-10}$ | -193 | -90 |
| 1 | $10^6$ | 9.61 | $10^6$ | .500 | $5\times10^{-7}$ | -163 | -60 |
| $10^{-3}$ | $10^3$ | 13.7 | $10^3$ | .500 | $5\times10^{-4}$ | -133 | -30 |
| $10^{-5}$ | 10 | 16.5 | 10 | .500 | $5\times10^{-2}$ | -114 | -9.9 |
| $10^{-6}$ | 1 | 17.8 | 1 | .500 | .5 | -114 | -9.9 |
| $10^{-7}$ | $10^{-1}$ | 19.2 | $10^{-1}$ | .500 | 5 | -123 | -19 |
| $10^{-9}$ | $10^{-3}$ | 22.0 | $10^{-3}$ | .500 | 500 | -143 | -39 |

* Resonators from examples 1 and 2 of Section III(b)
Note: A heavier contaminant molecule will raise the phase noise levels.

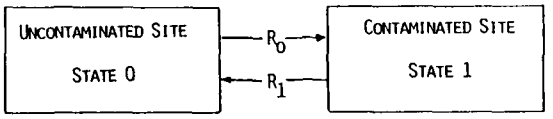



FIG.1 TRANSISTION DIAGRAM FOR THE TWO STATES OF A MOLECULAR SITE.

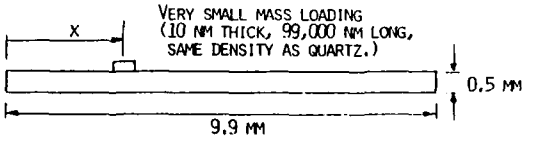



FIG.2 SC-CUT STRIP SUBJECTED TO A VERY SMALL MASS LOADING.

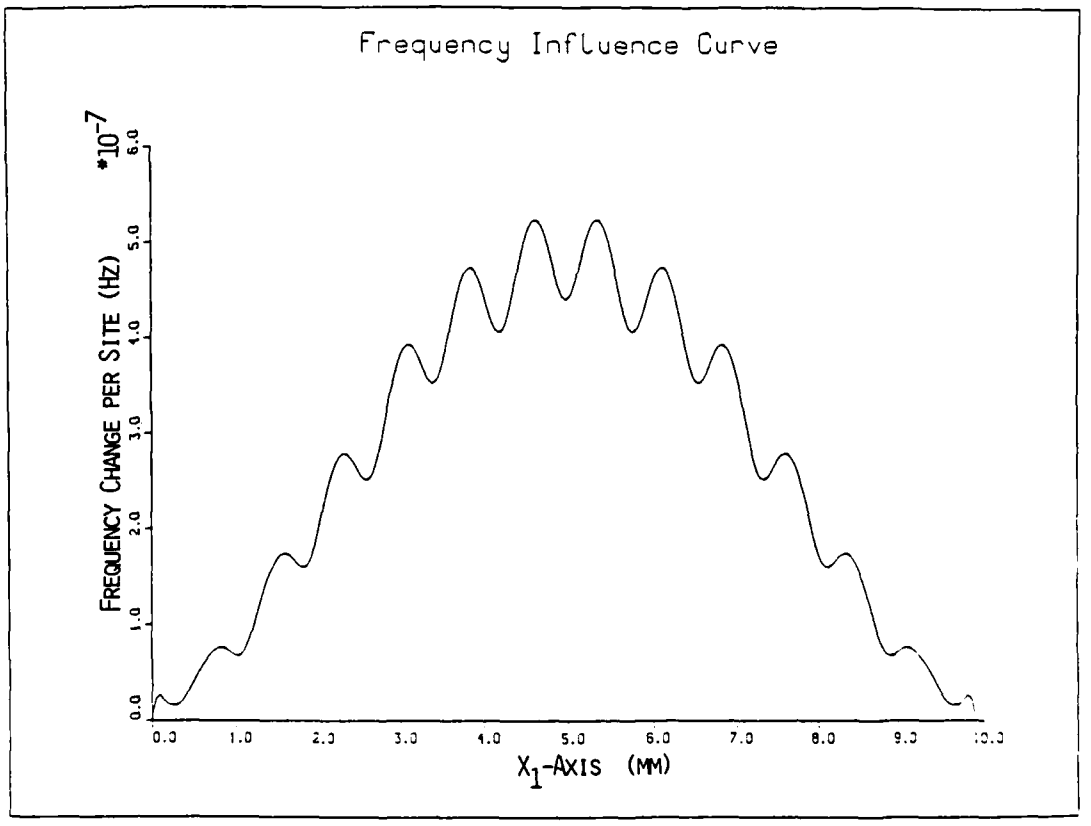



FIG.3 FREQUENCY INFLUENCE CURVE OF THE SC-CUT STRIP IN FIG.2.

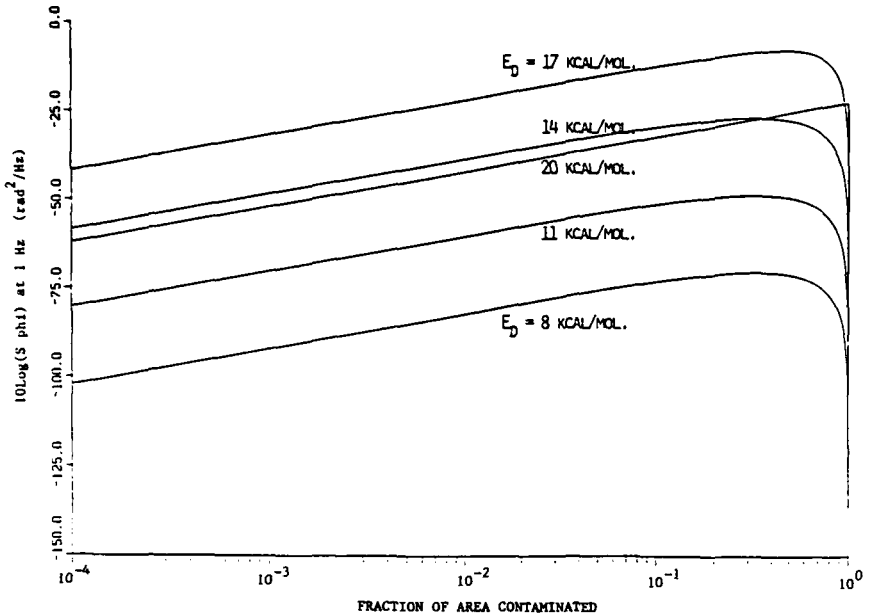



FIG.4 EFFECTS OF FRACTION OF AREA CONTAMINATED ON THE SPECTRAL DENSITY OF PHASE FLUCTUATIONS AT 1 Hz(525 RESONATOR)

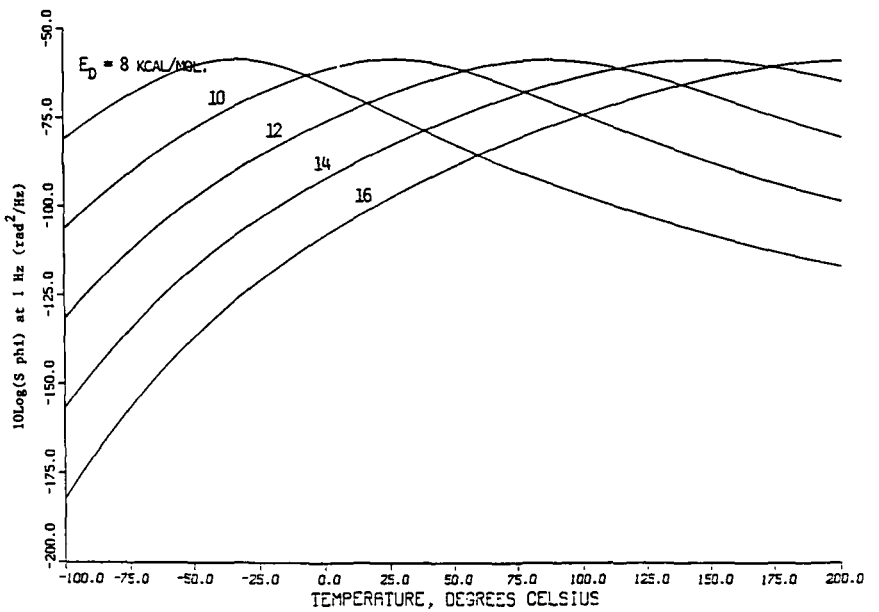



FIG.5 PHASE NOISE-TEMPERATURE CURVES FOR THE 525 MHZ RESONATOR.



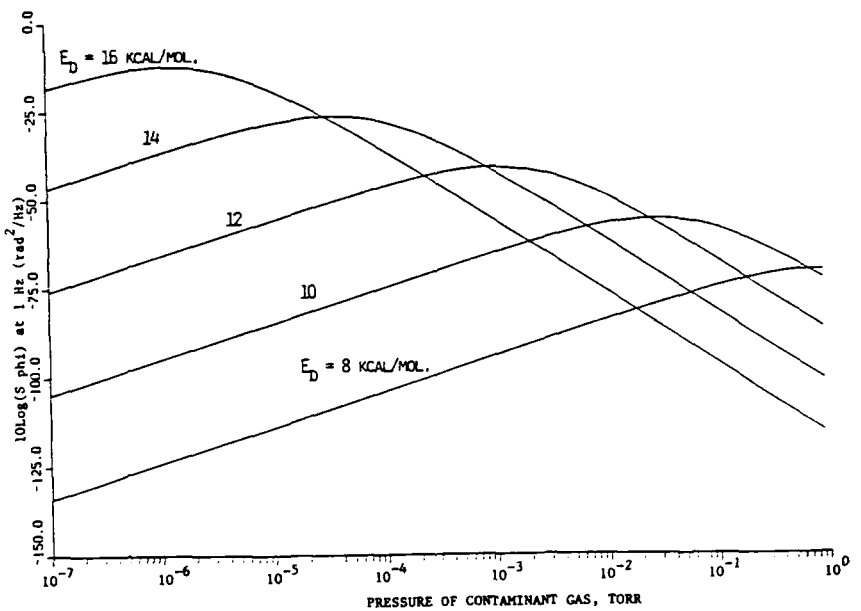



FIG.6 PHASE NOISE-PRESSURE CURVES FOR THE 525 MHZ RESONATOR.

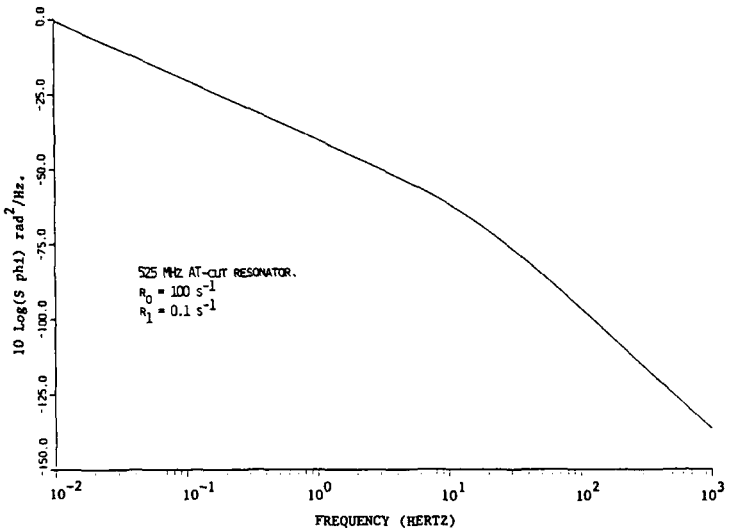



FIG.7 SPECTRAL DENSITY OF PHASE FLUCTUATIONS FOR THE 525 MHZ AT-CUT RESONATOR.

42nd Annual Frequency Control Symposium - 1988

# AGING PREDICTION OF QUARTZ CRYSTAL UNITS

Miodrag R. Miljković, Gordana Lj. Trifunović and Vladimir J. Brajović

Institut "Mihailo Pupin", Volgina 15, 11000 Beograd, Yugoslavia

## Summary

The paper deals with mechanisms which influence crystal unit stability with respect to time as well as their mathematical models. It is shown that it is possible to forecast aging over a longer period of time using extrapolation based upon these models, if good correlation exists between the frequency changes measured over a short period of time and the calculated values. This method is used for high-stability components. It points out the dominant aging mechanism, which can be used for technological process control. The method of accelerated aging at higher temperatures can, provided the activation energy of the process is known, also be utilized for prediction purposes, especially for large-scale manufacturing with moderate technology. Several examples of the results obtained by these analyses are given.

## Introduction

Long-term stability or aging of quartz crystal resonators is defined as the change of their resonant frequency in time. This change is observed by measurement, so it is necessary to create such measurement conditions as to eliminate the influence of all external factors on frequency stability. These factors are understood to be temperature (static and dynamic f-T changes), shock, vibration, radiation and the change of operating conditions in the oscillator circuit (supply voltage level, load reactance, continuity of operation, etc.).

Aging is expressed as the change in frequency after a period of time relative to the frequency at the beginning of the period, or the aging rate at a certain point in time during this process. The aging information in both cases refers to a certain temperature, noting whether the crystal unit was constantly in operation during the process or not.

One of the most important characteristics of devices which employ crystal units is their frequency stability in time. Aging knowledge is therefore especially important in assessing the quality of a device or module. Exact aging prediction enables, for example, acurrate determination of oscillator recalibration periods or crystal filter lifetimes, based on the allowed shift of the critical point in the filter characteristic, which depends on crystal unit frequency.

The aging characteristics of a certain type of crystal unit can be determined with absolute certainty only by measuring its frequency over a sufficient period of time. Since this period cannot be too long for practical reasons, least of all equal to the planned lifetime of the device (say 20 years), two problems arise:

- what is the minimum period of time necessary to obtain a reliable estimate of aging for a considerably longer period and how to obtain it
- what will the aging be during that considerably longer period.

Starting from the nature of the aging process, in our work we have attempted to apply two aging prediction methods:

a) the extrapolation method
b) the accelerated aging method.

## The extrapolation method

The main part of the crystal unit is the quartz resonator plate, having a specific shape and dimensions, cut from a quartz crystal bar, then processed by aggressive mechanical and chemical methods, coated with thin film metal electrodes, mounted and finally hermetically sealed in a holder. These operations disturb the quartz structure which tends to reach a state of equilibrium, whereas the materials on and around the quartz plate undergo physical and chemical transformations in time, also striving to attain a state of equilibrium. All of this is transferred to the resonator, changing its physical characteristics (elastic properties, density, mass distribution), and therefore its resonant frequency.

Reaching an equilibrium of the resonant frequency, i.e. of the relevant physical proper-

CH2588-2/88/0000-404 $1.00 C 1988 IEEE

ties, is explained by the kinetic theory of rate processes[1].

Researchers generally concur that the change in frequency is primarily influenced by two factors: the change in mass distribution and stress relaxation[1,2,3].

The changes in mass distribution arise during the processes of adsorption and desorption of physisorbed monolayers, as well as first-order chemical reactions, characterized by a specific aging constant $K_1$, and are expressed by the relation[1,4]:

$$\frac{df(t)}{dt} = -K_1 f(t) \qquad (1)$$

where f is frequency and t time. Integrating (1), the dependence of frequency on time takes the form

$$f(t) = f_\infty \pm (f_o - f_\infty)\, e^{-K_1 t} \qquad (2)$$

where the "-" sign refers to the units with "+" aging (whose frequency change is positive) and vice versa.

Other processes which involve mass changes are chemisorption and oxidation, described by[1,2]:

$$f(t) = f(o) + K_r \ln(1+Bt) \qquad (3)$$

where $K_r$ is the recovery rate constant for a particular type of crystal unit and B represents a constant related to the initial conditions.

Finally, processes exist which may influence mass distribution without being connected to mass transfer between the electroded plate and the ambient gas-filled holder - diffusion of the electrode material into the quartz or interdiffusion in the case of several different electrode metals. They are power law dependent, with the exponent approximately 1/2[2,5]:

$$f(t) = At^B, \quad B \approx 0.5 \qquad (4)$$

where A = const.

Stress changes are related to the relaxation of internal stresses in the quartz as well as in the electrodes. Assuming that the activation energy of the recovery process is a decreasing function of the stress, the relaxation rate is given by[1,6]:

$$\frac{d\sigma_e}{dt} = -a \cdot e^{-(U-b\sigma_e)/k_B T} \qquad (5)$$

where $\sigma_e$ is the mean value of the internal stresses in the electrodes, a, b and U are constants, T is the absolute temperature in K and $k_B$ is Boltzmann's constant. Integration of (5) gives

$$\sigma_e(t) = \sigma_e(o) + A \cdot \ln(1+Bt) \qquad (6)$$

$$\text{where} \quad A = -\frac{k_B T}{b}, \quad B = \frac{ab}{k_B T}\, e^{-(U-b\cdot\sigma_e(o))/k_B T} \qquad (7)$$

For $Bt \gg 1$ a simplified expression may be used

$$\sigma_e(t) = K_r \ln(t) \qquad (8)$$

where $K_r$ is the recovery rate constant.

Assuming that the stresses in plastically deformed electrodes are balanced by perfectly elastic stresses in the quartz, and that the stress in the quartz is directly related to the frequency, logarithmic laws (6) and (8) may also be applied for frequency changes

$$f(t) = f(o) + A \cdot \ln(1+Bt)$$

and (for $Bt \gg 1$): $f(t) = A\ln t + B', \; B' = A\ln B. \qquad (9)$

All the above processes and their mathematical models are summarized in the following table.

Table 1.

| Aging model | Aging mechanisms or physical interpretation of mathematical models | Mathematical model |
|---|---|---|
| LOG 1 (logarithmic) | - chemisorption<br>- oxidation<br>- internal stress relief and relaxation processes in the electrode film | $\Delta f/f = A\ln(Bt+1)+C$ |
| LOG 2 (logarithmic) | | $\Delta f/f = A\ln(t)+B'$, $B' = A\ln(B)+C$ (for $Bt \gg 1$) |
| EXP (exponential) | - first-order chemical reactions<br>- adsorption and desorption of a physisorbed monolayer | $\Delta f/f = A[1-\exp(-t/B)]$ |
| POW (power) | - diffusion processes | $\Delta f/f = At^B$, $B \approx 1/2$ |

The extrapolation method consists of measuring the aging of a number of samples of a certain type of crystal unit and, after a sufficient period of time, approximating the frequency changes by the mathematical models given in Table 1 or their combinations, i.e. determing the constants (A,B,C,etc.) by a least squares curve fit. A brief review of the computer program employed is given in the Appendix. The formula yielding the best correlation between the measured and computed values of frequency is adopted as the mathematical aging law, which enables prediction of aging for a period of time longer than the duration of the test.

We have applied the method to two types of AT crystal units belonging to different classes: high-precision for OCXO and moderate-precision for TCXO applications. The latter were 5472 kHz nominal frequency, vacuum coldweld units with gold electrodes with a chrome base. The test period was two years, with the units in constant operation in a TCXO. Six samples were

tested at a room temperature which varied from 20 to 30°C. The frequency was measured by a counter with a stability of $5 \cdot 10^{-11}$/day and a resolution of 0.01 Hz i.e. $2 \cdot 10^{-9}$. The data was processed using a computer program for fitting arbitrary functions to data which is described in the Appendix. The correlation was expressed by the determination coefficient $R^2 \leqq 1$, representing the ratio between the explained and total variations of the measured frequency, relative to the adopted mathematical law.

The LOG 1, LOG 2, EXP and POW models (Table 1) were used in the analysis of measurement results. Curves were fit for all samples, for periods of 1, 3 and 6 months, and aging was predicted, based on the constants obtained, for a period of 2 years and then compared to the actual values.

Prediction results are given in Table 2.

Table 2.

| Sample No. | Measured two years' aging (ppm) | The best approximation | Two years' aging prediction based on the best approximation for a period of 30days | 90days | 180days | Aging prediction for a period of 10 y. (ppm) |
|---|---|---|---|---|---|---|
| 1 | -1.16 | LOG 1 | -1.14 | -1.40 | -1.23 | -1.8 |
| 2 | -2.62 | LOG 1 | -1.65 | -3.21 | -2.78 | -4.1 |
| 3 | -1.64 | LOG 1 | -1.10 | -1.28 | -1.45 | -2.1 |
| 4 | -1.83 | POW | -2.45 | -1.89 | -1.80 | -3.6 |
| 6 | -0.88 | POW | -0.78 | -0.80 | -0.77 | -1.4 |
| 8 | -1.76 | POW | -2.33 | -1.98 | - | |
| | | LOG 1 | - | - | -1.56 | -2.2 |

It can be seen that beside the LOG 1 model the POW law is also dominant for some samples. This is believed to be a consequence of the chrome base's diffusion into the gold part of the electrode. This hypothesis was verified by an Auger analysis of the electrode film.

Fig. 1 shows the relative prediction error $\Delta P$ for two years' aging, as dependent on the period of time upon which the approximation was based.

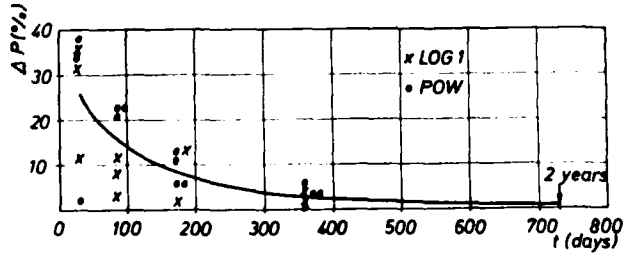



Fig. 1 Relative prediction error.

Fig.2 depicts the application of various models for approximating the aging of a single unit.

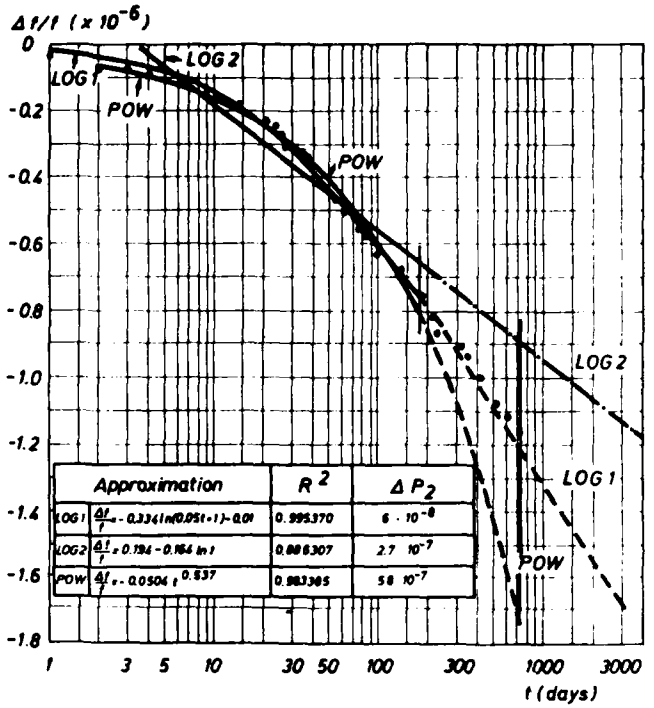



Fig. 2 Application of various models to unit No.1; $\Delta P_2$ - error of two years' aging prediction.

Aging diagrams, including measured values and their approximations, are given in Fig. 3 for two characteristic samples.

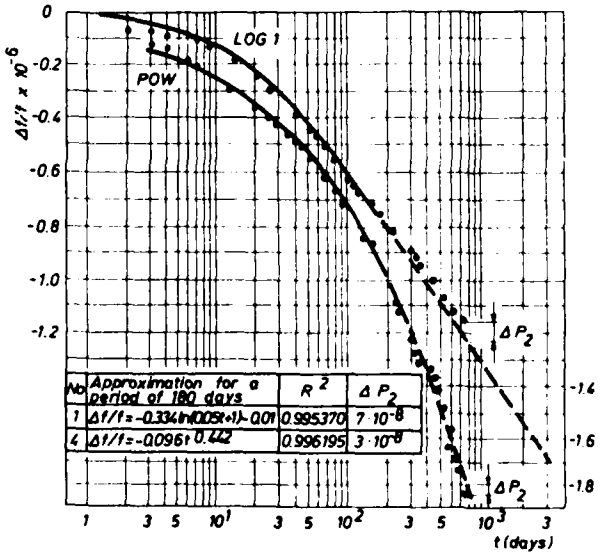



Fig. 3 The best approximation for samples No.1 and No.4.

The correlation coefficients $R^2$ are very high, ranging from 0.974 to 0.996 in the optimum approximations.

Aging prediction for a period of 10 years is given in Table 2, based on constants obtained for an approximation period of 6 months.

The other type of crystal unit tested were high-precision 4332 kHz units for OCXO applications, utilizing a polished quartz plate with gold electrodes, enclosed in vacuum coldweld holders. The frequency was measured for over a year, with the units in constant operation in OCXOs at a temperature of about 75°C. The frequency change was approximated only by the LOG 1 law[6,7] and curve fitting was carried out for periods of 1,3 and 6 months. The aging prediction for a period of one year was compared to the aging which actually took place.

Aging diagrams for two typical samples are given in Fig. 4. The determination coefficients ranged from 0.995 to 0.998.

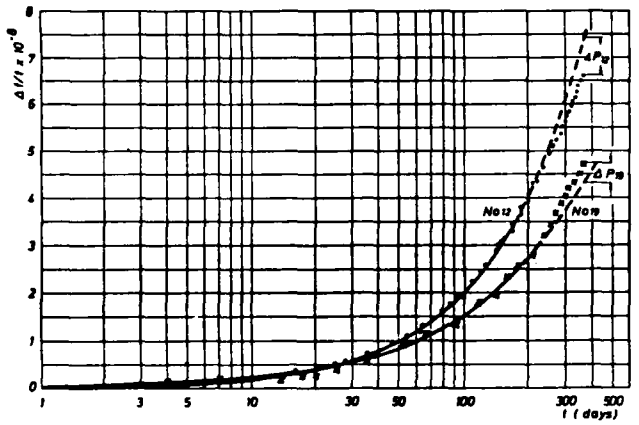



Fig.4 Aging prediction for a one-year period for two samples of high-precision crystal units.

## The accelerated aging method

Aging prediction by the accelerated aging method is based on the assumption that aging of quartz crystal units is faster under severe conditions than under normal operating conditions. Severe conditions may be higher drive levels or, especially, higher temperature.

Contradictory opinions regarding this assumption can be found in papers published so far. On one hand, there is a substantiated argument that "it is virtually impossible to predict the aging at a temperature other than the one at which the aging is actually performed"[6], or that the statement "the higher the temperature the higher the aging rate" is not correct[8]; or, a somewhat more moderate assertion "that the aging is not primarily due to a single thermally activated process"[9]. On the other hand, papers can be found dealing with aging prediction for very long periods of time (20 to 25 years) based on aging measured over a relatively short period at temperatures substantially higher than those at which the device is to operate[10,11]. The results of our experiments, under certain conditions, support the second hypothesis, although we believe that these results may not be sufficient to resolve the dilemma.

The aging law can be generally represented by[12]:

$$\frac{\Delta f}{f}(i, T, t) = R(i)\,R(T)\,R(t) \qquad (10)$$

where $\frac{\Delta f}{f}$ is the relative frequency change, R(i) is the dependence on drive level, R(T) the dependence on temperature and R(t) the dependence on time, given by one of the functions in Table 1.

In our work we have assumed that R(i) may be neglected and that the R(t) function may be very reliably taken to be the LOG 1 aging law, so Eq. 10 may take the form

$$\frac{\Delta f}{f}(T,t) = R(T)\ln(1+bt) \qquad (11)$$

where $R(T) = C\,e^{-\frac{E_a}{k_B T}}$ is Arrhenius's law for stress relaxation or sorption process rates, $E_a$ is the activation energy of the aging process, $k_B = 8.616\cdot10^{-5}$ eV/K - Boltzmann's constant and b, C - constants. If $bt \gg 1$, which is always the case after a sufficient period of time t, Eq.(11) simplifies to:

$$\frac{\Delta f}{f} = R(T)\ln b + R(T)\ln t \qquad (12)$$

If the first term of this equation is negligible, which was shown to be correct in our case (Fig. 6a and Fig. 6b), the final aging equation, which we have employed for aging prediction, is

$$\frac{\Delta f}{f} = C\,e^{-\frac{E_a}{k_B T}}\ln t \qquad (13)$$

Let $T_L$ be the operating temperature of the crystal unit under normal operating conditions, $T_H$ the elevated test temperature, $t_L$ and $t_H$ the respective periods of time and $\delta f_L = \frac{\Delta f}{f_L}$ and $\delta f_H = \frac{\Delta f}{f_H}$ the respective frequency changes at the temperatures $T_L$ and $T_H$ after periods of time $t_L$ and $t_H$, where $t_L > t_H$. Based on Eq. (13) we obtain

$$\Delta = \frac{\ln t_H}{\ln t_L}\,e^{\frac{E_a}{k_B}\left(\frac{1}{T_L}-\frac{1}{T_H}\right)} = \frac{\ln t_H}{\ln t_L}\,A \qquad (14)$$

where $\Delta = \frac{\delta f_H}{\delta f_L}$ is the ratio of the frequency change at a higher temperature over a shorter period of time to the frequency change

at a lower temperature over a longer period of time. The quantity A is called the acceleration factor.

Equation (14) makes it possible, if the activation energy $E_a$ is known, to calculate the aging at a lower temperature over a longer period of time, having beforehand measured the aging of units of the same type at a higher temperature over a shorter period of time. The activation energy of the aging process is basically the resultant of the activation energies of several primary processes which cause aging.

The activation energy is calculated from the aging data at two or more temperatures. If only two temperatures are used, Eq. (14) yields the expression for the activation energy

$$E_a[eV] = \frac{8.616 \cdot 10^{-5}}{\frac{1}{T_L} - \frac{1}{T_H}} \ln \left( \frac{\delta f_H}{\delta f_L} \frac{\ln t_L}{\ln t_H} \right) \cdot \quad (15)$$

The aforesaid demonstrates the importance of knowing the activation energy for application of this method of aging prediction. It is our impression that the available data on activation energy determined on the basis of crystal aging is not sufficient. Based on aging data at different temperatures, found in papers published so far, we have calculated, with some freedom of interpretation, by applying expression (15), the respective activation energies. The results are given in Table 3.

Table 3.

| | Reference | $T_H$ (°C) | $T_L$ (°C) | $t_H$ (days) | $t_L$ (days) | $\delta f_H$ | $\delta f_L$ | $\Delta$ | A | $E_a$ (eV) | Note |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Briese /13/ | 85 | 25 | 30 | 365 | - | - | 1 | 1.7 | 0.08 | (2) |
| 2 | Grata /14/ | 85 | 25 | 365 | 365 | - | - | 1.8 | 1.8 | 0.09 | (1) |
| 3 | Grata /14/ | 120 | 25 | 365 | 365 | - | - | 2.6 | 2.6 | 0.10 | (1) |
| 4 | Brice /3/ | 85 | 25 | 25 | 365 | - | - | 1 | 1.8 | 0.09 | (2) |
| 5 | Brice /3/ | 130 | 85 | 30 | 365 | - | - | 1 | 1.7 | 0.15 | (2) |
| 6 | MIL-C-XXX /16/ | 125 | 85 | 7 | 30 | $2 \cdot 10^{-6}$ | $2 \cdot 10^{-6}$ | 1 | 1.7 | 0.17 | |
| 7 | Olster /10/ | 137 | 91 | 142 | 142 | $125 \cdot 10^{-6}$ | $0.125 \cdot 10^{-6}$ | 10 | 10 | 0.65 | (3) |
| 8 | Feinberg /11/ | 120 | 75 | - | - | - | - | 2.6 | - | 0.25 | (4) |
| 9 | Filler /8/ | 90 | -5 | 30 | 30 | $2 \cdot 10^{-10}$ | $6 \cdot 10^{-11}$ | 3.3 | 3.3 | 0.1 | (5) |
| 10 | Warner /15/ | - | - | - | - | - | - | - | - | 0.04 | (6) |

Notes:

(1) $\delta f_H / \delta f_L$ are mean values for 600 samples of 8 types of units

(2) $\delta f_H / \delta f_L = 1$ means the authors claim that aging at a higher temperature is equal to aging at a lower temperature

(3) Median aging of lower 3 dB frequency of 8 MHz A6 monolithic filters

(4) Surface acoustic wave filter 215.6 MHz

(5) $\delta f_H$ and $\delta f_L$ are aging rates per day on the 30th day, estimated mean values of 25 units

(6) $E_a$ calculated from the fact that $Q = 1 \frac{kcal}{mol}$

It can be seen that $E_a$ values vary over a wide range.

In our experiment we have calculated the activation energy for two types of crystal units, having frequencies of 4000 kHz and 4194 kHz. The former type belongs to the TCXO class while the latter is a microprocessor crystal. We applied the isothermal method at temperatures +25°C, +85°C and +125°C. Five units of each type were tested at each temperature. The units were constantly at the mentioned temperatures, with a stability of ±0.1°C, for the duration of the test. The test lasted 55 days or approximately 1300 h. The frequency was measured by the IEC-444 zero phase technique. The units were operated only during measurement.

The activation energy was calculated from the expression $E_a\ [eV] = -8 \cdot 616 \cdot 10^{-5} \cdot B$, where $B = -\frac{E_a}{k_B}$ is the slope of the function $\ln \frac{\Delta f}{f} = \varphi(\frac{1}{T})$:

$$\ln \frac{\Delta f}{f} = \ln C - \frac{E_a}{k_B} \cdot \frac{1}{T} \quad (16)$$

which was obtained by a linear fit of three points (Fig. 5).

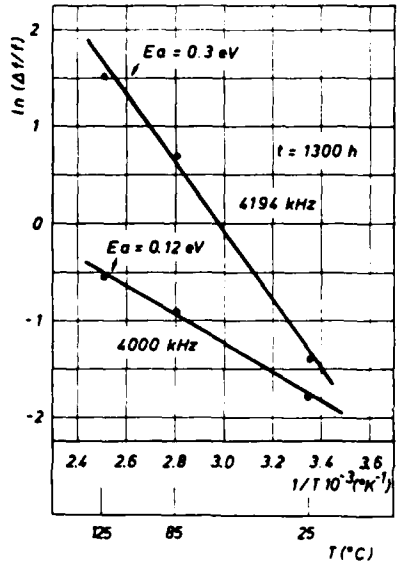



Fig.5 Activation energy calculation.

The measurement results are given in Table 4.

Table 4.

| Frequency (kHz) | 4000 | | | 4194 | | |
|---|---|---|---|---|---|---|
| Holder | HC-36 | | | HC-49 | | |
| Temp. ($^{o}$C) | +25 | +85 | +125 | +25 | +85 | +125 |
| $\Delta f/f$ ($10^{-6}$) | -0.18 | -0.4 | -0.6 | -0.25 | -2 | -4.6 |
| B | -1422 | | | -3500 | | |
| $E_a$(eV) | 0.12 | | | 0.3 | | |

Figs. 6a and 6b depict the aging curves for both cases. Each point in the diagram is the mean value of five samples. The curves were fit by the function $\frac{\Delta f}{f} = a+b\cdot \ln t$.

The relatively high values of the determination coefficients ($R^2$=0.90÷0.98) and $a \approx 0$ show the assumption that these curves may be represented by an ordinary LOG function (13) to be valid.

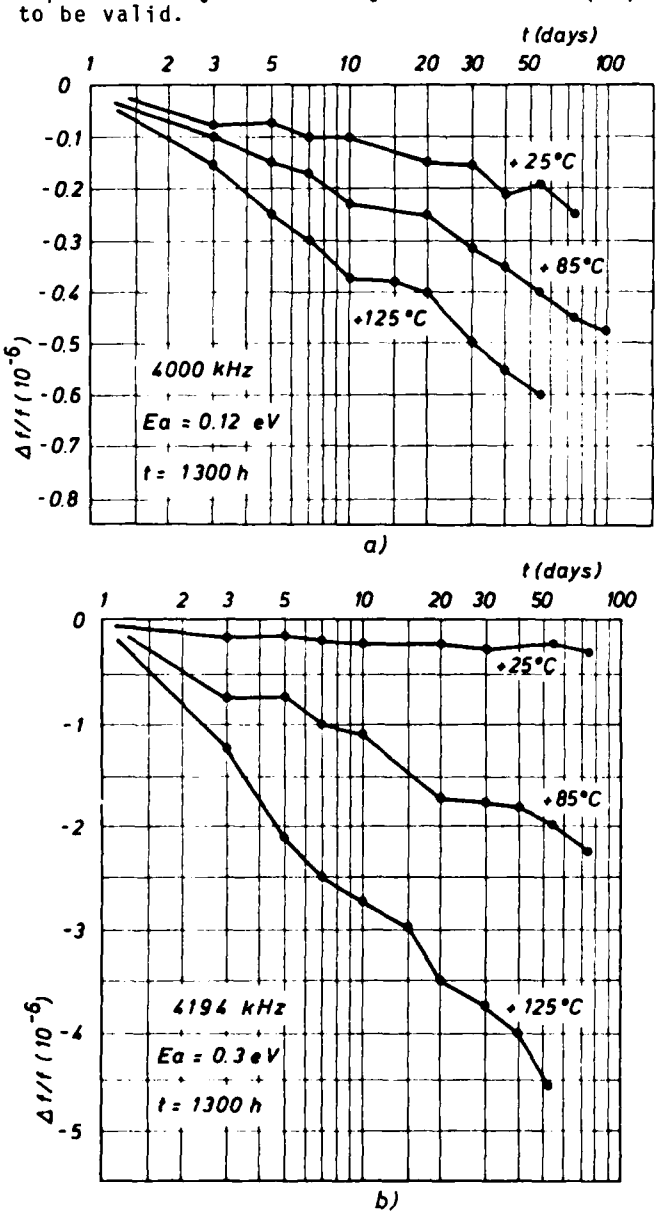



Fig.6 Aging curves.

Diagrams of the aging ratio $\Delta$ versus $E_a$, calculated using Eq.(14), are given in Fig. 7.

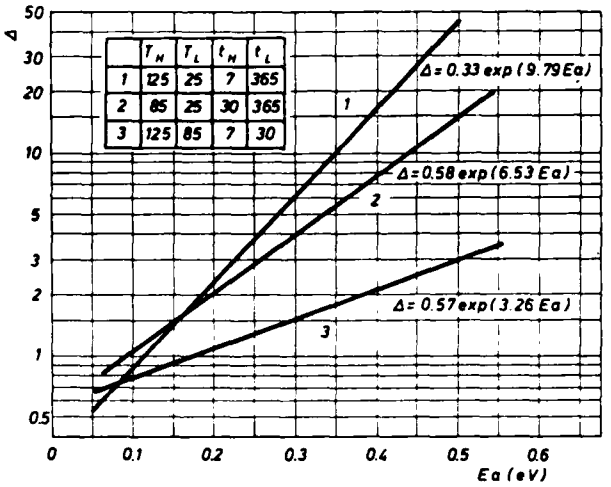



Fig.7 Aging ratio $\Delta(E_a)$.

For example, aging of crystal units having an activation energy of 0.2 eV will be, at +125$^{o}$C during 7 days and +85$^{o}$C during 30 days approximately twice ($\Delta \cong 2$) the aging at +25$^{o}$C during 1 year.

The application of the accelerated aging method expounded above is possible only if the activation energy of the particular crystal unit type is known and if the LOG 2 aging law can be applied with adequate precision.

In order to verify the above calculations, the aging curves of the 4000 kHz and 4194 kHz units at 25$^{o}$C were extrapolated to 1 year. This yielded aging of $-0.3\cdot10^{-6}$ and $-0.42\cdot10^{-6}$ respectivaly. The $\Delta$ factor can now be calculated and compared to the values in Fig. 7, for $E_a$=0.12eV and $E_a$=0.3 eV. Thus, for example, the measured factor for $E_a$ = 0.3 eV for the +125$^{o}$C/7 days aging process, in relation to the extrapolated +25$^{o}$C/1 year aging, according to Fig. 6b, is $\Delta$= 5.7. The calculated factor for the same case (Fig. 7) is $\Delta$= 6.2. Other combinations also show good agreement.

The procedure for the application of the accelerated aging method consists of determining $E_a$ based on a larger number of aging measurements at several temperatures and over longer periods of time. The $\Delta$ factor is then calculated for the specified accelerated aging test conditions $T_H$ and $t_H$ and the specified normal aging conditions $T_L$ and $t_L$. Crystal units which do not satisty the condition $\delta f_H \leq \delta f_L$, where $\delta f_H$ is the measured aging in the accelerated process and $\delta f_L$ the allowed aging over a longer period, do not meet the specified requirements.

## Conclusion

The results obtained show that under certain conditions crystal unit long-term frequency stability can be predicted. The approximation method can also point out the dominant mechanism influencing aging, which, in turn, can be utilized for control and correction

of the technological process. The application of this method is especially important for high-precision units. The accelerated aging method is usually employed in routine quality control. Besides yielding the activation energy value, this method also enables aging prediction over a longer period of time.

## Acknowledgements

The authors would like to express their gratitude to B.Milić for initiating the systematic study of the problem of aging prediction, to M.Djuričić for the organization of the high-precision units' aging tests and Miodrag Stojanović, Milan Stojanović and Z.Tepavac for their endless patience and efforts in measuring and collecting aging data.

## Appendix

### Curve fit program

This program fits curves to data by an arbitrary function, using a modified Simplex algorithm[17], which is essentially a mixture of the Newton-Ralphson and steepest-descent methods. It is based on creating a simplex (a geometrical figure which has one more vertex than the space in which it is defined has dimensions) in the space described by the fitting parameters and then substituting the "worst" vertex (the one where the sum of squared residuals is the largest) with a new one, according to four mechanisms: reflection, expansion, contraction and shrinkage, until the lowest sum of squares is reached. In actual implementation there are two termination criteria: maximum number of iterations and required accuracy of parameters and sum of squares (accuracy being defined as the difference between values in two successive iterations). The advantages of the algorithm are: a) the fitting function can have any number of parameters and variables, as well as any (finite) number of discontinuities; b) divergence is impossible and c) there is no numerical differentiation or matrix operation involved, which minimizes rounding-off errors and speeds up the program. The flowchart is given in Fig. A1[17].

The program we use runs in VMS BASIC on a VAX 11/730 minicomputer and has been extensively tested for functions of one and two variables with up to nine parameters and 250 data points. It has been proven to usually converge in 20 $n^2$ iterations, n being the number of parameters, with the termination criteria relaxed so only one of the parameters need reach the accuracy required since by that time the others are usually very close to it also. The accuracies commonly employed are $10^{-4}$ to $10^{-6}$, with the accuracy for the sum of squares being an order of magnitude higher than that of the parameters. The execution time on our machine is typically 30s for 500 iterations of a three-parameter function fit with an average user priority. The program output consists of the number of iterations, the parameter values and their estimated errors, the input and calculated data points and their difference, the standard deviation and the determination coefficient.

The two most frequently encountered problems were found to be: a) reaching a point where the function is not defined (log of negative number, division by zero, etc.); changing the starting point (the initial simplex) several times may be necessary to find a "route" which "avoids" the undefined points, or, alternatively, inserting a conditional branch into the program for the particular function involved, and b) obtaining practically the same determination coefficients with widely differing sets of parameters. This simply means that the sum of squares has several minima and that the adequate set is to be chosen according to additional criteria.

Owing to the versatility of the algorithm, many other features can be implemented, such as contour conditions, parameter weighting, modifying the coefficients for reflection, expansion and contaction, etc. We have found the program to be a reliable, efficient and very useful tool.

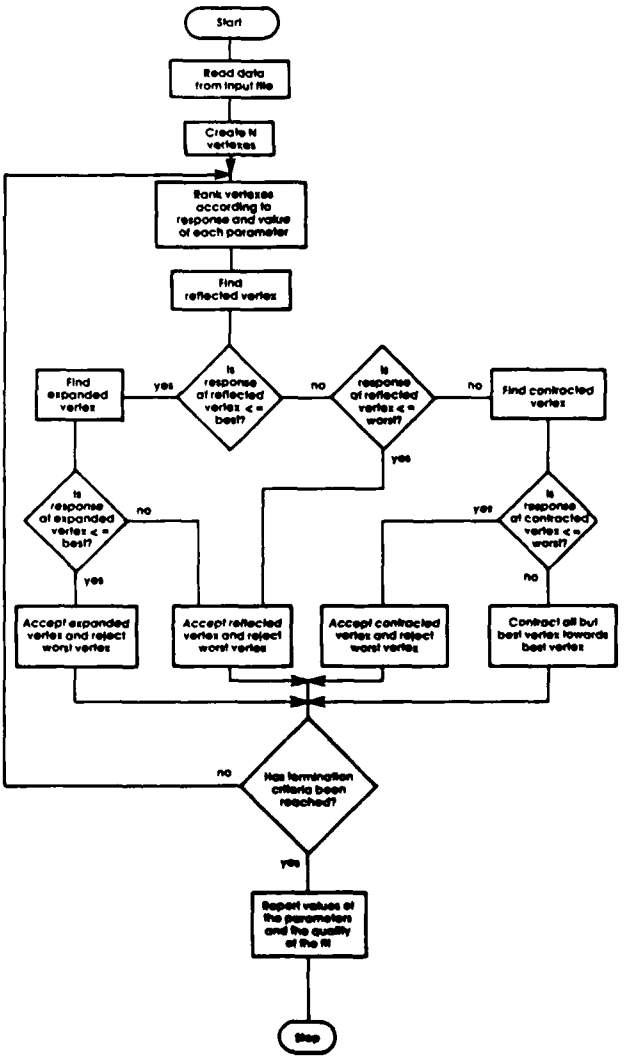



Fig. A1 Computer program flowchart.

## References


/1/ A.W.Warner,D.B.Fraser and C.D.Stockbridge: "Fundamental Studies of Aging in Quartz Resonators", IEEE Trans.Sonics Ultrason., SU-12,1965.,pp 52

/2/ T.E.Parker: "Analysis of Aging Data on SAW Oscillators",Proc.34th AFCS,1980., pp 292

/3/ J.C.Brice: "Crystals for quartz resonators", Reviews of Modern Physics,Vol.57, No, 1, Jan. 1985. pp 105

/4/ J.Zelenka: "A Note on the Ageing of High Precision Piezoelectric Crystal Units" Czech.J.Phys. B 16, 1966. pp 437

/5/ G.L.Dybwad: "Aging Analysis of Quartz Crystal Units with TiPdAu Electrodes", Proc.31st AFCS,1977.,pp 144

/6/ R.Filler:"Aging specification, measurement, and analyses" Proc. 7th Quartz Devices Conference and Exhibition, 1985. pp 93

/7/ R.Filler: "The New Military Specification for Precision Quartz Crystal Units", Proc. 7 th Quartz Devices Conference and Exhibition, 1985, pp 7

/8/ R.L.Filler et al.: "Aging Studies on Quartz Crystal Resonators and Oscillators", Proc.38th AFCS,1984.,pp 225

/9/ R.Filler et al. "The aging of resonators and oscillators under various test conditions" Proc 41st AFCS, 1987, pp 444

/10/ S.H.Olster et al.:"A6 Monolithic Crystal Filter Design for Manufacture and Device Quality",Proc.29th AFCS, 1975., pp 107

/11/ A.A.Feinberg: "Parametric failure rate model for quartz crystal device aging with application to surface acoustic wave filters" Proc 41 st AFCS, 1987, pp 360

/12/ D.Epstein: "Application and Use of Acceleration Factors in Microelectronics Testing" Solid State Technology, Nov. 1982., pp 116

/13/ W.Briese:"Aufbau,Eigenschaften und Anwendung von Schwingquarzen", Tagungs-Dokumentation, Quarz-Symposium '85, Berlin,Vistas Verlag 1986.,pp 23

/14/ K.Grata: "Long term crystal stability study" Proc 5th Quartz Crystal Conference, 1983, pp 214

/15/ A.W.Warner and C.D.Stockbridge: "Quartz Resonators;Reduction of Transient Frequency Excursion Due to Temperature Change", Journal of Appl.Physics,Vol.34., Number 2.,Feb,1963.,pp 437

/16/ MIL-C-XXX: "Crystal Units, Quartz, Precision, General Specification for".

/17/ M.S.Caceci, W.P.Cacheris. "Fitting Curves to Data", BYTE, May 1984, pp 340

42nd Annual Frequency Control Symposium - 1988

# EXPERIMENTAL RESULTS ON AGING OF AT-CUT STRIP RESONATORS

J. R. Gehrke and R. Klawitter
Motorola Inc.
2553 N. Edgington St.
Franklin Park, Illinois 60131

## ABSTRACT

Aging in a crystal unit is generally defined as any changes in the operating parameters as a function of time. More specifically here it is defined as the time dependent change in frequency. This paper describes the results of experiments designed to determine the primary causes of aging in AT-cut strip resonators operating in the thickness-shear mode of vibration.

The crystal units used were 16 MHz AT-cut rectangular blanks cantilever mounted on glass/metal bases using epoxy and vacuum sealed in 3x8 mm tubular packages. The crystal units were heat treated at various temperatures to induce accelerated aging and then measured at 25 degrees Celsius in a temperature controlled chamber.

Accelerated aging at temperatures 125 degrees C and above showed an initial positive frequency shift of short duration followed by a dominant negative shift. Units treated at temperatures lower than 125 deg C showed only positive frequency shifts. This temperature dependent negative aging implied that a mechanism was involved which had a minimum threshold energy of activation.

A series of experiments were conducted to establish the mechanism(s) which caused this type of aging. The results of the experiments indicated that the cause of the aging was dominated by first stress relief followed by aging due to contaminants adhering to the crystal.

The extraneous stresses applied to the crystal were found to come from two sources. There were stresses induced in the blank during the sealing process due to the force fitting of the shield over the base of the crystal holder. In addition there were stresses induced in the blank which were developed during the epoxy cure process. These stresses when relaxed over time cause a frequency shift due to the non-linear stress-frequency effect. The time constant for this stress relief is relatively small compared to most aging processes in crystals so that this type of aging can be easily eliminated by a pre-age during the manufacturing process.

The rate of negative aging was observed to be dependent on the type of conductive epoxy used. In addition, the rate of aging was dependent upon the temperature at which the crystals were heat treated. With one epoxy there was evidence of a minimum threshold temperature necessary to induce the negative aging.

## INTRODUCTION

Aging is defined as a change in frequency relative to the initial frequency of the unit. If the crystal enclosure is assumed to be hermetic then the aging is caused by mass sorption and/or stress recovery of the crystal. The mass sorption can be in the form of chemisorption, desorption or oxidation. By using the Kinetic Theory of Rate and assuming that the stress is directly related to the frequency then we have the relationship for change in frequency as a function of time at constant temperature

$$(\text{del } F)/F = A + B*\ln( 1 + T/To )$$

where A, B and To are constants.[1] This shows the well known relationship that change in frequency is logarithmic with time.

This logarithmic relationship fails to describe the type of aging which was exhibited in this series of experiments. When these crystals were aged at elevated temperatures two dominant mechanisms were apparent in driving the frequency shifts of the devices. These mechanisms were opposite in their effect on change of frequency and had different time constants. The mechanism with the long time constant incurred negative frequency shifts while the mechanism with the short time constant incurred positive shifts.

In order to investigate the behavior and cause of these frequency shifts, a series of experiments were performed. Each of the experiments consisted of 10 to 20 crystals per group. The resulting aging plots show only the mean frequency shifts of the groups. Analyses of variance were performed to determine the level of separation between groups.

The first experiments investigated the short term positive aging. The first experiment attempted to isolate the aging due to stresses incurred in the mounting of the crystal. The second experiment showed that positive aging can arise from relaxation of stresses incurred during the package sealing operation. A further study indicated that these sealing stresses are transmitted via the nonconductive epoxy.

The last series of experiments investigated the behavior of the long term negative aging as a function of temperature. In addition, two epoxy related experiments were done to attempt to determine the source of the negative aging.

CH2588-2/88/0000-412 $1.00 C 1988 IEEE

## The AT-Cut Strip Resonator Crystal

The type of quartz crystal used for the aging studies described here is a 16.8 MHz fundamental thickness shear mode device. The vibrating plate is rectangular with length 0.24" and width 0.07". The quartz resonator has an electrode covering about 70% of the plate area. (See figure 1.) The electrode was gold with a thin sublayer of chrome. The crystal blanks were fabricated using photolithographic techniques from a 2" x 2" wafer that had been mechanically lapped and chemically deep etched. All the crystals are mounted on a Toyo Dempa IBP glass metal base using the non-conductive epoxy Amicon A-329-1. The electrical connections to the crystal were made using either Amicon C-850 or Ablebond 71-1 epoxies (here the term epoxy is used to include polyimides). The control groups for each experiment consisted of crystals mounted using the Amicon C-850 and A-329-1. All crystals were sealed using a cold-weld shield supplied by Toyo Dempa.

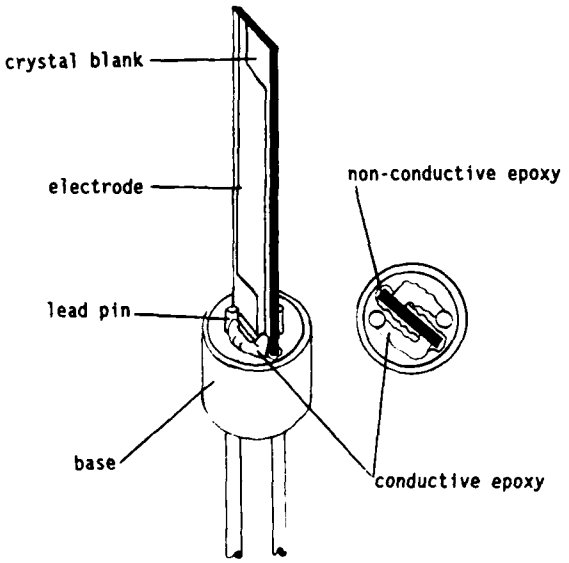



**FIGURE 1**

## CRYSTAL MEASUREMENT

The measurement system is the critical link in any aging experiment. The measurement system must be capable of accurately and repeatably measuring crystals over a time frame of several months to several years. The system assembled for this series of experiments has that capability.

### Measurement system hardware

The measurement system consists of a Hewlett-Packard 3577 network analyzer with S-parameter test set, a Tenny Jr. temperature chamber, a Hewlett-Packard 2804A quartz thermometer, a Hewlett Packard series 300 computer and some miscellaneous switching equipment.

The fixture that holds the crystals during measurement has a capacity of 48 crystals and employs pin diodes to switch from port to port. Reflection measurements are used to establish crystal electrical parameters, utilizing the network analyzer and its full 3-term error correction. Using impedance measurements, the series resonant frequency of the crystal was taken as the point of zero reactance.[2] Measurements were made in accordance with the EIA 512 standard. A 26 point, 2nd order circle fit was used to determine the crystal motional parameters. Below is the block diagram of the system.

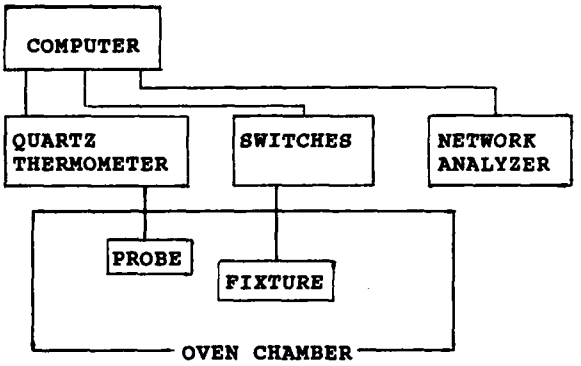



**Figure 2**

### Measurement Error

The error associated with measuring the frequency of the crystals falls into two categories, systematic error and random error. The systematic error arises from the differences in line length, capacitance and inductance between the first crystal test fixture, which is used to calibrate the network analyzer, and the other fixtures linked in the switching network. Also, there is systematic error introduced by calibration error and off frequency calibration. Off frequency calibration occurs when a crystal's frequency fall outside of the calibration window. All of these systematic errors are removed by normalizing to each crystal's initial frequency reading and then analyzing the behavior in terms of relative frequency shift. In order to achieve this, care must be taken to place each crystal in its respective port for each subsequent measurement.

The random error comes from the network analyzer, the fixture and from fluctuations in crystal temperature. The analyzer error is specified at 0.02 db and 0.2 degrees.[3] Also, associated with the analyzer is the error due to sweep time (0.1% of frequency span at 1 second) and the error related to the curve fit of reactance. Fixture error is caused largely by changes in the crystal position in the port between measurements.

The last source of error is differences in the measurement temperature of the crystal between measurements. The nature of the AT-cut strip resonator is to have a frequency-temperature relation which is roughly described by a third order polynomial. The slope of the frequency curve as a function of temperature at 25 degrees Celsius varies from crystal to crystal. By taking the steepest slope as the worst contributor to the error in frequency, the tolerance on the temperature can be determined. The worst case slope for the crystals used in these experiments was 1ppm/deg C and the oven stability was +/- 0.03 degrees Celsius resulting in an estimated error of 1 Hz.

These random errors combine to result in an error in the repeatability of the measurements. To determine the repeatability a group of crystals was measured ten subsequent times with removal from the fixture after each measurement cycle. The sigma for the normalized frequency is 15.75 Hz with a mean of 0.66 Hz. Since the errors are random, they result in a spreading of the population and an uncertainty in the mean of the population. By weighting the mean reading for each crystal by the sigma of its measurements a weighted mean for the population can be calculated along with an uncertainty in that mean. The resulting weighted mean is 0.65 Hz with an uncertainty of 0.32 Hz. Since the mean and its uncertainty are smaller by an order of magnitude than the changes in the mean that are observed during these aging experiments, it is safe to assume that the results are reliable in terms of the measurement error.

## STRESS RELAXATION AND AGING

It is known that stresses in a quartz resonator can change its frequency as the external stresses change. This nonlinear stress-frequency effect has been investigated experimentally and theoretically [4,5]. Previous theoretical work has related stress-frequency effects to the higher order elastic constants of quartz[6]. Here we are concerned with stress relaxation in strip resonator crystals and its effect on frequency aging. The frequency change due to the application of an external stress can be either positive or negative depending on the orientation of the applied stress with respect to the crystal planes and is approximately linearly proportional to the applied force[5]. In our studies we found the stress related component of resonator aging to give rise to positive frequency aging.

### Stresses Due to Mount and Seal

The mounting structure of a crystal can apply stresses to the crystal to cause frequency shifts. The mounting scheme used for these experimental units had both a nonconductive cement for structural support and conductive cement to make electrical contact with the crystal (see Figure 1). There are two ways in which the mounting structure of the crystal unit may impose a stress on the crystal. During the cement curing operations stresses may be locked in as the cement cures with a resulting strain in the crystal. This condition will relax over time and so cause a frequency shift. The sealing operation may also present an external stress to the crystal. The package used to house this crystal is solder sealed by compression fit of a solder coated shield to a gold flashed crystal base. We found that the force fitting of the shield to the base causes a frequency shift that relaxes over time and so contributes to aging.

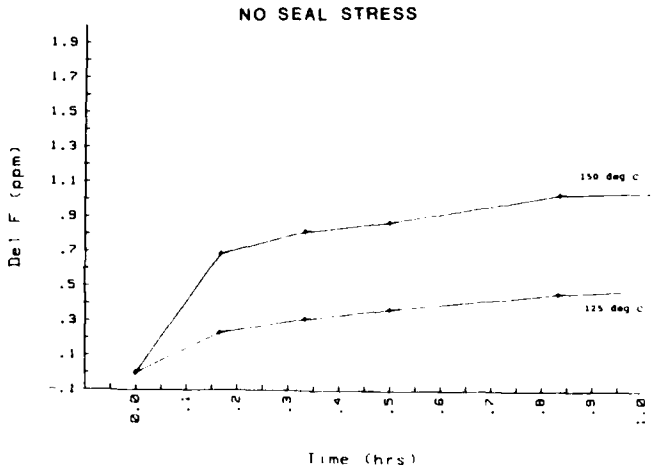



**FIGURE 3**

**Stresses Incurred at Mount.** To isolate the stress-frequency effect due to mounting alone from that due to sealing, we mounted a group of standard crystals in the usual way as shown in Figure 1. Instead of sealing the crystals with standard 3x8 mm tubular shields, the cylindrical bases with crystals were mounted on HC-6 crystal bases so that stresses to the crystals would not be incurred during sealing of the HC-6 cans. Figure 3 is the plot of mean accelerated aging at 125 and 150 degrees centigrade for these units. The plot shows initial positive aging which, as will be seen, is indicative of aging due to stress relief for the crystal units of our studies. At this time we cannot say with certainty that the aging shown in Fig. 3 was due to mounting stress relief. Earlier studies by Belser and Hicklin [7] showed that Cr-Au plated AT-cut crystal units exhibited positive aging while units with other electrode metals could show negative aging. The metallizations of the units in our studies were also Cr-Au so that the initial positive aging we saw may be due to effects other than mounting stress relaxation. Further studies could clear up this point. Obvious stress relaxation effects were found in the study described below.

**Stresses Incurred at Package Seal.** Stresses are superposed on the crystal during package sealing via the mounting structure. The shield is compression fitted to the base during sealing and a frequency shift before and immediately after this operation can be measured. This frequency shift due to stress effects was negative so subsequent short term aging due to relaxation was positive.

In one study we aged strip crystal units at 150 degrees centigrade after sealing in vacuum. Half the units served as an aging "control" group while the others underwent several cycles of aging, then opening the crystal package, re-sealing, then more aging. This latter group is labeled the "re-seal" group in Figure 4. The plot shows dramatically the effect of stresses induced during the package sealing operation. (The effect can be seen without opening and re-sealing the package by simply applying suitable forces to the crystal base while monitoring crystal frequency.) The two groups begin with similar aging curves including short term positive aging followed by longer term negative aging. However, the "re-seal" group shows a burst of positive aging after each re-sealing cycle. It is clear that the sealing operation causes forces to be transmitted to the crystal which affect the resonator frequency. As the stresses induced during sealing relax the resonator frequency changes and so provides an important though short term contribution to the overall aging of the crystal unit.

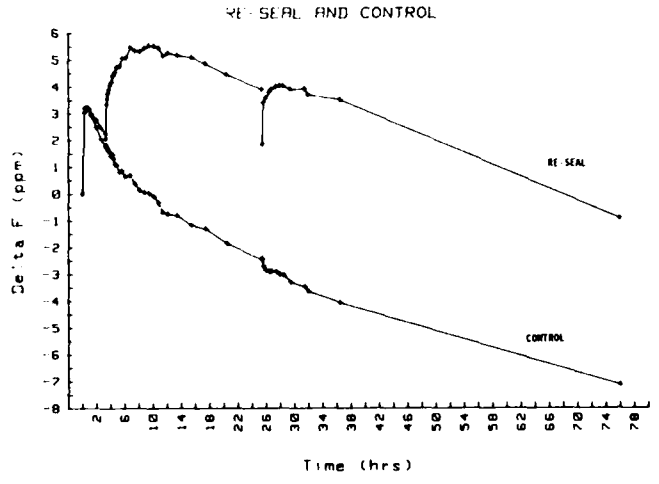



**FIGURE 4**

We found that the structural, nonconductive cement rather than the conductive cement transmits most of the sealing stress which causes the short term positive aging common to the crystal type of this report. In one study a group with the electrical connection from crystal to package lead was made by wire bonding and compared to a control group with the connection made by conductive cement. The wire bonds were placed from the leads of the crystal base to the electrode tab regions where normally the epoxy electrical connections are made. Both the wire bonded group and the control group were mounted with a nonconductive cement to secure the cantilevered crystal to the base. Both groups were accelerated aged at 150 degrees centigrade. If the conductive cement contact to the crystal was transmitting most of the sealing stress then the wire bonded crystals should have shown less stress related (positive) aging since the .0015" diameter wire bonded to the electrode tab could not exert the same magnitude forces that cured cement could. Our results showed, however, that both the wire bonded and control groups have a similar amount of positive, stress related aging. Figure 5 shows the short term aging of the wire bonded and control groups. Though the short term (positive) aging of the two groups is similar an analysis of variance shows the groups have distinguishable aging. The difference is most likely due to a negative aging component present in the control group caused by the conductive cement (see discussion on negative aging). We conclude that most of the stresses transmitted to the crystal during the sealing operation are due to the nonconductive, structural epoxy.

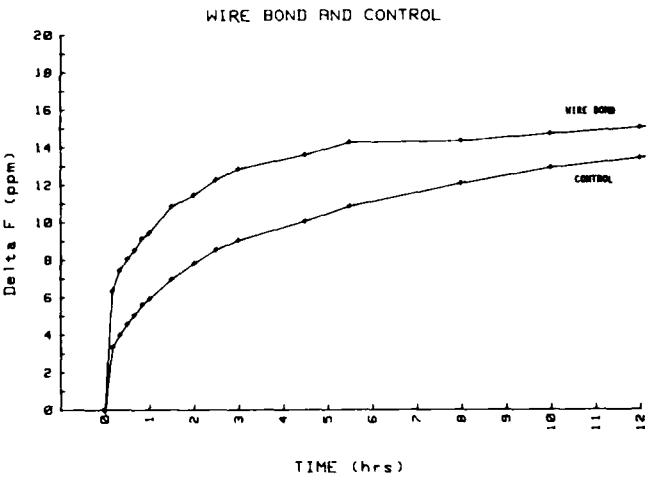



**FIGURE 5**

**Frequency Shifts Due to Stress Relief**

In one study we aged four groups of crystal units at different temperatures to determine the magnitude of the positive aging shift and more importantly the relaxation time for aging due to stress relief. Units were aged at 85, 100, 125, and 135 degrees centigrade. The short term aging plot (Fig. 6) of these four groups shows the initial positive frequency shifts which have been determined to be due to stress relaxation. The rate of the initial positive aging is, of course, temperature dependent though it is present at all temperatures. Negative aging, however, was observed only for temperatures above 100 degrees centigrade after the initial positive shifts.

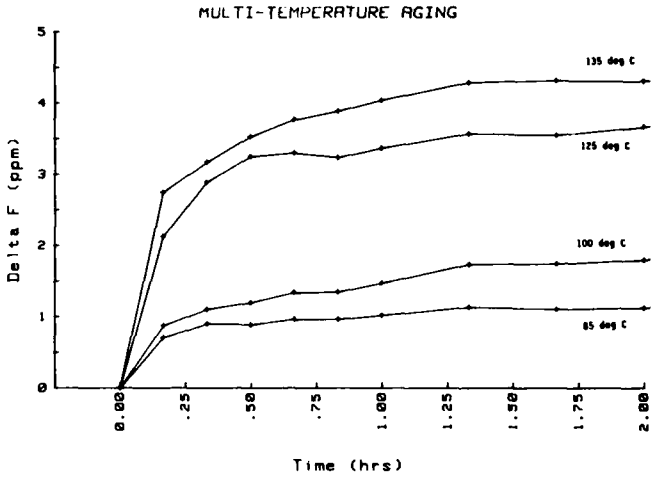



**FIGURE 6**

Figure 7 shows the results, beginning at 100 hours, for long term (almost 3000 hours) accelerated aging for the same four temperature groups of Figure 6. It can be seen that the long term behavior of the groups at 85 and 100 degrees is identical while those at 125 and 135 degrees have a negative aging "tail" which is much longer in duration than the positive aging portion of the curves. The positive aging is characterized by a relatively short time constant for the relaxation process.

Figure 8 is a plot of the same groups as Figures 6 and 7 but of intermediate duration (80 hours aging). It can be seen from Figure 8 that a "preage" of duration on the order of a few days would be sufficient to stress relieve and stabilize crystal units for aging temperatures between 85 and 100 degrees centigrade.

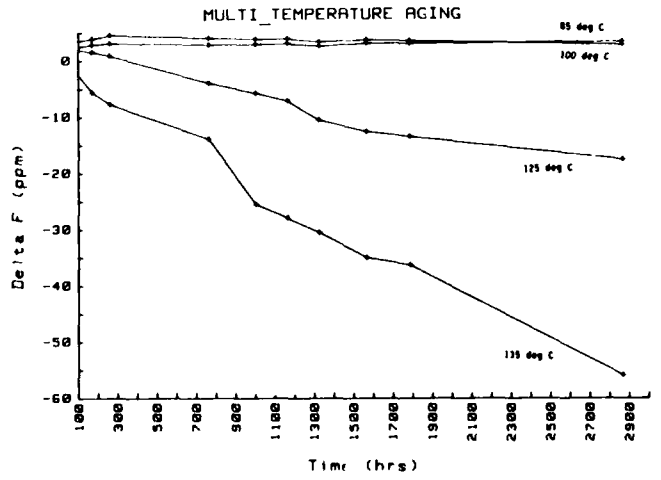



**FIGURE 7**

## NEGATIVE AGING

### Negative aging theories

There are several theories to explain the mechanism of negative aging. If we make the assumption that the crystal is hermetically sealed then we can eliminate leakage as a source of aging. The remaining potentially interrelated factors associated with negative aging are then electrode oxidation, chemisorbtion and contamination.

Given that the crystal is a closed system, the molecular source for the chemical activity or contamination must come from within the package. This leaves the package walls, residual gas or oils in the package, mounting epoxy and conductive epoxy as potential candidates. Gasses absorbed by the package walls can be desorbed under the condition of reduced pressure and subsequently absorbed by the crystal. The converse can also be true if the package has a higher getter characteristic than the crystal. During vacuum processing and vacuum sealing the backstreamed hydrocarbons from the oil diffusion pumps can be trapped in the package and eventually migrate either on or off the crystal. Lastly, the epoxy can give off gasses associated with decomposition or residual curing.

### Aging versus temperature

In order to establish the aging characteristics of AT-cut strip resonator crystals as a function of temperature and gain some insight into the aging mechanisms, several experiments were done. The first experiment investigated the effect of temperature on aging. The experiment consisted of four groups of crystals. The crystals in each group were processed simultaneously and then randomly selected for each group, thus insuring the statistical integrity of each group. The groups consisted of ten crystals, each of which had been processed by standard production procedures except that none had been final frequency adjusted. The final frequency adjustment had been eliminated in order to maintain the similarity of each group and not introduce another variable into the experiment. The crystals were all mounted using Amicon C-850. The groups frequency measurements were taken using the method previously described.

Once each group's frequency had been measured they were separated and placed in aging ovens of different temperatures. The four ovens were at 85, 100, 125, and 135 degrees Celsius. In order to resolve any short term aging effects (see the discussion on stress effect), the duration of the heat treatment was initially 10 minutes and lengthened to several hours. After each heat treatment the crystals were allowed to temperature stabalize at ambient temperature for about 30 minutes and then soaked prior to measurement for a minimum of 30 minutes in the measurement chamber at 25.05 degrees Celsius.

If we plot the mean shift in frequency for each group relative to its initial mean reading as function of time we have for the first 74 hours of aging the results shown in Figure 8. There are several points of interest in the results. After the first 8 hours of aging the 125 and 135 degree groups had changed from positive to negative aging while the 85 and 100 degree groups had not. In addition, at the 8 hour mark the 125 and 135 degree groups were placed in each others ovens (an error the experimenters will attribute to serendipity). This change caused each group's aging slope to switch and then subsequently return after they were placed back into their correct ovens at the 11 hour mark. These results point to two facts: that there is a threshold temperature between 100 and 125 degrees celsius at which the negative aging mechanism is activated and that once activated, the rate of aging is dependent upon the temperature. Figure 7 shows that even after 2800 hours of aging in the ovens the negative aging rate of the 125 and 135 degree groups has remained relatively constant as has the 85 and 100 degree groups. Notice how the 85 and 100 degree groups have maintained stable aging that is statistically equal in magnitude and near zero slope, thus showing that the negative aging mechanism has not been activated.

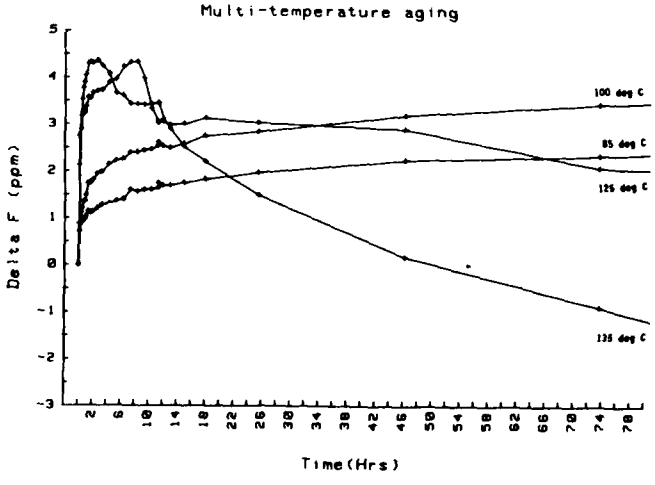



**FIGURE 8**

There is no separation between the 85 and 100 degree C groups while there is a significant separation between groups 125 and 135.

## Aging versus conductive epoxy

In the course of investigating conductive epoxy as a source for the stress related aging it was discovered that the type of conductive epoxy had a dramatic effect on the long term negative aging characteristics. Two groups of 20 crystals were processed together except for the application of conductive epoxy and its subsequent curing cycle. One group was fabricated using Ablebond conductive epoxy and the other group used Amicon C-850. The crystals were then aged at 150 degrees Celsius after the initial readings were taken.

Figure 9 shows that the Ablebond epoxy aged negative significantly worse than the Amicon C-850. This would point to the conductive epoxy as the major contributor to the negative aging. Again, since the aging is negative, the epoxy must be outgassing. In order to determine if the outgassing is due to incomplete epoxy curing a second experiment was performed.

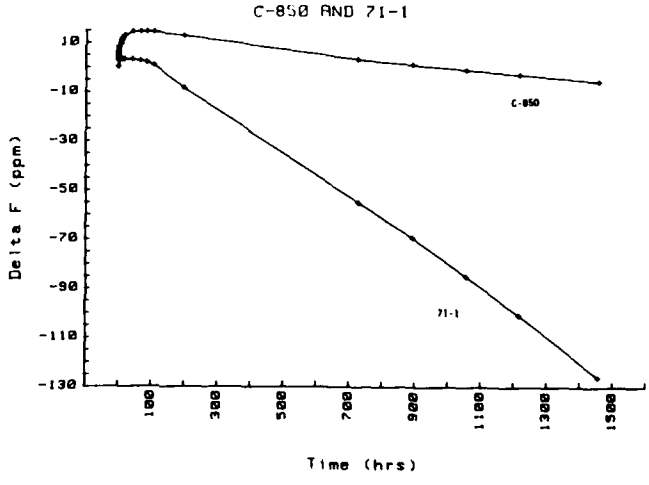



**FIGURE 9**

## ANOVA of Multi-Temperature Groups

In order to establish the confidence on the group separation an analysis of variance was performed (ANOVA). The analysis was done on the group populations at the 1568 hour mark. Table 1 shows the separation of the groups to a 99.5% confidence level (alpha of 0.005).

**Table #2: ANOVA**

| Group | Mean(ppm) | Standard Dev. | Seprtn |
|---|---|---|---|
| 85 C | 2.82 | 1.25 | A |
| 100 C | 3.64 | 1.76 | A |
| 125 C | -12.19 | 6.47 | B |
| 135 C | -34.30 | 10.73 | C |

Two groups of ten crystals were fabricated identical to the above defined method. Both the Amicon C-850 and the Ablebond 71-1 were then heat treated for 24 hours at 150 degrees Celsius prior to sealing. The intention of the heat treat cycle was to drive off any gasses associated with curing before sealing the crystal. Figure 10 shows the results of ensuing accelerated aging at 150 degrees Celsius. The Ablebond 71-1 group is now aging less negative than the non-heat treated Ablebond, while the Amicon is aging more negative. Table 2 has the slopes for the negative aging portion of each curve.

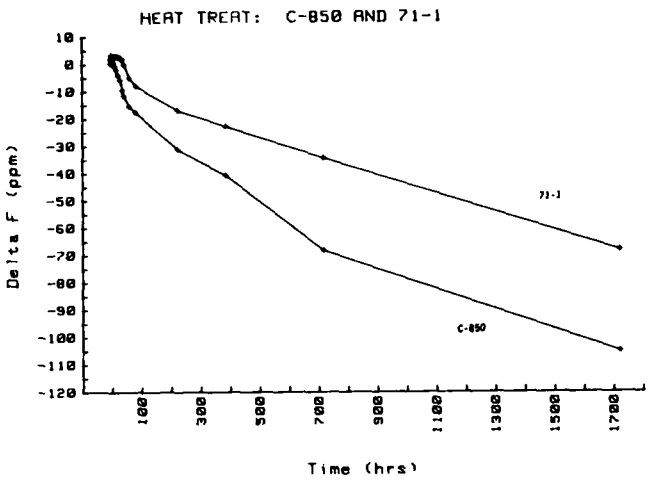



**FIGURE 10**

**Table #3: Aging Slopes**

| Group | Heat Treat | Slope (ppm/hr) |
|---|---|---|
| C-850 | No | -.015 |
| 71-1 | No | -.093 |
| C-850 | Yes | -.050 |
| 71-1 | Yes | -.033 |

The change in slope would imply that the initial negative aging of the Ablebond 71-1 was caused by the outgassing of the epoxy, since the aging rate was reduced by a factor of three with the addition of the heat treat cycle. The Amicon C-850 aging did not improve with heat treatment but worsened by a factor of three. After 1400 hours of aging the package lids were removed from several units of the non-heat treated groups. Visual inspection yielded no difference between these units and units that had not been aged. There was no apparent damage to the crystal or decomposition of the epoxy. These results suggest that the epoxy outgassing is not the only contributor to the negative aging mechanism. For completeness the following table shows the analysis of variance at the 108 degree Celsius point for each of the groups and there separations with an alpha of 0.005 (99.5% confidence). Again we find good statistical separation between the populations.

**Table #4: ANOVA Epoxy**

| Group | Mean(ppm) | Standard Dev. | Seprtn |
|---|---|---|---|
| No Heat Treat | | | |
| C-850 | 14.13 | 3.38 | A |
| 71-1 | 0.66 | 2.45 | B |
| Heat Treat | | | |
| C-850 | -9.59 | 5.45 | A |
| 71-1 | 1.55 | 2.63 | B |

## CONCLUSION

Our results have shown that the AT-cut quartz strip resonator can exhibit both positive and negative frequency shifts in the course of accelerated aging. Aging studies at 150 degrees C. show initial positive shifts due to stress relaxation followed by negative aging. The initial positive aging is of sufficient magnitude that the net resonator frequency change is positive in the early portion of the aging cycle. However, the time constant for the relaxation of the positive aging process is sufficiently small (on the order of a few days at temperatures, say, 85 - 150 degrees C.) that net negative aging soon becomes apparent. The negative aging process is not seen, however, at temperatures below 100 degrees centigrade. This means that in the manufacturing process a preage of relatively short duration is sufficient to stabilize the frequency provided the crystal operating temperatures remain below 100 degrees centigrade which is the usual case.

## REFERENCES

[1] Warner, A. W., Fraser, D.B., and Stockbridge, C.D. (1965) "Fundamental Studies of Aging in Quartz Resonators", IEEE Transactions on Sonics and Ultrasonics.

[2] Masini, B. (1986) "Crystal Measurement and Error Analysis using the HP3577 Network Analyzer with S-Parameter", Proc. 8th Annual Quartz Crystal Conf.

[3] Hewlett Packard HP3577 Operating Manual 03577-90000, 1983.

[4] Ballato, A.D. (1960) "Effects of Initial Stress on Quartz Plates Vibrating in Thickness Modes", Proc. AFCS 14, pp 89-114.

[5] Mingins, Barcus, and Perry (1962) "Effects of External Forces on the Frequency of Vibrating Crystal Plates", Proc. AFCS 16, pp 47-76.

[6] Smythe and Morely (1985) "Experimental Evaluation of the Effective Non-linear Elastic Constant For Trapped Energy and Contoured Resonators", Proc. AFCS 39, pp 317-324.

[7] Belser, and Hicklin (1969) "Comparison of Aging Performance of 5 MHz Resonators", Proc. AFCS 23, pp 132-142.

42nd Annual Frequency Control Symposium - 1988

# STANDARD TERMINOLOGY FOR FUNDAMENTAL FREQUENCY AND TIME METROLOGY

David Allan, National Bureau of Standards, Boulder, CO 80303
Helmut Hellwig, National Bureau of Standards, Gaithersburg, MD 20899
Peter Kartaschoff, Swiss PTT, R&D, CH 3000 Bern 29, Switzerland
Jacques Vanier, National Research Council, Ottawa, Canada K1A 0R6
John Vig, U.S. Army Electronics Technology and Devices Laboratory, Fort Monmouth, NJ 07703
Gernot M.R. Winkler, U.S. Naval Observatory, Washington, DC 20390
Nicholas F. Yannoni, Rome Air Development Center, Hanscom AFB, Bedford, MA 01731

## 1. Introduction

Techniques to characterize and to measure the frequency and phase instabilities in frequency and time devices and in received radio signals are of fundamental importance to all manufacturers and users of frequency and time technology.

In 1964, a subcommittee on frequency stability was formed within the Institute of Electrical and Electronics Engineers (IEEE) Standards Committee 14 and, later (in 1966), in the Technical Committee on Frequency and Time within the Society of Instrumentation and Measurement (SIM), to prepare an IEEE standard on frequency stability. In 1969, this subcommittee completed a document proposing definitions for measures on frequency and phase stabilities (Barnes, et al., 1971). These recommended measures of instabilities in frequency generators have gained general acceptance among frequency and time users throughout the world.

In this paper, measures in the time and in the frequency domains are reviewed. The particular choice as to which domain is used depends on the application. However, the users are reminded that conversions using mathematical formulations (see Appendix I) from one domain to the other can present problems.

Most of the major manufacturers now specify instability characteristics of their standards in terms of these recommended measures. This paper thus defines and formalizes the general practice of more than a decade.

## 2. Measures of Frequency and Phase Instability

Frequency and phase instabilities shall be measured in terms of the instantaneous, normalized frequency departure $y(t)$ from the nominal frequency $\nu_0$ and/or by phase departure $\phi(t)$, in radians, from the nominal phase $2\pi\nu_0 t$ as follows:

$$y(t) = \frac{1}{2\pi\nu_0}\,\frac{d\phi(t)}{dt} = \frac{\dot{\phi}(t)}{2\pi\nu_0}$$

$$x(t) = \frac{\phi(t)}{2\pi\nu_0}$$

where $x(t)$ is the phase departure expressed in units of time.

## 3. Characterization of Frequency and Phase Instabilities

### a. Frequency Domain:

In the frequency domain, frequency and phase instability is defined by any of the following one-sided spectral densities (the Fourier frequency ranges from 0 to $\infty$):

$S_y(f)$ of $y(t)$

$S_\phi(f)$ of $\phi(t)$

$S_{\dot{\phi}}(f)$ of $\dot{\phi}(t)$

$S_x(f)$ of $x(t)$.

These spectral densities are related by the equations:

$$S_y(f) = \frac{f^2}{\nu_0^2}\,S_\phi(f)$$

$$S_{\dot{\phi}}(f) = (2\pi f)^2\,S_\phi(f)$$

$$S_x(f) = \frac{1}{(2\pi\nu_0)^2}\,S_\phi(f)\ .$$

A device or signal shall be characterized by a plot of spectral density vs. Fourier frequency or by tabulating discrete values or by equivalent means such as a statement of power law(s) (Appendix I).

According to the conventional definition (Kartaschoff, 1978) of $\mathcal{L}(f)$ (pronounced "script ell"), $\mathcal{L}(f)$ is the ratio of the power in one sideband due to phase modulation by noise (for a 1 Hz bandwidth) to the total signal power (carrier plus sidebands), that is,

$$\mathcal{L}(f) = \frac{\text{Power density, one phase-noise modulation sideband/Hz}}{\text{Total signal power}}$$

The conventional definition of $\mathcal{L}(f)$ is related to $S_\phi(f)$ by

$$\mathcal{L}(f) \simeq \tfrac{1}{2}S_\phi(f)$$

only if the mean squared phase deviation, $\langle\phi^2(f)\rangle$ = the integral of $S_\phi(f)$ from f to $\infty$, is much smaller than one radian. In other words, this relationship is valid only for Fourier frequencies f far enough from the carrier frequency and is always violated near the carrier.

Since $S_\phi(f)$ is the quantity that is generally measured in frequency standards metrology, and $\mathcal{L}(f)$ has become the prevailing measure of phase noise among manufacturers and users of frequency standards, $\mathcal{L}(f)$ is redefined as

$$\mathcal{L}(f) \equiv \tfrac{1}{2}S_\phi(f)\ .$$

This redefinition is intended to avoid erroneous use of $\mathcal{L}(f)$ in situations where the small angle approximation is not valid. In other words, $S_\phi(f)$

CH2588-2/88/0000-419 $1 00 C 1988 IEEE

is the preferred measure, since, unambiguously, it always can be measured.

b. Time-Domain:

In the time domain, frequency instability shall be defined by the two-sample deviation $\sigma_y(\tau)$ which is the square root of the two-sample variance $\sigma_y^2(\tau)$. This variance, $\sigma_y^2(\tau)$, has no dead-time between the frequency samples and is also called the Allan variance. For the sampling time $\tau$, we write:

$$\sigma_y^2(\tau) = \frac{1}{2}\left\langle (\bar{y}_{k+1} - \bar{y}_k)^2 \right\rangle$$

where

$$\bar{y}_k = \frac{1}{\tau}\int_{t_k}^{t_k+\tau} y(t)dt = \frac{x_{k+1} - x_k}{\tau}.$$

The symbol < > denotes an infinite time average. In practice, the requirement of infinite time average is never fulfilled; the use of the foregoing terms shall be permitted for finite time averages. The $x_k$ and $x_{k+1}$ are time residual measurements made at $t_k$ and $t_{k+1} = t_k+\tau$, k=1,2,3,..., and $1/\tau$ is the nominal fixed sampling rate which gives zero dead time between frequency measurements. "Residual" implies the known systematic effects have been removed.

If dead time exists between the frequency departure measurements and this is ignored in the computation of $\sigma_y(\tau)$, resulting instability values will be biased (except for white frequency noise). Some of the biases have been studied and some correction tables published [Barnes, 1969; Lesage, 1983; Barnes and Allan, 1988]. Therefore, the term $\sigma_y(\tau)$ shall not be used to describe such biased measurements. Rather, if biased instability measures are made, the information in the references should be used to report an unbiased estimate.

If the initial sampling rate is specified as $1/\tau_0$, then it has been shown that, in general, we may obtain a more efficient estimate of $\sigma_y(\tau)$ using what is called "overlapping estimates." This estimate is obtained by computing

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

where N is the number of original time residual measurements spaced by $\tau_0$ (N=M+1, where M is the number of original frequency measurements of sample time $\tau_0$) and $\tau = m\tau_0$.

From the above equation, we see that $\sigma_y^2(\tau)$ acts like a second-difference operator on the time deviation residuals--providing a stationary measure of the stochastic behavior even for nonstationary processes. Additional variances, which may be used to describe frequency instabilities, are defined in Appendix II.

c. Clock-Time Prediction

The variation of the time difference between a real clock and an ideal uniform time scale, also known as time interval error, TIE, observed over a time interval starting at time $t_0$ and ending at $t_0+t$ shall be defined as:

$$TIE(t) = x(t_0+t) - x(t_0) = \int_{t_0}^{t_0+t} y(t')dt'.$$

For fairly simple models, regression analysis can provide efficient estimates of the TIE (Draper and Smith, 1966; CCIR, 1986). In general, there are many estimators possible for any statistical quantity. Ideally, we would like an efficient and unbiased estimator. Using the time domain measure $\sigma_y^2(\tau)$ defined in (b), the following estimate of the standard deviation (RMS) of TIE and its associated systematic departure due to a linear frequency drift (or its uncertainty) can be used to predict a probable time interval error of a clock synchronized at $t=t_0=0$ and left free running thereafter:

$$RMSTIE_{est} = t\left[\frac{a^2}{4}t^2 + \sigma_{y_0}^2 + \sigma_y^2(\tau = t) + \left(\frac{x(t_o)}{t}\right)^2\right]^{1/2},$$

where "a" is the normalized linear frequency drift per unit of time (aging) or the uncertainty in the drift estimate, $\sigma_{y_0}$ the two-sample deviation of the initial frequency adjustment, $\sigma_y(\tau)$ the two-sample deviation describing the random frequency instability of the clock at $t=\tau$, and $x(t_0)$ is the initial synchronization uncertainty. The third term in the brackets provides an optimum and unbiased estimate (under the condition of an optimum (RMS) prediction method) in the cases of white noise FM and/or random walk FM. The third term is too optimistic, by about a factor of 1.4, for flicker noise FM, and too pessimistic, by about a factor of 3, for white noise PM.

This estimate is a useful and fairly simple approximation. In general, a more complete error analysis becomes difficult; if carried out, such an analysis needs to include the methods of time prediction, the uncertainties of the clock parameters, using the confidence limits of measurements defined below, the detailed clock noise models, systematic effects, etc.

4. Confidence Limits of Measurements

An estimate for $\sigma_y(\tau)$ can be made from a finite data set with M measurements of $y_j$ as follows:

$$\sigma_y(\tau) \simeq \left| \frac{1}{2(M-1)} \sum_{j=1}^{M-1} (\bar{y}_{j+1} - \bar{y}_j)^2 \right|^{1/2}$$

or, if the data are time readings $x_j$:

$$\sigma_y(\tau) \simeq \left| \frac{1}{2\tau^2(M-1)} \sum_{j=1}^{M-1} (x_{j+2} - 2x_{j+1} + x_j)^2 \right|^{1/2}.$$

The 68 percent confidence interval (or error bar), $I_\alpha$, for Gaussian noise of a particular value $\sigma_y(\tau)$ obtained from a finite number of samples can be estimated as follows:

$$I_\alpha \simeq \sigma_y(\tau)\kappa_\alpha M^{-1/2}$$

where:

M = total number of data points used in the estimate,
$\alpha$ = an integer as defined in Appendix I,
$\kappa_2 = \kappa_1 = 0.99$,

$\kappa_0 = 0.87$,
$\kappa_{-1} = 0.77$,
$\kappa_{-2} = 0.75$.

As an example of the Gaussian model with M=100, $\alpha = -1$ (flicker frequency noise) and $\sigma_y(\tau = 1 \text{ second}) = 10^{-12}$, we may write:

$$I_\alpha \simeq \sigma_y(\tau) \bullet (0.77) \bullet (100)^{-\frac{1}{2}} = \sigma_y(\tau) \bullet (0.077)$$

which gives:

$$\sigma_y(\tau = 1 \text{ second}) = (1 \pm 0.08) \times 10^{-12}.$$

If M is small, then the plus and minus confidence intervals become asymmetric and the $\kappa_\alpha$ coefficients are not valid; however, these confidence intervals can be calculated (Lesage and Audoin, 1973).

If "overlapping" estimates are used, as outlined above, then the confidence interval of the estimate can be shown to be less than or equal to $I_\alpha$ as given above (Howe, Allan, Barnes, 1981).

## 5. Recommendations for Characterizing or Reporting Measurements of Frequency and Phase Instabilities

a. Nonrandom phenomena should be recognized, for example:

- o any observed time dependency of the statistical measures should be stated;
- o the method of modeling systematic behavior should be specified (for example, an estimate of the linear frequency drift was obtained from the coefficients of a linear least-squares regression to M frequency measurements, each with a specified averaging or sample time $\tau$ and measurement bandwidth $f_h$);
- o the environmental sensitivities should be stated (for example, the dependence of frequency and/or phase on temperature, magnetic field, barometric pressure, vibration, etc.);

b. Relevant measurement or specification parameters should be given:

- o the method of measurements;
- o the characteristics of the reference signal;
- o the nominal signal frequency $\nu_0$;
- o the measurement system bandwidth $f_h$ and the corresponding low pass filter response;
- o the total measurement time and the number of measurements M;
- o the calculation techniques (for example, details of the window function when estimating power spectral densities from time domain data, or the assumptions about effects of dead-time when estimating the two-sample deviation $\sigma_y(\tau)$);
- o the confidence of the estimate (or error bar) and its statistical probability (e.g. "three-sigma");
- o the environment during measurement;
- o if a passive element, such as a crystal filter, is being measured in contrast to a frequency and/or time generator.

## 6. References


Barnes, J.A., Tables of bias functions, $B_1$ and $B_2$, for variances based on finite samples of processes with power law spectral densities, NBS, Washington, DC, Tech. Note 375, (Jan. 1969).

Barnes, J.A. and Allan, D.W., "Variances Based on Data with Dead Time Between the Measurements: Theory and Tables," NBS Tech Note 1318 (1988).

Barnes, J.A., Chi, A.R., Cutler, L.S., Healey, D.J., Leeson, D.B., McGunigal, T.E., Mullen, J.A., Smith,W.L., Sydnor, R., Vessot, R.F. and Winkler, G.M.R., Characterization of frequency stability, IEEE Trans. Instr. and Meas., Vol. IM-20, 105-120, (May 1971).

CCIR Report 898, Dubrovnik, (1986).

Draper, N.R. and Smith, H., Applied Regression Analysis, John Wiley and Sons,(1966).

Howe, D. A., Allan, D. W.. and Barnes, J.A., Properties of signal sources and measurement methods. Proc. 35th Annual Symposium on Frequency Control, 669-717, (May 1981).

Kartaschoff, P., Frequency and Time, Academic Press, New York (1978).

Lesage, P., and Audoin, C. Characterization of frequency stability: uncertainty due to the finite number of measurements. IEEE Trans. Instr. and Meas., Vol. IM-22, 157-161, (June 1973).

Lesage, P., Characterization of Frequency Stability: Bias due to the juxtaposition of time interval measurements. IEEE Trans. on Instr. and Meas, IM-32, 204-207, (1983).

Stein, S. R, Frequency and Time: Their measurement and characterization. Precision Frequency Control, Vol. 2, Academic Press, New York, 191-232, (1985).


## APPENDIX I

### 1. Power-Law Spectral Densities

Power-law spectral densities are often employed as reasonable and accurate models of the random fluctuations in precision oscillators. In practice, these random fluctuations can often be represented by the sum of five independent noise processes, and hence:

$$S_y(f) = \begin{cases} \sum_{\alpha=-2}^{+2} h_\alpha f^\alpha & \text{for } 0 < f < f_h \\ 0 & \text{for } f \geq f_h \end{cases}$$

where $h_\alpha$'s are constants, $\alpha$'s are integers, and $f_h$ is the high frequency cut-off of a low pass filter. High frequency divergence is eliminated by the restrictions on f in this equation. The identification and characterization of the five noise processes are given in Table 1, and shown in Fig. 1.

## 2. Conversion Between Frequency and Time Domain

The operation of the counter, averaging the frequency for a time $\tau$, may be thought of as a filtering operation. The transfer function, H(f), of this equivalent filter is then the Fourier transform of the impulse response of the filter. The time domain frequency instability is then given by

$$\sigma^2 (M,T,\tau) = \int_0^\infty S_y(f) \mid H(f) \mid^2 df,$$

where $S_y(f)$ is the spectral density of frequency fluctuations. 1/T is the measurement rate (T-$\tau$ is the dead time between measurements). In the case of the two-sample variance $|H(f)|^2$ is $2(\sin^4 \pi\tau f)/(\pi\tau f)^2$. The two-sample variance can thus be computed from

$$\sigma_y^2(\tau) = 2 \int_0^{f_h} S_y(f) \frac{\sin^4 \pi\tau f}{(\pi\tau f)^2} df .$$

Specifically, for the power law model given, the time domain measure also follows a power law.

**TABLE 1 - The functional characteristics of the independent noise processes used in modeling frequency instability of oscillators**

| Description of Noise Process | Slope characteristics of log log plot | | | | |
|---|---|---|---|---|---|
| | Frequency domain | | Time-domain | | |
| | $S_y(f)$ or $S_\phi(f)$ | $S_\phi(f)$ or $S_x(f)$ | $\sigma_y^2(\tau)$ | $\sigma_y(\tau)$ | Mod.$\sigma_y(\tau)$ |
| | $\alpha$ | $\beta$ | $\mu$ | $\mu/2$ | $\mu'$ |
| Random Walk Freq Modulation | -2 | -4 | 1 | 1/2 | 1 |
| Flicker Frequency Modulation | -1 | -3 | 0 | 0 | 0 |
| White Frequency Modulation | 0 | -2 | -1 | -1/2 | -1 |
| Flicker Phase Modulation | 1 | -1 | -2 | -1 | -2 |
| White Phase Modulation | 2 | 0 | -2 | -1 | -3 |

$$S_y(f) = \frac{(2\pi f)^2}{(2\pi\nu_0)^2} S_\phi(f) = h_\alpha f^\alpha \qquad \sigma_y^2(\tau) \sim |\tau|^\mu$$

$$S_\phi(f) = \nu_0^2 h_\alpha f^{\alpha-2} = \nu_0^2 h_\alpha f^\beta \quad (\beta \equiv \alpha-2) \qquad \sigma_y(\tau) \sim |\tau|^{\mu/2}$$

$$S_x(f) = \frac{1}{4\pi^2} h_\alpha f^{\alpha-2} = \frac{1}{4\pi^2} h_\alpha f^\beta \qquad \text{Mod.}\sigma_y(\tau) \sim |\tau|^{\mu'}$$

**TABLE 2 - Translation of frequency instability measures from spectral densities in frequency domain to variances in time domain and vice versa (For $2\pi f_h \tau \gg 1$)**

| Description of noise process | $\sigma_y^2(\tau)$ = | $S_y(f)$ = | $S_\phi(f)$ = |
|---|---|---|---|
| Random Walk Frequency Modulation | $A\|f^2 \; S_y(f)\|_{\tau^1}$ | $\frac{1}{A}\left(\tau^{-1}\sigma_y^2(\tau)\right) f^{-2}$ | $\frac{\nu_0^2}{A}\left(\tau^{-1}\sigma_y^2(\tau)\right) f^{-4}$ |
| Flicker Frequency Modulation | $B\|f \; S_y(f)\|_{\tau^0}$ | $\frac{1}{B}\left(\tau^0 \; \sigma_y^2(\tau)\right) f^{-1}$ | $\frac{\nu_0^2}{B}\left(\tau^0 \; \sigma_y^2(\tau)\right) f^{-3}$ |
| White Frequency Modulation | $C\|f^0 \; S_y(f)\|_{\tau^{-1}}$ | $\frac{1}{C}\left(\tau^1 \; \sigma_y^2(\tau)\right) f^0$ | $\frac{\nu_0^2}{C}\left(\tau^1 \; \sigma_y^2(\tau)\right) f^{-2}$ |
| Flicker Phase Modulation | $D\|f^{-1} S_y(f)\|_{\tau^{-2}}$ | $\frac{1}{D}\left(\tau^2 \; \sigma_y^2(\tau)\right) f^1$ | $\frac{\nu_0^2}{D}\left(\tau^2 \; \sigma_y^2(\tau)\right) f^{-1}$ |
| White Phase Modulation | $E\|f^{-2} S_y(f)\|_{\tau^{-2}}$ | $\frac{1}{E}\left(\tau^2 \; \sigma_y^2(\tau)\right) f^2$ | $\frac{\nu_0^2}{E}\left(\tau^2 \; \sigma_y^2(\tau)\right) f^0$ |

$$A = \frac{4\pi^2}{6} \qquad C = 1/2 \qquad E = \frac{3f_h}{4\pi^2}$$

$$B = 2\log_e 2 \qquad D = \frac{1.038 + 3 \log_e(2\pi f_h \tau)}{4\pi^2}$$

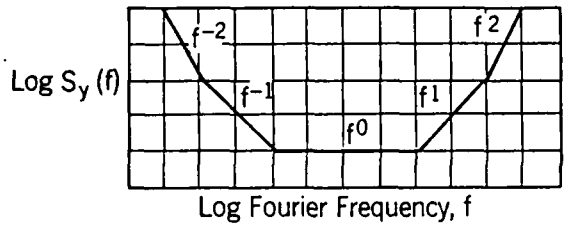
$$\sigma_y^2(\tau) = h_{-2}\frac{(2\pi)^2}{6}\tau + h_{-1}\,2\log_e 2 + h_0\frac{1}{2\tau} + h_1\frac{1.038 + 3\log_e(2\pi f_h\tau)}{(2\pi)^2\tau^2} + h_2\frac{3f_h}{(2\pi)^2\tau^2}\,.$$

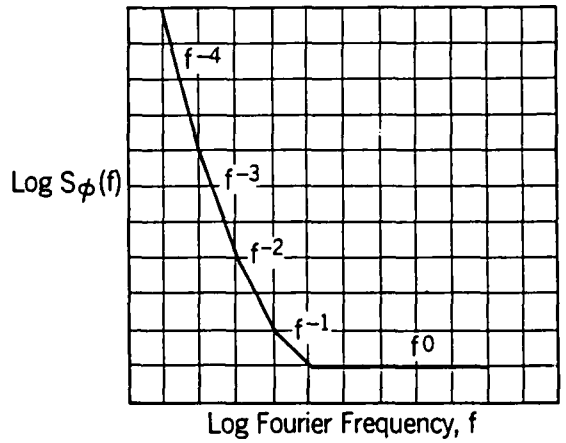
This implicitly assumes that the random driving mechanism for each term is independent of the others. In addition, there is the implicit assumption that the mechanism is valid over all Fourier frequencies, which may not always be true.

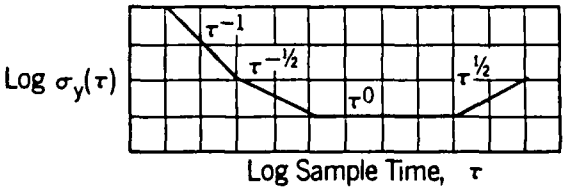
The values of $h_\alpha$ are characteristic models of oscillator frequency noise. For integer values (as often seems to be the case for reasonable models), $\mu = -\alpha - 1$, for $-3 \le \alpha \le 1$, and $\mu \approx -2$ for $\alpha \ge 1$, where $\sigma_y^2(\tau) \sim \tau^\mu$.

Table 2 gives the coefficients of the translation among the frequency stability measures from time domain to frequency domain and from frequency domain to time domain.

The slope characteristics of the five independent noise processes are plotted in the frequency and time domains in Fig. 1 (log-log scale).




Slope characteristics of the five independent noise processes.

FIGURE 1

## APPENDIX II

## ADDITIONAL VARIANCES THAT MAY BE USED TO DESCRIBE FREQUENCY INSTABILITIES IN THE TIME DOMAIN

### 1. Modified Allan, or Modified Two-Sample, Variance, Mod $\sigma_y^2(\tau)$

Instead of the use of $\sigma_y^2(\tau)$, a "Modified Variance" Mod $\sigma_y^2(\tau)$ may be used to characterize frequency instabilities (Stein, 1985; Allan, 1987). It has the property of yielding different dependence on $\tau$ for white phase noise and flicker phase noise. The dependence for Mod $\sigma_y(\tau)$ is $\tau^{-3/2}$ and $\tau^{-1}$ respectively. Mod $\sigma_y^2(\tau)$ is defined as:

$$\text{Mod }\sigma_y^2(\tau) = \frac{1}{2\tau^2 m^2 (N-3m+1)}\sum_{j=1}^{N-3m+1}\left[\sum_{i=j}^{m+j-1}(x_{i+2m} - 2x_{i+m} + x_i)\right]^2$$

where N is the original number of time measurements spaced by $\tau_o$ and $\tau = m\tau_o$ the sample time of choice (N=M+1). A device or signal shall be characterized by a plot of $\sigma_y(\tau)$ or $\sigma_y^2(\tau)$ or Mod $\sigma_y(\tau)$ or Mod $\sigma_y^2(\tau)$ vs. sampling time $\tau$, or by tabulating discrete values or by equivalent means such as a statement of power laws (Appendix I).

### 2. Other Variances

Several other variances have been introduced by workers in this field. In particular, before the introduction of the two-sample variance, it was standard practice to use the sample variance, $s^2$, defined as

$$s^2 = \int_o^{f_h} S_y(f)\left(\frac{\sin \pi f\tau}{\pi f\tau}\right)^2 df.$$

In practice it may be obtained from a set of measurements of the frequency of the oscillator as

$$s^2 = \frac{1}{N}\sum_{i=1}^{N}(y_i - \bar{y})^2 \quad .$$

The sample variance diverges for some types of noise and, therefore, is not generally useful.

Other variances based on the structure function approach can also be defined (Lindsey and Chi, 1976). For example, there are the Hadamard variance, the three-sample variance and the high pass variance (Rutman 1978). They are occasionally used in research and scientific works for specific purposes, such as differentiating between different types of noise and for dealing with systematics and sidebands in the spectrum.

## APPENDIX III

## Bibliography


Allan, D. W., Statistics of atomic frequency standards, Proc. IEEE, Vol. 54, 221-230, (Feb 1966).

Allan, D. W., et al., Performance, modeling, and simulation of some cesium beam clocks. Proc. 27th Annual Symposium on Frequency Control, 334-346, (1973).

Allan, D. W., The measurement of frequency and frequency stability of precision oscillators. Proc. 6th Annual Precise Time and Time Interval Planning Meeting, 109-142, (Dec 1974).

Allan, D. W., and Daams, H., Picosecond time difference measurement system. Proc. of the 29th Annual Symposium on Frequency Control, 404-411, (May 1975).

Allan, D. W., and Hellwig, H. W., Time Deviation and Time Prediction Error for Clock Specification, Characterization, and Application, Proc. IEEE Position Location and Navigation Symposium, 29-36, (1978).

Allan, D. W., and Barnes, J. A., A modified "Allan Variance" with increased oscillator characterization ability. Proc. 35th Annual Symposium on Frequency Control, 470-476, (May 1981).

Allan, D. W., Time and Frequency (Time Domain) Characterization, Estimation and Prediction of Precision Clocks and Oscillators. IEEE Trans. UFFC Vol. 34, 647-654 (Nov 1987).

Atkinson, W. K., Fey, L., and Newman, J., Spectrum analysis of extremely low-frequency variations of quartz oscillators. Proc. IEEE, Vol. 51, 379-380, (Feb 1963).

Babitch, D., and Oliverio, J., Phase noise of various oscillators at very low Fourier frequencies. Proc. 28th Annual Symposium on Frequency Control, 150-159, (1974).

Barnes, J. A., Models for the interpretation of frequency stability measurements. NBS Tech. Note 683, (Aug 1976).

Barnes, J. A., Jones, R. H., Tryon, P. V., and Allan, D.W., Noise models for atomic clocks. Proc. 14th Annual Precise Time and Time Interval Planning Meeting, 295-307, (Dec 1982).

Baugh, R. A., Frequency modulation analysis with the Hadamard variance. Proc. 25th Annual Symposium on Frequency control, 222-225, (1971).

Bernier, L.G., Linear Prediction of the Non-Stationary Clock Error Function, Proc. 2nd European Frequency and Time Forum, Neuchâtel (1988).

Blackman, R. B., and Tukey, J. M., The measurement of power spectra, Dover Publication, Inc., New York, NY, (1959).

Blair, B. E., Time and frequency: Theory and fundamentals. NBS Monograph No. 140, US Government Printing Office, Washington, DC 20402, (May 1974).

Boileau, E., and Picinbono, B., Statistical study of phase fluctuations and oscillator stability. IEEE Trans. Instr. and Meas., Vol. 25, 66-75, (Mar 1976).

Brandenberger, H., Hadorn, F., Halford, D., and Shoaf, J. H., High quality quartz crystal oscillators: frequency-domain and time-domain stability. Proc. 25th Annual Symposium on Frequency Control, 226-230, (1971).

Chi, A. R., The mechanics of translation of frequency stability measures between frequency and time-domain measurements. Proc. 9th Annual Precise Time and Time Interval Planning Meeting, (Dec 1977).

Cutler, L. S., and Searle, C. L., Some aspects of the theory and measurement of frequency fluctuations in frequency standards, Proc. IEEE, Vol. 54, 136-154, (Feb 1966).

De Prins. J., and Cornelissen, G., Analyse spectrale discrete. Eurocon (Lausanne, Switzerland), (Oct 1971).

De Prins, J., Descornet, G., Gorski, M., and Tamine, J., Frequency-domain interpretation of oscillator phase stability. IEEE Trans. Instr. and Meas., Vol. IM-18, 251-261, (Dec 1969).

Fischer, M. C., Frequency stability measurement procedures. Proc. 8th Annual Precise Time and Time Interval Planning Meeting, 575-618, (Dec 1976).

Greenhall, C.A., Initializing a Flicker Noise Generator, IEEE Trans. Instr. and Meas., Vol. IM-35, 222-224 (June 1986).

Groslambert, J., Oliver, M., and Uebersfeld, J., Spectral and short-term stability measurements. IEEE Trans. Instr. and Meas., Vol. IM-23, 518-521, (Dec 1974).

Halford, D., A general mechanical model for |f| spectral density noise with special reference to flicker noise 1/|f|. Proc. IEEE Vol. 56, 251-258, (Mar 1968).

IEEE, Special issue on time and frequency. Proc. IEEE, Vol. 60, (May 1972).

IEEE-NASA, Proceedings of Symposium on short-term frequency stability. NASA Publication SP 80, (1964).

Jones, R. H., and Tryon, P. V., Estimating time from atomic clocks. NBS Journal of Research, Vol. 88, November 17-24, (Jan-Feb 1983).

Kartaschoff, P., Computer Simulation of the Conventional Clock Model, IEEE Trans. Instr. and Meas., Vol. IM-28, 193-197 (Sept 1979).

Kroupa, V. F., Frequency stability: Fundamentals and Measurement. IEEE Press. IEEE Selected Reprint Series. Prepared under the sponsorship of the IEEE Instrumentation and Measurement Society, 77-80, (1984).

Lesage P., and Audoin, C., Correction to: Characterization of frequency stability: Uncertainty due to finite number of measurements. IEEE Trans. Instr. and Meas., Vol. IM-23, 103, (Mar 1974).

Lesage, P., and Audoin, F., A time domain method for measurement of the spectral density of frequency fluctuations at low Fourier frequencies. Proc. 29th Annual Symposium on Frequency Control, 394-403, (May 1975).

Lesage, P., and Audoin, C., Correction to: Characterization of frequency stability: Uncertainty due to the finite number of measurements. IEEE Trans. Instr. and Meas., Vol. IM-25, 270, (Sept 1976).

Lesage, P., and Ayi, T., Characterization of frequency stability: Analysis of modified Allan variance and properties of its estimate. IEEE Trans. on Instr. and Meas., Vol. IM-33, 332-336, (Dec 1984).

Lindsey, W. C., and Chie, C. M., Theory of oscillator instability based upon structure functions. Proc. IEEE, Vol. 64, 1662-1666, (Dec 1976).

Mandelbrot, B., Some noises with 1/f spectrum; a bridge between direct current and white noise, IEEE Trans. Inf. Theory, Vol. IT-13, 289-298, (Apr 1967).

Meyer, D. G., A test set for the accurate measurements of phase noise on high-quality signal sources. IEEE Trans. Instr. and Meas., Vol. IM-19, 215-227, (Nov 1970).

Percival, D. B., A heuristic model of long-term atomic clock behavior. Proc. 30th Annual Symposium on Frequency Control, (June 1976).

Ricci, D. W., and Peregrino, L., Phase noise measurement using a high resolution counter with on-line data processing. Proc. 30th Annual Symposium on Frequency Control, (June 1976).

Rutman, J., Comment on characterization of frequency stability. IEEE Trans. Instr. and Meas., Vol. IM-21, 85, (Feb 1972).

Rutman, J., Characterization of frequency stability: A transfer function approach and its application to measurements via filtering of phase noise. IEEE Trans. Instr. and Meas., Vol. IM-23, 40-48, (Mar 1974).

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

Rutman, J., and Sauvage, G., Measurement of frequency stability in the time and frequency domains via filtering of phase noise. IEEE Trans. Instr. and Meas., Vol. IM-23, 515-518, (Dec 1974).

Rutman, J., and Uebersfeld, J., A model for flicker frequency noise of oscillators. Proc. IEEE, Vol. 60, 233-235, (Feb 1972).

Sauvage, G., and Rutman, J., Analyse spectrale du bruit de frequence des oscillateurs par la variance de Hadamard. Ann. des Telecom., Vol. 28, 304-314, (July-Aug 1973).

Vanier, J., and Tetu, M., Time domain measurement of frequency stability: A tutorial approach, Proc. 10th Annual Precise Time and Time Interval Planning Meeting, 247-291, (1978).

Vessot, R., Frequency and Time Standards. Methods of Experimental Physics, Academic Press, New York, 198-227, (1976).

Vessot, R., Mueller, L., and Vanier, J., The specification of oscillator characteristics from measurements made in the frequency domain. Proc. IEEE, Vol. 54, 199-207 (Feb 1966).

Von Neumann, J., Kent, R. H., Bellinson, H. R., and Hart, B. I., The mean square successive difference, Ann. Math. Stat., Vol. 12, 153-162, (1941).

Walls, F.L., and Allan, D.W., Measurements of Frequency Stability, Proc. IEEE, Special Issue on Radio Measurements and Standards, Vol. 74, No. 1, pp. 162-168 (Jan 1986).

Walls, F. L., Stein, S. R., Gray, J. E., and Glaze, D. J., Design considerations in state-of-the-art signal processing and phase noise measurement systems. Proc. 30th Annual Symposium on Frequency Control, (June 1976).

Winkler, G. M. R., A brief review of frequency stability measures. Proc. 8th Annual Precise Time and Time Interval Planning Meeting, 489-528, U.S. Naval Research Laboratory, Washington, D.C., (Dec 1976).

Winkler, G. M. R., Hall, R. G., and Percival, D. B., The U.S. Naval Observatory clock time reference and the performance of a sample of atomic clocks. Metrologia, Vol. 6, 126-134, (Oct 1970).

Yoshimura, K., Characterization of frequency stability: Uncertainty due to the autocorrelation of the frequency fluctuations. IEEE Trans. Instr. and Meas., IM-27, 1-7, (Mar 1978).

---

The authors are members of the Technical Committee TC-3, Time and Frequency, of the IEEE Instrumentation and Measurement Society. This paper is part of this Committee's effort to develop an IEEE Standard. To this end, the IEEE authorized a Standards Coordinating Committee, SCC 21, and the development of the IEEE Standard under PAR-P-1139. SCC 21 is formed out of the I&M, UFFC and MTT Societies; the authors are also members of SCC 21.

42nd Annual Frequency Control Symposium - 1988

# NOISE AND TIME AND FREQUENCY --- A POTPOURRI

James A. Barnes
Austron, Inc.

## ABSTRACT

There exist several sources of noise encountered in time and frequency related equipment. The noises can be characterized by their spectral type and mode of influence (e.g., additive or multiplicative). The noises have various consequences in processing equipment that can seriously affect performance. This tutorial paper reviews the spectral types of noise, the physical origins in various devices (where known), and the consequences of signal processing (i.e., frequency addition, subtraction, multiplication, and division). Noises also affect the stabilities of oscillators and frequency standards.

## Introduction

One can classify noises in several ways. For example, power spectral densities (PSD) are a powerful and common means of description but often not complete. One can have a given noise type modulating the RF carrier, the frequency, or the phase of a signal. Further, the noise could be some filtered form of a random, uncorrelated (i.e., "white") noise or it could be a "popcorn noise" with essentially the same spectrum -- there are an infinite number of sample noise functions with the same PSD. Thus, we can speak of the spectral type of a noise, its origin, and the carrier on which it rides to name a few. Recently, one hears of "Fractals", flicker noise, chaos, and sporadic noises also.

The following is a partial list of noises and some of their properties. I do not pretend that this is an exhaustive list but it does cover many of the noise models in common use today and some of their consequences in electronic systems.

The word "frequency" is used in two separate contexts: The instantaneous frequency of a signal in the time domain; and Fourier frequencies in the frequency domain. These are two very different concepts. The instantaneous frequency, $\nu$, is defined to be the time rate of change of the instantaneous phase of a signal. The Fourier frequency, f, labels a spectral component in the frequency domain.

## Noise Sources and Types

### A. Thermal Noise

Thermal noise is a white noise associated with the thermodynamic properties of a simple resistance. The voltage power spectral density has the value 4kTR, where k is Boltzman's constant, T is the absolute temperature in Kelvin, and R is the resistance in Ohms. Thermal noise has a Gaussian (i.e., normal or "bell shaped") distribution.

CH2588-2/88/0000-426 $1.00 C 1988 IEEE

### B. Shot Noise

Shot noise arises from the discrete nature of an electronic current. For example, a cesium beam atomic frequency standard uses a flux of cesium atoms through the device. The detector current is sporadic in nature as the individual atoms are ionized and collected. The PSD of shot noise is white (i.e., constant) over a frequency range below the reciprocal transit time of the ion within the detector. Shot noise is also present in vacuum tubes and semi-conductors. Any quantized current displays shot noise.

### C. Power-Law Noises

A power-law noise by definition is a noise whose PSD varies as $f^{\alpha}$ over a large, but important range of Fourier frequencies, f. It is obvious that the range of Fourier frequencies must be bounded either on the high end or near zero, or both since the total power must be finite. That is,

$$\int_{0}^{\infty} S(f)\, df < \infty \qquad (1)$$

where S(f) is the PSD as a function of the Fourier frequency, f. The exponent, $\alpha$, is assumed to be a constant.

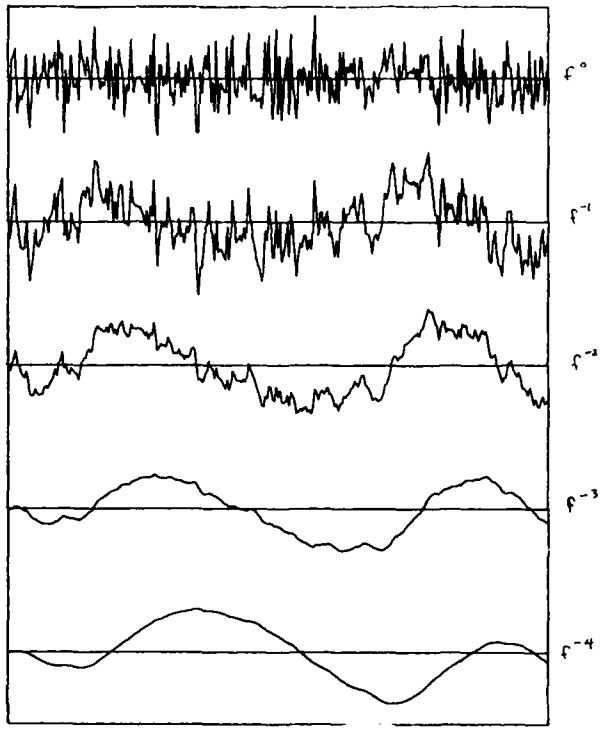
If one has a signal whose frequency is modulated by a white noise, then the phase must be the integral of the white noise. Such an integral is often called a Brownian motion, or a random walk. One can show (e.g., by the Laplace transform), that the PSD of a random walk, then,is proportional to $f^{-2}$ -- a power-law PSD.

The power-law PSD's play an important role in the measurement and characterization of frequency stability of clocks and oscillators for integer exponents [Kroupa, '83]. In particular there is a set of names often given to some of the noises:

| NAME | $S_{\phi}(f)$ | $S_{y}(f)$ |
|---|---|---|
| White noise phase mod (PM) | $f^{0}$ | $f^{2}$ |
| Flicker noise PM | $f^{-1}$ | $f^{1}$ |
| White noise frequency mod (FM) (same as random walk PM) | $f^{-2}$ | $f^{0}$ |
| Flicker noise FM | $f^{-3}$ | $f^{-1}$ |
| Random walk FM | $f^{-4}$ | $f^{-2}$ |

where $S_{\phi}(f)$ is the PSD of the phase fluctuations, $S_{y}(f)$ is the PSD of the frequency fluctuations, and Flicker noise is discussed below (see Fig. 1).

Typical oscillator noise models often assume some sub-set of these noises along with a linear frequency drift. There is no fundamental reason that the exponent in the power-law could not be a fractional value, but few researchers have used them. The




natural tendency of scientists is to use the integral values.

The PSD is a very useful diagnostic tool for identifying noise types in the frequency domain. In the time domain, one often uses the two-sample (or Allan) variance [Kroupa, '83).

D. Flicker noise

Flicker noise is a power-law noise whose PSD varies as 1/ f. The subject of flicker noise has received much attention with flicker noise being found almost everywhere -- from the height of the river Nile to the music of Bach and the Beatles. In particular, flicker noise seems to be present at some level in almost all electronic devices such as vacuum tubes and semi-conductors. [van der Ziel, '88], [Halford '66], [Mandelbrot, '67].

The physical origin(s) of flicker noise are not certain but it is known that diffusion processes can give rise to flicker noise. Again using the Laplace transform, a flicker noise could be generated from a white noise with a filter whose transfer function is proportional $1/\sqrt{s}$. Note that such a filter has a constant argument of 45 degrees. (A "whole" integrator, 1/s, has a constant argument of 90 degrees.) It is not surprising that the conventional seasons (Spring, Summer, Fall, and Winter) lag the annual north-south motion of the sun by 45 degrees. The source of this lag is the heat diffusion into the ground. Shorter term random (weather) variations could similarly give rise to flicker noise variations in addition to the annual term. Computer programs approximate the half-ordered integrals to provide flicker noise simulations of clocks and oscillators [Barnes, '87].

E. Popcorn Noise

Occasionally, a device will show sporadic bursts of "noise". Sometimes this appears as a sudden shift in value between two rather fixed levels. This has the appearance of a random telegraph signal. If one connects this signal to headphones or a speaker, then one hears a "noise" which can sound very much like popcorn popping. Sometimes, the noise involves more than two quasi-stable levels [Card, '65]. Popcorn noise can have a 1/f PSD [Barnes, '76]. (See Fig. 2.)

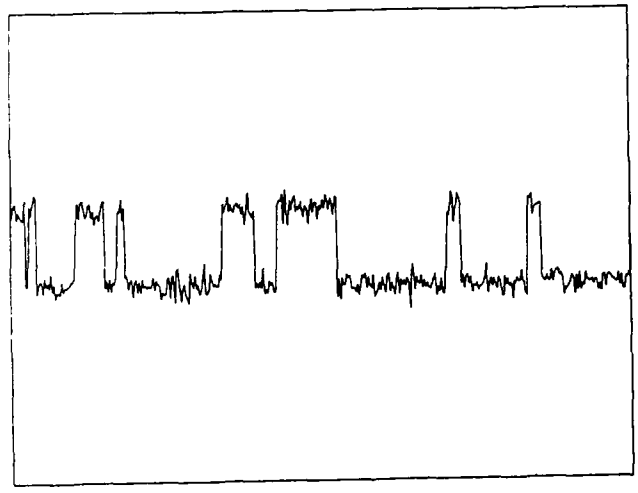


Fig. 2. Simulated Popcorn Noise

The physical mechanism(s) underlying this noise are unknown (as far as this author knows). However, popcorn noise has been found in quartz crystal oscillators and also capacitors. In the case of oscillators, the time between steps can exceed one or two hours. This time interval often grows as the oscillator ages. It has been suggested that replacement of the crystal often cures the problem. There is some indication that popcorn noise is associated with some surface contamination and that improved fabrication techniques have resulted in reduced levels of popcorn noise, generally.

F. Chaos

Cauchy proposed the concept of "The well-posed problem". The key condition of the well-posed problem is that small changes in the input result in small changes in the results. Chaotic behavior is exactly the opposite. One of the most familiar chaotic processes is the pseudo-random noise generator in a computer. Each seed number given to a computer generates a series of random numbers which is uncorrelated with series begun with even slightly different seeds. Turbulence is a classic example of chaotic behavior in nature. (See Fig. 3.) [Chua, '87]

G. Fractals

Many of the noises considered here have the property of "self-similarity". Assume that we have a segment of a sample time series

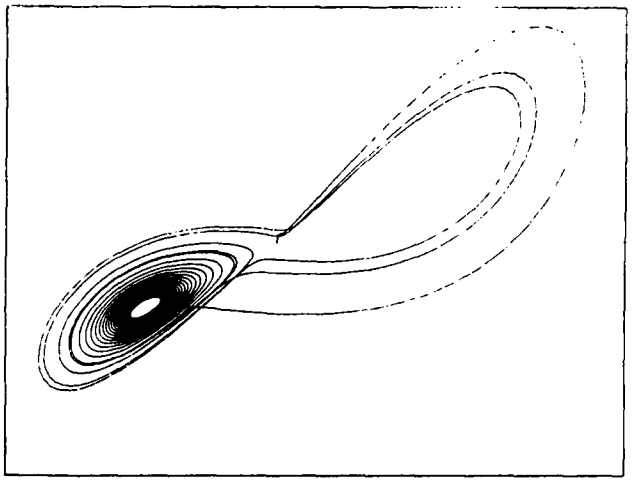


Fig. 3. A "Strange Attractor" See [Chua, '87]

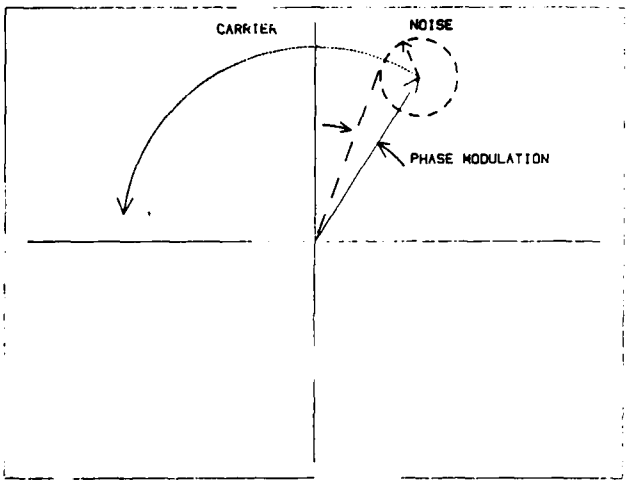



Fig. 4. Phase and Amplitude Modulation with Additive Noise

obtained from a given population (i.e., a given noise). If the noise is self-similar, then I can select an arbitrary scaling factor, H, to scale the time axis of the noise sample. Given this horizontal factor there exists a vertical scaling factor, V, which, when applied to the time series already scaled by H, yields a noise process whose statistics are the same as the given population. For example, a Brownian motion is self-similar and the vertical scaling factor is just the square root of the horizontal scaling factor. Another example comes from an oscillator whose frequency is modulated by a flicker noise. The phase is self-similar and the vertical scaling factor is identical to the horizontal factor.

All of these noises belong to the study of "Fractional Brownian motions" or "Fractals" as developed by Benoit Mandelbrot. [Mandelbrot, '71, '77], [Peitgen, '86]

## Additive and Multiplicative Noises

Consider the situation where we have a radio receiver, a distant transmitter and a lot of local noise. The local noise power added to the received signal is half amplitude modulation and half phase modulation (see Fig. 4). We can improve the received signal-to-noise ratio by increasing the transmitter power.

Assume next that there is no local (additive) noise sources but that the transmission path is not constant and varies with, say, the height of the ionosphere. The variations of path delay appear as phase modulation sidebands on the received signal. Now however, increasing the transmitted power does NOT improve the received signal-to-noise ratio. No matter what the signal level, the delay (and, hence the phase of the signal) will be the same.

The first example is that of additive noise and the second that of multiplicative noise.

For whatever reason, vacuum tubes as well as semi-conductors used as the active element in an amplifier, exhibit a (multiplicative) phase noise whose spectrum is that of flicker noise. It is surprising that over wide choices of semi-conductors or vacuum tubes, the phase noise is remarkably close to the same level when expressed in dB/Hz relative to the carrier. The level is about -115 dB/Hz +/- 10 dB/Hz at one Hz. [Halford, '68] [Healey, '74], [Parker, '87]. Of course, at some suitable modulation frequency, the flicker noise will fall below additive, white, thermal noise. It is worth noting that for small phase modulation indices ( <1 ) the PSD of the RF carrier about the carrier frequency mirrors the phase PSD (see Fig. 5). This is useful in measuring frequency stabilities of oscillators and amplifiers.

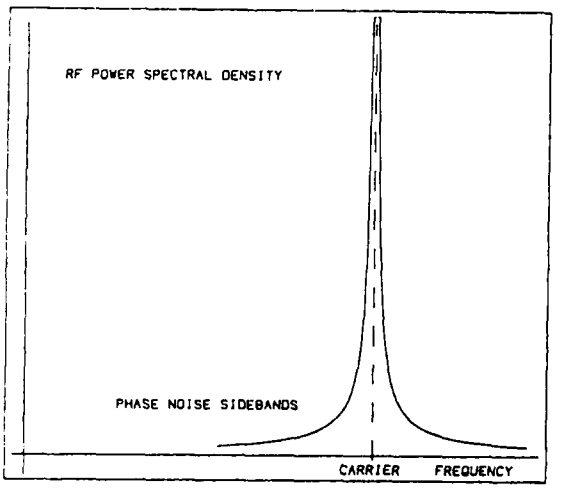



Fig. 5. Power Spectral Density of a Phase Modulated RF Carrier

It has been found that adding local negative feedback to the amplifier can reduce the flicker noise by as much as 30 dB/Hz. I am not aware of any physical explanation for these noises or for their constancy over many different devices. Since this is a multiplicative noise, raising the signal level does not help. Further, using a low-noise ( i.e., low phase noise PM) pre-amplifier doesn't help. The system phase noise will be no better than that of the worst amplifier in the system.

## Noise and Arithmetic Operations on Freq.

Two RF signals can be mixed or "beat" together using non-linear elements (e.g., mixers). A piano tuner listens to beats between notes on a piano. These beat notes are the difference frequencies between two notes (or their harmonics). When two RF signals are mixed in a non-linear device one gets all possible sums and differences of the RF frequencies. That is,

$$V(t) = a\,X(t) + b\,Y(t) + c\,X^2(t) + d\,Y^2(t) + e\,X(t)\,Y(t) + g\,X^2(t)\,Y(t) + \ldots \quad (2)$$

where the two input RF signals are

$$X(t) = X_0 \sin(\omega\, t + \phi),$$

$$Y(t) = Y_0 \sin(\omega\, t + \phi),$$

$$\omega = 2\pi\nu$$

and the a, b, c, ... are coefficients of the non-linearity. If one uses a double balanced mixer, the first four coefficients (a, b, c, and d) are relatively small and the coefficient e, of the cross term, predominates.

Thus, we are left with

$$V(t) = e \cdot X(t) \cdot Y(t)$$

$$= X_0 \sin(\omega_x t + \phi_x)\; Y_0 \sin(\omega_y t + \phi_y)$$

$$= X_0 Y_0 [\cos((\omega_x + \omega_y)t + \phi_x + \phi_y) + \cos((\omega_x - \omega_y)t + \phi_x - \phi_y)] \quad (3)$$

where use has been made of some trigonometric identities. The piano tuner hears the difference frequency, $\omega_x - \omega_y$, and he can set this beat note to the desired value (not always zero in the "equally tempered scale"). Whichever term is desired can often be selected by filtering the signal, V(t). It is important to notice that the phase relationships are retained in the sum and difference components -- it is not just the frequencies which are added and subtracted.

Not only can one add and subtract frequencies (and phases), but one can also multiply and divide. One can multiply by the integer 2 simply by feeding the same signal to both inputs of the double balanced mixer discussed above and band-pass filter the output signal centered on the second harmonic. The frequency doubling process also doubles the phase variations. The doubling process is exact by virtue of the trigonometric identities used in Eq. 3.

If we send a sinusoidal signal through a simple non-linear device (e.g., a diode), the output is a distorted sinewave, and from Fourier analysis we know that such a distorted signal can be represented by a Fourier series containing all integer multiples of the frequency of the original signal. All that remains is to filter out the desired harmonic. Again, the output phase is multiplied by the same integer. That is, if we multiply a noisy signal by the integer N, then the phase excursions are also multiplied by the same integer, N. Frequency multiplication degrades the spectral purity of a noisy signal. In effect, as the phase modulation index grows with frequency multiplication at the expense of carrier power, one can reach a point where the carrier is lost entirely. This becomes especially important when one wants to generate a microwave signal controlled by a lower frequency quartz oscillator [Walls, '75].

Although phase noise and amplitude noise are two distinct concepts, real life devices often suffer from conversions of phase noise to amplitude noise and the reverse.

## Noise and Oscillators

An oscillator is basically an amplifier and a resonance (e.g., a quartz crystal) in a feedback loop. (Oscillators can function without a sharp resonance, but most high-stability oscillators include a high Q resonator.) Considering the transfer functions of the amplifier and the resonator, the oscillation frequency will be that frequency for which the net phase shift around the loop is zero (actually N2$\pi$ ). If there is even a momentary change in the phase shift in the amplifier, the instantaneous frequency (NOT a Fourier frequency) will also change no matter what the source of the phase shift. These phase shifts might come from additive, white, thermal noise, or from flicker noise as noted above for most amplifiers.

Note that a PHASE shift within the oscillator loop causes a FREQUENCY change in the output. Thus if the amplifier phase noise is white, then the oscillator output frequency would also be white noise, or the output phase would be a random walk. This is true for noises within the band pass of the resonator. Outside the resonator bandwidth the phase noise would be essentially undisturbed. It is interesting to note that since most amplifiers impart a flicker noise PM to the signal, an oscillator would have a flicker noise FM spectrum -- or the output phase noise would be the integral of a flicker noise.

Quartz crystal oscillators are often phase locked to a reference in order to improve spectral purity of the reference. At first thought it would seem that the bandpass of the resonator might significantly restrict just how fast one could change the oscillator frequency. This is not true if the changes directly affect the frequency determining parameters (in contrast to the noise within the loop). As an example, voltage controlled crystal oscillators (VCXO's) can be tuned rapidly by changing the bias on a small varicap in the oscillator circuit.

As a thought experiment, consider a pendulum of ten to twenty centimeters in length and a heavy weight at the end to make a resonant device. It is reasonable to imagine that such a pendulum might make in excess of twenty or more swings before it fell to 1/e of its initial amplitude. We could say it has a Q of

greater than 20 or that the resonance bandwidth is 1/20 th of the resonators normal frequency. Now, suppose we start the pendulum swinging. If I should suddenly pinch the pendulum's string mid way down its length, and hold that point of the string fixed, then the pendulum would almost instantaneously assume a new frequency characteristic of its new length. Notice that the pendulum would take on its new frequency in a time short in comparison to a single period of its swing even though it has a "high" Q.

In principle, frequency division should reverse the affects of multiplication on phase noise sidebands. Unfortunately, this is often not the case. A frequency divider is typically a digital counter which puts out a pulse after N cycles of the input have gone by. The phase fluctuations of the input are, indeed, scaled down in the division, but new phase noise sources (e.g., amplifiers and resistors) are often present to contaminate again the lower frequency output signal from the divider.

## Observation

When one adds or subtracts two frequencies, the relative phases (expressed in radians or degrees) are also added or subtracted. When one multiplies or divides the frequency of a signal, the delay time is unaffected. As an example, consider a divider from 1 MHz to 1 pps. If one delays the input by, say, 10 nanoseconds, then the output 1 pps will be delayed by exactly 10 nanoseconds also. In this latter example, the phase (expressed in radians or degrees) IS divided but NOT the delay time (i.e., phase expressed in time units -- nanoseconds).

## Hetroschedasticty

The level of noise contamination of a signal often changes with time (see Figs. 6 & 7). This is especially true of radio propagation where the short-term noise LEVEL (i.e., noise intensity) has a diurnal variation. The time and frequency community has not treated the subject of time-varying variances to any great extent. Many of the models and their associated analysis techniques are not appropriate for variances which change with time of day. Kalman filters are probably the best means of treating the subject since they explicitly allow for time varying model parameters, including state covariances and their up-dates. One must develop a method to estimate the various model parameters, however.

It's interesting to note that an amplitude modulated white noise is still "white" in the sense that the covariance function is a delta function, and hence, the estimated PSD is constant (white). Still, the forecast error variance can be very far off depending on the autocorrelation function of the signal that modulates the white noise. We are back to one of the early comments made in this paper: There is an infinite number of distinct noises with the same PSD, and these noises may need their own analysis and estimation techniques -- the PSD does not tell it all! Another observation is that one person's noise is another person's signal.

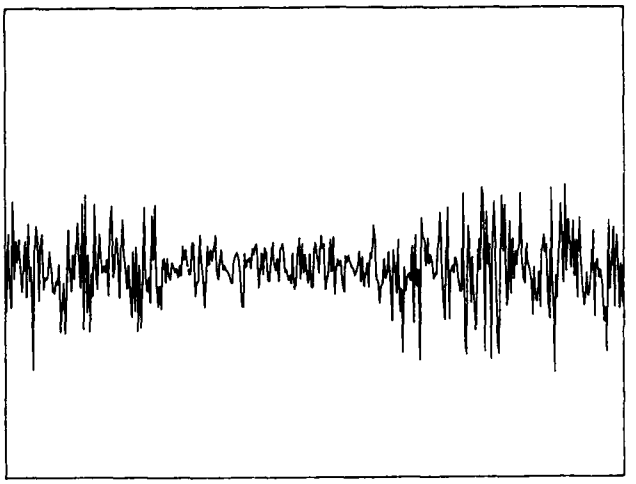


Fig. 6. Simulated White Noise with Time-Varying Variance

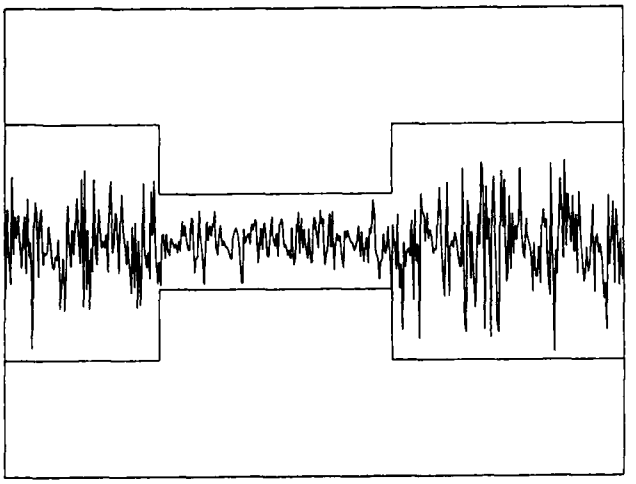


Fig. 7. Simulated Noise of Fig. 6 with ±3σ Limits. (Known std. of 8db)

## Marginal Distributions

Imagine a large ensemble of independent but statistically identical clocks (i.e., they were drawn from the same population). We can imagine that they were all synchronized at the same instant of time. At some later time the clocks will have accumulated errors and there will be a distribution of clock errors. Typically, the distribution of clock errors at an instant (the "marginal distribution") is bell-shaped (or "normal", or "Gaussian"). The width of the marginal distribution usually grows with running time: for example it will grow as the square root of the running time if the clocks are affected by white FM.

Unfortunately, not all marginal distributions are Gaussian. Examples of non-Gaussian distributions include sinusoidal modulations (PM) and popcorn noise. There exist some distributions for which neither the first nor the second moment of the distribution converges. That is, there is no mean frequency or variance. These are often referred to as "fat-tailed" distributions an example of which is a Cauchy distribution. A

Cauchy distribution is the statistician's equivalent of the physicist's Lorentz line shape. For these "fat-tailed" distributions special statistical treatment is necessary such as the use of the median rather than the mean to estimate a central tendency of the distribution [Winkler, '87]. Fortunately, the Gaussian distribution fits most clock performances. One can test for normality by plotting the observed cumulative distribution function on standard probability paper. A truly normal distribution will appear as a straight line on the plot.

## REFERENCES


J.A. Barnes, "Large Sample Simulation of Flicker Noise," Proceedings of the Nineteenth Annual Precise Time and Time Interval (PTTI) Applications Planning Meeting, Dec. 1987.

J.A. Barnes, "Models for the Interpretation of Frequency Stability Measurements", National Bureau of Standards Technical Note 683, US Government Printing Office, Washington, D.C., 1976.

W.H. Card, and P.K. Chaudhari, "Characteristics of Burst Noise", Proc. IEEE, 652-653, June, 1965.

L.O. Chua et al., ed., IEEE Special Issue on Chaotic Systems, Proc IEEE, Aug. 1987.

D. Halford, "A General Mechanical Model for $f^{\alpha}$ Spectral Density Random Noise with Special reference to Flicker Noise 1/f", Proc IEEE, Vol 56, No. 3, 251-258, March 1966.

D. Halford, et al., "Flicker Noise of Phase in RF Amplifiers: Characterization, Cause, and Cure", Proceedings of the Annual Symposium on Frequency Control, No. 22, 1968.

D.J. Healey III, "L(f) Measurements on UHF Sources Comprising VHF Crystal Controlled Oscillator Followed by a Frequency Multiplier", Proceedinngs of the Annual Symposium on Frequency Control No. 28, 1974.

V.F. Kroupa, ed., "Frequency Stability: Fundamentals and Measurement", IEEE Press, NY, 1983. (Contains reprints of many important papers dealing with frequency stability.)

B. Mandelbrot, "A Fast Fractional Gaussian Noise Generator", Water Resources Research, Vol. 7, No.3, June 1971.

B. Mandelbrot, "Fractals; Form, Chance, and Dimension", W.H. Freeman and Co. San Francisco, 1977.

B. Mandelbrot, "Some Noises with 1/f Spectrum, a Bridge Between Direct Current and White Noise," IEEE Trans on Info. Theory Vol IT-13, NO. 2, 289-298, April 1967.

T.E. Parker, "Characteristics and Sources of Phase Noise in Stable Oscillators", Proc 41st Annual Symposium on Frequency Control, 99-110, 1987.

H.-O. Peitgen, and P.H. Richter, "The Beauty of Fractals; Images of Complex Dynamical Systems", Springer-Verlag, Berlin, Heidelberg, New York, Tokyo, 1986.

F.L. Walls and A. DeMarchi, "RF Spectrum of a Signal after Multiplication; Measurement and Comparison with Simple Calculations," IEEE Trans on I & M, Vol. IM-24, 210-217, Sept. 1975.

G.M.R. Winkler, "Introduction to Robust Statistics and Data Filtering", Proc. Nineteenth Annual Precise Time and Time Interval Applications and Planning Meeting, Dec. 1987.

A. van der Ziel, "Unified Presentation of the 1/f Noise in Electronic Devices: Fundamental 1/f Noise Sources", Proc. IEEE, Vol. 76, No. 3, 233-258, March 1988.

42nd Annual Frequency Control Symposium - 1988

# EXTENDING THE RANGE AND ACCURACY OF PHASE NOISE MEASUREMENTS

F.L. Walls, A.J.D. Clements, C.M. Felton, M.A. Lombardi, and M.D. Vanek

Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

## Summary

This paper describes recent progress in extending high accuracy measurements of phase noise in oscillators and other devices for carrier frequencies from the rf to the millimeter region and Fourier frequencies up to 10 % of the carrier (or a maximum of about 1 GHz). A brief survey of traditional precision techniques for measuring phase noise is included as a basis for comparing their relative performance and limitations. The single oscillator techniques, although conceptually simple, require a set of 5 to 10 references to adequately measure the phase noise from 1 Hz to 1 GHz from the carrier. The two oscillator technique yields excellent noise floors if, for oscillator measurements, one has a comparable or better oscillator for the reference and, for other devices, either pairs of devices or a reference oscillator with comparable or better noise. We have developed several new calibration techniques which, when combined with previous two oscillator techniques, permits one to calibrate all factors affecting the measurements of phase noise of oscillator pairs to an accuracy which typically exceeds 1 dB and in favorable cases can approach 0.4 dB. In order to illustrate this expanded two oscillator approach, measurements at 5 MHz and 10 GHz are described in detail. At 5 MHz we achieved accuracies of about ±0.6 dB for phase noise measurements from 20 Hz to 100 kHz from the carrier. At 10 GHz we achieved an accuracy of ±0.6 dB for phase noise measurements a few kHz from the carrier degrading to about ±1.5 dB, 1 GHz from the carrier.

## I. Introduction

This paper describes recent progress at the National Bureau of Standards (NBS) in extending high accuracy measurements of phase noise in oscillators, amplifiers, frequency synthesizers, and passive components at carrier frequencies from the rf to the millimeter region and Fourier frequencies up to 10 % of the carrier (or a maximum of about 1 GHz). An examination of existing techniques for precision phase noise measurements of oscillators[1-11] showed that present approaches which don't require a second "reference" oscillator have good resolution or noise floor for Fourier frequencies extending over only 1 or 2 decades.[7-10] Consequently, these approaches require a set of 5 to 10 references, either delay lines or high Q factor cavities, in order to adequately measure the phase noise from 1 Hz to 1 GHz from the carrier. Using the cavity approach would require a entire set of reference cavities for each carrier frequency measured. Similar considerations also apply to phase noise measurements of the other devices unless one can measure pairs of devices.

The limitations of the single oscillator techniques led us to adopt a two oscillator method for all measurements. This approach yields good resolution or noise floor from essentially dc to the bandwidth of the mixer if, for oscillator measurements, one has a comparable or better oscillator for the reference and, for other devices, either pairs of devices or a reference oscillator with comparable or better noise.

The major limitations in the accuracy of the two oscillator method (which also apply to the single oscillator methods) are the calibration of: the mixer phase-to-voltage conversion factor, the amplifier gain versus Fourier frequency, and the accuracy of the spectrum analyzer. We have developed several new calibration techniques which, when combined with previous techniques, allow us to address each of these limitations. The net result is the development of a complete measurement concept that permits one to calibrate all factors affecting the measurements of phase noise of oscillator pairs to an accuracy which typically exceeds 1 dB and in favorable cases can approach 0.4 dB. For other types of devices the limitations are similar if the noise of the reference oscillator can be neglected. The ultimate accuracy that can be easily achieved with this approach is now limited by the accuracy of the attenuators in available spectrum analyzers.

Measurements at 5 MHz and 10 GHz are described in detail, in order to illustrate this expanded two oscillator approach. Specifically we measure the mixer phase-to-voltage conversion factor multiplied by amplifier gain on all channels versus Fourier frequency using a new ultra-wideband phase modulator. We then measure the absolute mixer conversion factor multiplied by the amplifier gain at one Fourier frequency, the effect of the phase-lock loop on the measured noise voltage, and spectral density function of the spectrum analyzers. These measurements are used to normalize the relative gains of the noise measurements on the various channels. At 5 MHz we achieved accuracies of about ±0.6 dB for phase noise measurements from 20 Hz to 100 kHz from the carrier. At 10 GHz we achieved an accuracy of ±0.6 dB for phase noise measurements a few kHz to 500 MHz from the carrier. The accuracy degrades to about ±1.5 dB, 1 GHz from the carrier.

## II. Model of a Noisy Signal

The output of an oscillator can be expressed as

$$V(t) = [V_o + \varepsilon(t)] \sin(2\pi\nu_o t + \varphi(t)), \qquad (1)$$

where $V_o$ is the nominal peak output voltage, and $\nu_o$ is the nominal frequency of the oscillator. The time variations of amplitude have been incorporated into $\varepsilon(t)$ and the time variations of the instantaneous frequency, $\nu(t)$, have been incorporated into $\varphi(t)$. The instantaneous frequency is

$$\nu(t) = \nu_o + \frac{d[\varphi(t)]}{2\pi dt} . \qquad (2)$$

The fractional frequency deviation is defined as

$$y(t) = \frac{\nu(t) - \nu_o}{\nu_o} = \frac{d[\varphi(t)]}{2\pi\nu_o dt} . \qquad (3)$$

Contribution of the U.S. Government, not subject to copyright.

Power spectral analysis of the output signal V(t) combines the power in the carrier $\nu_o$ with the power in $\varepsilon(t)$ and $\varphi(t)$ and therefore is not a good method to characterize $\varepsilon(t)$ or $\varphi(t)$.

Since in many precision sources understanding the variations in $\varphi(t)$ or $y(t)$ is of primary importance, we will confine the following discussion to frequency-domain measures of $y(t)$, neglecting $\varepsilon(t)$ except in cases where it sets limits on the measurement of $y(t)$. The amplitude fluctuations, $\varepsilon(t)$, can be reduced using limiters whereas $\varphi(t)$ can be reduced in some cases by the use of narrow band filters.

Spectral (Fourier) analysis of $y(t)$ is often expressed in terms of $S_\varphi(f)$, the spectral density of phase fluctuations in units of radians squared per Hz bandwidth at Fourier frequency f from the carrier $\nu_o$. Alternately $S_y(f)$, the spectral density of fractional frequency fluctuations in a 1 Hz bandwidth at Fourier frequency f from the carrier $\nu_o$ can be used [1]. These are related as

$$S_\varphi(f) = \frac{\nu_o^2}{f^2} S_y(f) \quad \text{rad}^2/\text{Hz} \quad 0<f<\infty \; . \quad (4)$$

$S_\varphi(f)$ can be intuitively understood as,

$$S_\varphi(f) = \frac{\Delta\varphi^2(f)}{BW} \quad (5)$$

where $\Delta\varphi(f)$ is the rms phase deviation measured at Fourier frequency f from the carrier in a bandwidth BW. It should be noted that these are single-sided spectral density measures containing the phase or frequency fluctuations from both sides of the carrier. The mean squared phase modulation in a measurement bandwidth, BW, is given by

$$\Delta\varphi^2(f) = \int_{-BW/2}^{+BW/2} S_\varphi(f)\, df \; . \quad (6)$$

Other measures sometimes encountered are $\mathcal{L}(f)$, dBC/Hz, and $S_{\Delta\nu}(f)$. These are related by [1-3]

$$S_{\Delta\nu}(f) = \nu_o^2 S_y(f) \quad \text{Hz}^2/\text{Hz}$$

$$S_\varphi(f) = \mathcal{L}(\nu_0-f) + \mathcal{L}(\nu_0+f) \equiv 2\,(f) \quad (7)$$

$$\text{dBC/Hz} = 10 \log \mathcal{L}(f)$$

$\mathcal{L}(f)$ and dBC/Hz are single sideband measures of phase noise. These are the revised definitions of $\mathcal{L}(f)$ and dBC as per the most recent recommendation of the Standards Coordinating Committee for Time and Frequency of the IEEE [3]. With these revised definitions $\mathcal{L}(f)$ and dBC/Hz are now defined for arbitrarily large phase modulation, whereas previously they were restricted to small angle modulation. In some situations, especially where f becomes a sizeable fraction of $\nu_0$, the phase noise spectrum is asymmetric about the carrier. In these cases one should specify whether the upper sideband noise $\mathcal{L}(\nu_0+f)$ or the lower sideband noise $\mathcal{L}(\nu_0-f)$ is being referenced. These distinctions are becoming increasingly important as users require the specification of phase noise near the carrier where the phase excursions are large compared to 1 radian, or at Fourier frequencies which exceed the bandwidth of typical circuit elements.

The above measures provide the most powerful (and detailed) analysis for evaluating types and levels of fundamental noise and spectral density structure in precision oscillators and signal handling equipment as it allows one to examine individual Fourier components of residual phase (or frequency) modulation.

## III. A. Two Oscillator Method

Fig. 1 shows the block diagram for a typical scheme used to measure the phase noise of a precision source using a double balanced mixer and a reference source. Fig. 2 illustrates a similar technique for measuring only the phase noise added in multipliers, dividers, amplifiers, and passive components. It is very important that the substitution oscillator be at the same drive level, impedance, and at the equivalent electrical length from the mixer as the signal coming from the reference oscillator. This is dramatically illustrated in Fig. 3, discussed below. The output voltage of the mixer as a function of phase deviation, $\Delta\varphi$, between the two inputs to the mixer is nominally given by

$$V_{out} = K \cos \Delta\varphi \quad (8)$$

Near quadrature (90° phase difference) this can be approximated by

$$V_{out} = K_d \delta\varphi, \text{ where } \delta\varphi \equiv [\Delta\varphi - \frac{2n-1}{2}\pi] < 0.1 \quad (9)$$

where n is the integer to make $\delta\varphi \sim 0$. The voltage to phase conversion ratio sensitivity, $K_d$, is dependent on the frequency, the drive level, and impedance of both input signals, and the IF termination of the mixer [4]. The combined spectral density of phase noise of both input signals, the noise of the mixer, and the amplitude noise from the IF amplifiers is

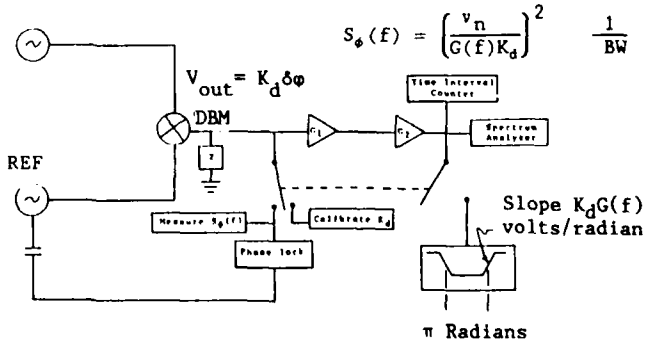



Fig. 1. Precision phase measurement system using a spectrum analyzer. Calibration requires a recording device to measure the slope at the zero crossing. The accuracy is better than 0.4 dB from dc to 0.1 $\nu_o$ Fourier frequency offset from the carrier $\nu_o$. Carrier frequencies from a few Hz to $10^{11}$ Hz can be accommodated with this type of measurement system [4].

given by

$$S_\varphi(f) = \left(\frac{V_n(f)}{G(f)K_d}\right)^2 \frac{1}{BW}, \qquad (10)$$

MEASUREMENT OF $S_\phi(f)$ BETWEEN TWO OSCILLATORS

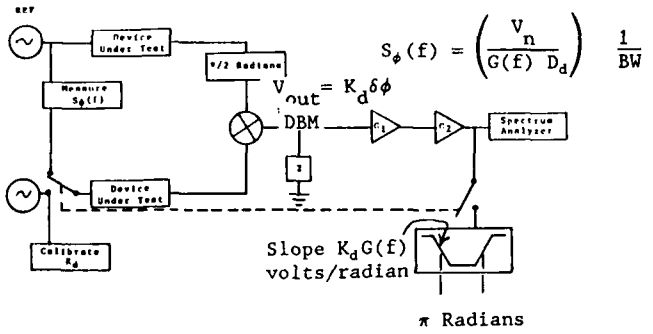



Fig. 2. Precision phase measurement system featuring self calibration to 0.4 dB accuracy from dc to 0.1 $\nu_o$ Fourier frequency offset from carrier. This system is suitable for measuring signal handling equipment, multipliers, dividers, frequency synthesizers, as well as passive components [4].

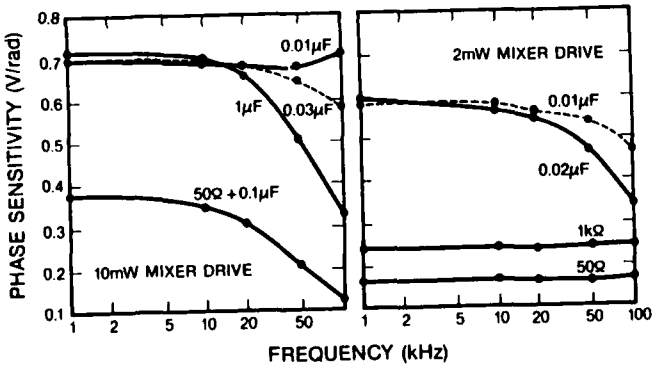



Fig. 3. Double-balanced mixer phase sensitivity at 5 MHz as a function of Fourier frequency for various output terminations. The curves on the left were obtained with 10 mW drive while those on the right were obtained with 2 mW drive. The data demonstrate a clear choice between constant, but low sensitivity or much higher, but frequency dependent sensitivity [4].

where $V_n(f)$ is the RMS noise voltage at Fourier frequency f from the carrier measured after IF gain G(f) in a noise bandwidth BW. Obviously BW must be small compared to f. This is very important where $S_\varphi(f)$ is changing rapidly with f, e.g., $S_\varphi(f)$ often varies as $f^{-3}$ near the carrier. In Fig. 1, the output of the second amplifier following the mixer contains contributions from the phase noise of the oscillators, the noise of the mixers, and the post amplifiers for Fourier frequencies much larger than the phase-lock loop bandwidth. In Fig. 2, the phase noise of the oscillator cancels out to a high degree (often more than 20 dB). Termination of the mixer IF port with 50 Ω maximizes the IF bandwidth, however, termination with reactive loads can reduce the mixer noise by ~6 dB, and increase $K_d$ by 3 to 6 dB as shown in Fig. 3 [4]. Accurate determination of $K_d$ can be achieved by measuring the slope of the zero crossing in volts/radian with an oscilloscope or other recording device when the two oscillators are beating slowly. For some applications the digitizer in the spectrum analyzer can be used to measure both the beat period and the slope in V/s at the zero crossing. The time axis is easily calibrated since one beat period equals $2\pi$ radians. The slope in volts/radian is then calculated with a typical accuracy of 0.2 dB. Estimates of $K_d$ obtained from measurements of the peak to peak output voltage induced can introduce errors as large as 6 dB in $S_\varphi(f)$ even if the amplitude of the other harmonics is measured unless the phase relationship is also taken into account [4]. $S_\varphi(f)$ can be made independent of the accuracy of the spectrum analyzer voltage reference by comparing the level of an externally IF signal (a pure tone is best), on the spectrum analyzer used to measure $V_n$ with the level recorded on the device used to measure $K_d$.

The noise bandwidth of the spectrum analyzer also needs to be verified. This calibration procedure is sufficient for small Fourier frequencies but looses precision and accuracy due to the problems illustrated in Fig. 3 and the variations of amplifier gain with Fourier frequency. If measurements need to be made at Fourier frequencies near or below the phase-lock-loop bandwidth, a probe signal can be injected inside the phase-lock-loop and the attenuation measured versus Fourier frequency.

Some care is necessary to assure that the spectrum analyzer is not saturated by spurious signals such as the power line frequency and its multiples. Sometimes aliasing in the spectrum analyzer is a problem. If narrow spectral features are to be measured it is usually recommended that a flat top window function (in the spectrum analyzer) be used. In the region where the measured noise is changing rapidly with Fourier frequency, the noise bandwidth should be much smaller than the measurement frequency. The approximate level of the noise floor of the measurement system should be measured in order to verify that it does not significantly bias the measurements or, if necessary, to subtract its effect from the results.

Typical best performance for various measurement techniques is shown in Fig. 4. The two oscillator approach exceeds the performance of almost all available oscillators from below 0.1 MHz to over 100 GHz and is generally the technique of first choice because of its versatility and simplicity. Figs. 10 - 12 give some examples. Phase noise measurements on pairs of signal sources can be made with an absolute accuracy better than 1 dB using the above calibration procedure. Such accuracy is not always attainable when the phase noise of the source exceeds that of the added noise of the components under test (see Fig. 2). The use of specialized high level mixers with multiple diodes per leg increases the phase to voltage conversion sensitivity, $K_d$ and therefore reduces the contribution of IF amplifier noise [5] as shown in Fig. 4. Phase noise measurements can generally be made at Fourier frequencies from approximately dc to 1/2 the source frequency. The

major difficulty is designing a mixer terminations to remove the source frequency from the output signal, which would generally saturate the low noise amplifiers following the mixer, without degrading the signal-to-noise ratio. As mentioned earlier the phase noise spectrum is quite likely asymmetric when f exceeds the bandwidth of the tuned circuits in the device under test. For example one expects that the phase noise at $1/2\nu_0$ is different than the phase noise at $3/2\nu_0$.

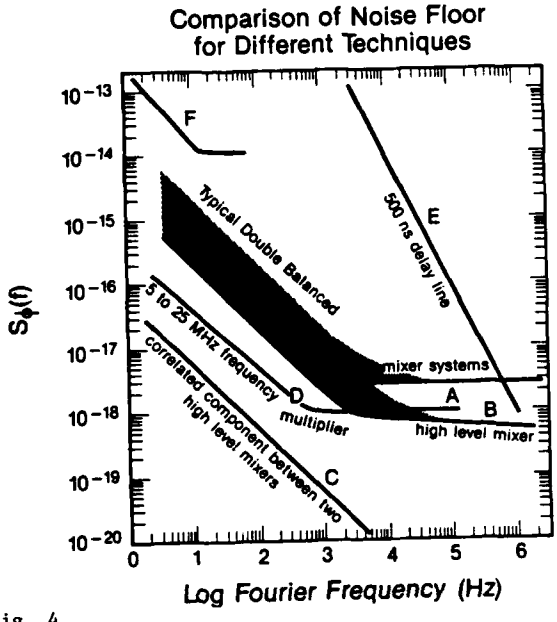



Fig. 4.

Curve A. The noise floor $S_\varphi(f)$ (resolution) of typical double balanced mixer systems (e.g. Fig. 1 and Fig. 2) at carrier frequencies from 0.1 MHz to 26 GHz. Similar performance possible to 100 GHz [5].

Curve B. The noise floor, $S_\varphi(f)$, for a high level mixer [5].

Curve C. The correlated component of $S_\varphi(f)$ between two channels using high level mixers [5].

Curve D. The equivalent noise floor $S_\varphi(f)$ of a 5 to 25 MHz frequency multiplier.

Curve E. Approximate phase noise floor of Fig. 8 using a 500 ns delay line.

Curve F. Approximate phase noise floor of Fig. 8 where a 1 ms delay has been achieved by encoding the signal on an optical carrier and transmitted it across a long optical fiber to a detector.

Most double balanced mixers have a substantial non-linearity that can be exploited to make phase comparison between the reference source and odd multiples of the reference frequency. Some mixers even feature internal even harmonic generation. The measurement block diagram looks identical to that of in Fig. 1, except that the source under test is at an odd (even) harmonic of the reference source. This method is relatively efficient (as long as the harmonics fall within the bandwidth of the mixer) for multiples up to x5 although multiples as high as 25 have been used. The noise floor is approximately degraded by the amount of reduction in the phase sensitivity of the mixer. The phase noise of the reference source is also higher at the multiplied frequency as shown in Section III.F below.

## III. B. Enhanced Performance Using Correlation Techniques

The resolution of the many systems can be greatly enhanced (typically 20 dB) by using correlation techniques to separate the phase noise due to the device under test from the noise in the mixer and IF amplifier [5, 11].

For purposes of illustration, consider the scheme shown in Fig. 5. At the output of each double balanced mixer there is a signal which is proportional to the phase difference, $\Delta\varphi$, between the two oscillators and a noise term, $V_N$, due to contributions from the mixer and amplifier. The voltages at the input of each bandpass filter are

$$V_1(\text{BP filter input}) = G_1\,\Delta\varphi(t) + C_1V_{N1}(t) \qquad (11)$$

$$V_2(\text{BP filter input}) = G_2\,\Delta\varphi(t) + C_2V_{N2}(t),$$

where $V_{N1}(t)$ and $V_{N2}(t)$ are substantially uncorrelated and $C_1$ and $C_2$ are constants. Each bandpass filter produces a narrow band noise function around its center frequency f:

$$V_1(\text{BP filter output}) = G_1[S_\varphi(f)]^{\frac{1}{2}}\,B_1^{\frac{1}{2}}\cos\,[2\pi ft + \varphi(t)]$$

$$+ C_1[S_{VN1}(f)]^{\frac{1}{2}}\,B_1^{\frac{1}{2}}\cos\,[2\pi ft + n_1(t)] \qquad (12)$$

$$V_2(\text{BP filter output}) = G_2[S_\varphi(f)]^{\frac{1}{2}}\,B_2^{\frac{1}{2}}\cos\,[2\pi ft + \psi(t)]$$

$$+ C_2[S_{VN2}(f)]^{\frac{1}{2}}\,B_2^{\frac{1}{2}}\cos\,[2\pi ft + n_2(t)]$$

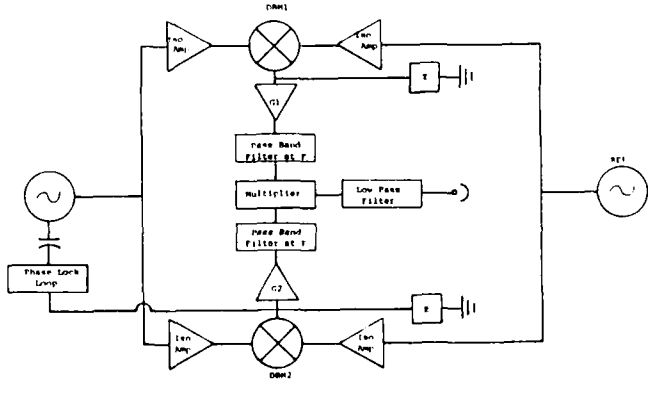



Fig. 5. Correlation phase noise measurement system.

where $B_1$ and $B_2$ are the equivalent noise bandwidths of filters 1 and 2 respectively. Both channels are bandpass filtered in order to help eliminate aliasing and dynamic range problems. The phases $\psi(t)$, $n_1(t)$ and $n_2(t)$ take on all values between 0 and $2\pi$ with equal likelihood. They vary slowly compared to 1/f and are substantially uncorrelated. When these two voltages are multiplied together and low pass filtered, only one term has finite average value.

The output voltage is

$$V_{out}^2 = \tfrac{1}{2} G_1 G_2 S_\varphi(f)\, B_1\, \tfrac{1}{2} B_2^{\frac{1}{2}} + D_1\langle\cos[7(t)]\rangle \qquad (13)$$

$$+ D_2\langle\cos[\psi(t) - n_2(t)]\rangle + D_3\langle\cos[n_1(t)-n_2(t)]\rangle,$$

so that $S_\varphi(f)$ is given by

$$S_\varphi(f) = \frac{(2)V_N^2(f)}{G_1 G_2 \sqrt{B_1 B_2}} . \qquad (14)$$

For times long compared to $B_1^{-\frac{1}{2}} B_2^{-\frac{1}{2}}$ the noise terms $D_1$, $D_2$ and $D_3$ tend towards zero as $\sqrt{t}$. Limits in the reduction of these terms are usually associated with harmonics of 60 Hz pickup, dc offset drifts, and non-linearities in the multiplier. Also if the isolation amplifiers have input current noise, they will pump current through the source resistance. The resulting noise voltage will appear coherently on both channels and cannot be distinguished from real phase noise between the two oscillators. One half of the noise power appears in amplitude and one half in phase modulation.

Obviously the simple single frequency correlator used in this example can be replaced by a fast digital system which simultaneously computes the correlated phase noise for a large band of Fourier frequencies. Typical results show a reduction in noise floor of order 20 dB over the noise floor of a single channel (See Fig. 4). The great power of this technique is that it can be applied at any carrier frequency where are available double balanced mixers. The primary limitations come from the bandwidth and non-linearities in the cross correlator [5,11].

### III. C. Reference Phase Modulation Method

Another method of determining $S_\varphi(f)$ uses phase modulation of the reference oscillator by a known amount. The ratio of the reference phase modulation to the rest of the spectrum then can be used for a relative calibration. This approach can save an enormous amount of time for measurements which are repeated a great many times. An adaptation of this approach is utilized in the new NBS phase noise system described in section IV below.

### III. D. Frequency Discriminator Methods

It is sometimes convenient to use a high-Q resonance directly as a frequency discriminator as shown in Fig. 6. The oscillator can be tuned 1/2 linewidth ($\nu_o/2Q$) away from line center yielding a detected amplitude signal of the form

$$V_{out} = G(f)k_d Q dy(f)\,[V + \varepsilon(t)] \qquad (15)$$

This approach mixes frequency fluctuations between the oscillator and reference resonance with the amplitude noise of the transmitted signal. By using amplitude control (e.g. by processing to normalize the data), one can reduce the effect of amplitude noise. [5] The measured noise at the detector is then related to the phase fluctuation of the reference resonance by

$$S_\varphi(f) = \left(\frac{\nu_o V_N(f)}{fQk_d G(f)}\right)^2 \frac{1}{BW} \quad \text{for} \quad f < \frac{\nu_o}{2Q} . \qquad (16)$$

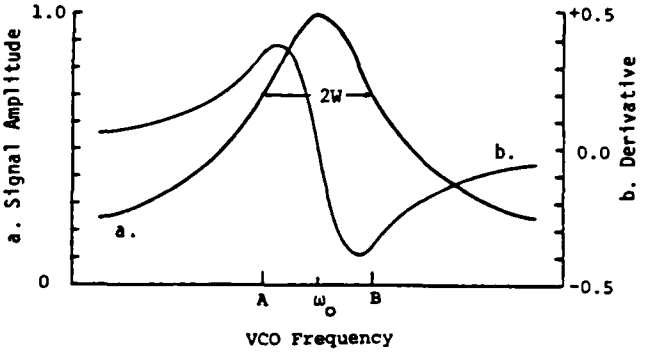



Fig. 6. High-Q resonance used as a frequency discriminator. Note that the peak response is displaced from the center of the resonance by about the half bandwidth.

This approach normally has the limitations that f must be small compared to the linewidth of the cavity, and the effect of residual amplitude noise is difficult to remove; however, no reference source is needed. The calibration factors G(f)KdQ can be measured even for Fourier frequencies larger than $\nu_0/2Q$ by stepping the source frequency an amount dy (which is small compared to 1/2Q) and measuring the output voltage versus the modulation frequency, f.

Differential techniques can be used to measure the inherent frequency (phase) fluctuations of two high-Q resonators as shown in Fig. 7 [7]. The output voltage is of the form $V_{out} = 2QK_dG(f)\,dy(f)$. The phase noise spectrum of the resonators is then obtained using Eq. 4.

$$S_\varphi(f) = \left(\frac{\nu_o V_N(f)}{2QfK_d G(f)}\right)^2 \frac{1}{BW} \qquad f < \frac{\nu_o}{2Q} \qquad (17)$$

Measurement of the Inherent Phase Noise in High-Q Resonators

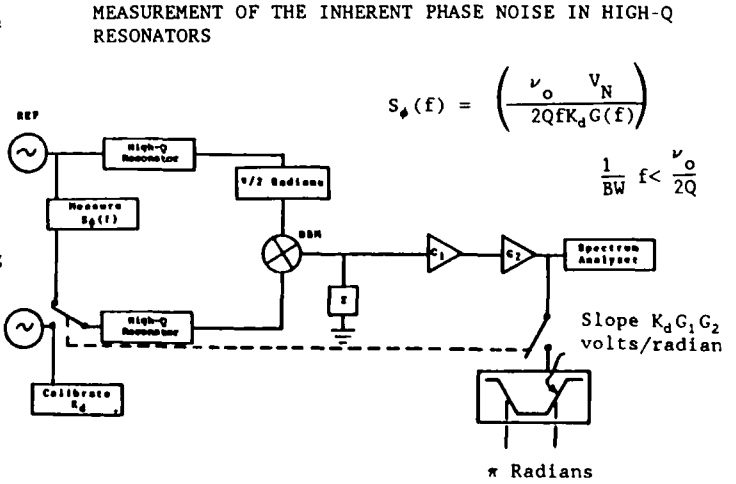



Fig. 7. Differential frequency discriminator using a pair of high-Q resonators. In this approach the phase noise of the source tends to cancel out.

The phase noise in the source can cancel out by 20 to 40 dB depending on the similarity of resonant frequencies Q's and the transmission properties of the two resonators. This approach was first used to demonstrate that the inherent frequency stability of precision quartz resonators exceeds the performance of most quartz crystal controlled oscillators [7].

If only one resonance is used, the output includes the phase fluctuations of both the source and the resonator. The calibration is accomplished by stepping the frequency of the source and measuring the output voltage, i.e., $\Delta V = G(f)K^1(f)\,\Delta\nu_1$. From this measurement the phase spectrum can be calculated as

$$S_\varphi(f) = \left( \frac{V_N^2(f)}{f\,G(f)K^1(f)} \right)^2 \left( \frac{1}{BW} \right) . \qquad (18)$$

Fig. 8A shows one method of implementing this approach at X-band. The cavity has a loaded quality factor of order 25,000. Fig. 8B shows the measured frequency discriminator curve. Note that $K^1(f)$ is constant for $f<\nu/(2Q)$ and decreases at values of f larger than the half bandwidth of the resonance as

$$K^1(f) \sim \frac{1}{1 + \left(\frac{2Qf}{\nu_o}\right)^2} \qquad (19)$$

The frequency dependence shown in eq. (19) can be accurately determined by measuring $\Delta V$ when stepping the source an amount $\Delta\nu$, which is small compared to $\nu_0/Q$, versus the modulation frequency, f. This approach has poor resolution near the carrier and limited high frequency response. Therefore the measurement of phase from close to the carrier out to 10% of the carrier could require a large set of cavities with different Q factors. In order to achieve the required Q factor for close in measurements it may even be necessary to use cryogenic techniques.

## III. E. Delay Line Method

Another different approach uses a delay line to make a pseudo reference which is retarded relative to the incoming signal [7-10] as shown in Fig. 9.

The mixer output is of the form

$$V_{out} = 2\pi t_d K_d \nu_o dy , \qquad (20)$$

and the input phase noise is given by

$$S_\varphi(f) = \left( \frac{V_N(f)}{2\pi f \tau_d G(f) K_d} \right)^2 \frac{1}{BW} , \quad f < \frac{1}{\tau_d} \qquad (21)$$

This approach is often used at microwave frequencies when only one oscillator is available. In this technique the ability to resolve phase noise close to the carrier depends on the delay time. For example, if f = 1 Hz and $\tau_d$ = 500 ns, then, $(2\pi f\tau_d)^2 \sim 10^{-11}$. The noise floor is 110 dB higher at f = 1 Hz than that of the two oscillator method, decreasing as $1/f^2$. Recent advances make it possible to encode the rf signal on an optical signal which then can be transmitted down an optical fiber to achieve delays up to the order of $10^{-3}$ s with an increase in the noise floor to approximately -140 dB relative to 1 rad$^2$/Hz. The noise floor can be reduced by ~20 to 40 dB using the correlation techniques described above [11]. Note that the range of Fourier frequencies is usually limited to less than $\sim 1/\tau_d$. This technique normally has good resolution over 1½ to 2 decodes in Fourier frequency. Therefore, measurements of phase noise from close to the carrier out to 10% of the carrier require a large set of different delay lines and hardware including optical delay lines, associated lasers, modulators, and detectors.

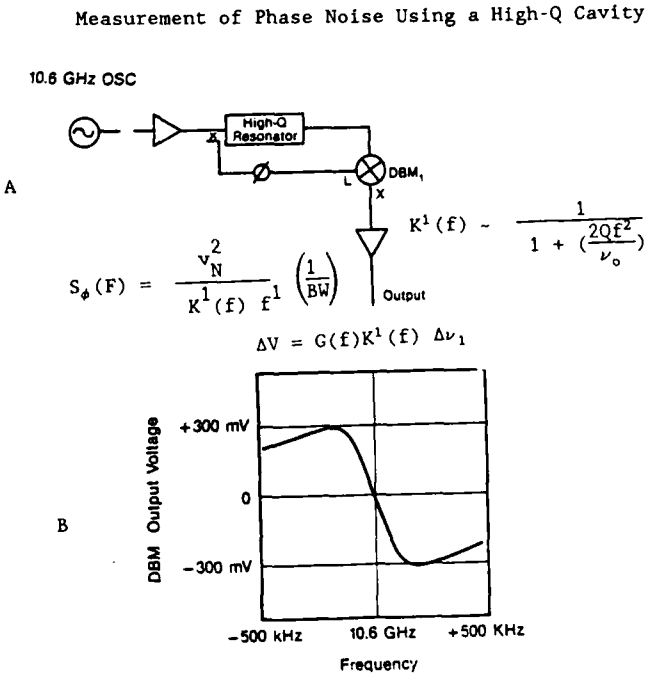



Fig. 8. A. block diagram of a high Q resonator used as a frequency discriminator. B. Frequency discriminator curve for the scheme shown in A used at X-band with a cavity having a loaded quality factor of approximately 25,000.

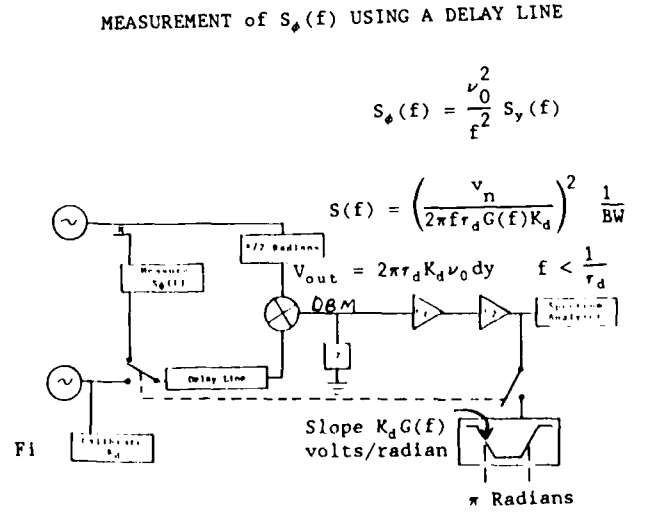



Fig. 9. Delay line frequency discriminator.

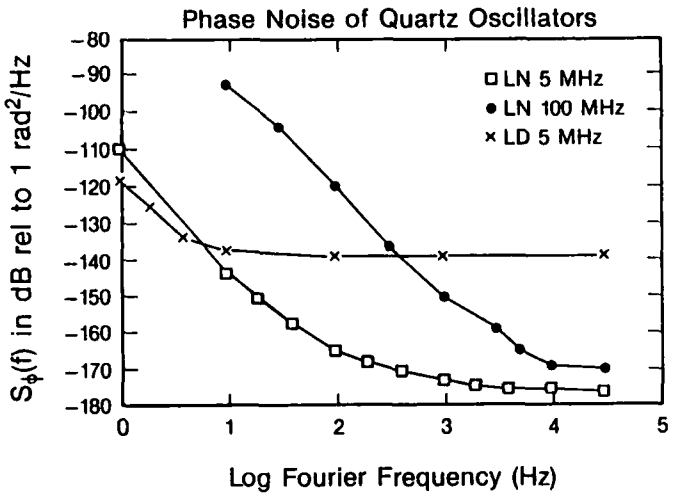



Fig. 10. Phase noise performance of selected quartz oscillators. The LN 5 MHz oscillator is driven at a high level to reduce the wideband noise while the LD 5 MHz oscillator is driven at a lower level to obtain low phase noise close to the carrier.

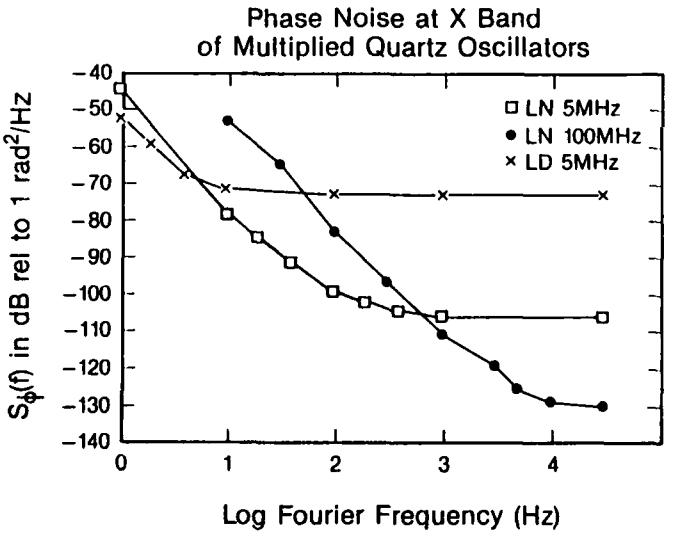



Fig. 11. Phase noise of the oscillators of Fig. 10 if multiplied to X-band in a perfect frequency multiplier.

### III. F. Multiplication/Division

The use of perfect frequency multipliers (or dividers) between the signal source and the double balanced mixer increases (decreases) the phase noise level [12] as

$$S_\varphi\nu_1(f) = \left(\frac{\nu_2}{\nu_1}\right)^2 S_\varphi\nu_1(f) \qquad (22)$$

where $\nu_1$ is the initial carrier frequency and $\nu_2$ is the final carrier frequency. This can be used to either increase or decrease the phase sensitivity of the mixer system. Fig. 4 shows the noise of a specialized 5 to 25 MHz multiplier referred to the 5 MHz input. A potential problem with the use of the multiplier approach comes from exceeding the linear range of the mixer. Once the phase excursion, $\Delta\varphi$, exceeds about 0.1 radian, non-linearities start to become important and at $\Delta\varphi$ ~1 radian, the measurement is no longer valid [12]. An additional practical problem is that low noise multipliers are usually narrowband devices. Each significantly new frequency generally requires a new set of frequency multipliers.

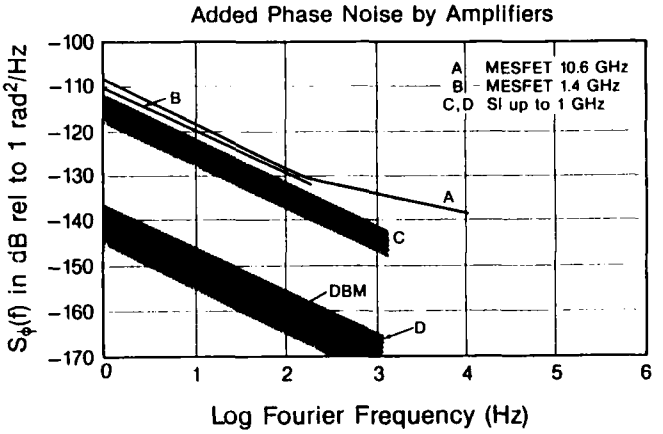



Fig. 12. Curves A and B show the phase noise added by selected GaAs MESFET amplifiers. Curve C shows the phase noise added by a typical common emitter silicon bipolar transistor with a "good" rf bypass on the emitter lead. Curve D shows the typical performance of the same amplifier with a small unbypassed impedance (approximately 1/transconductance) in the emitter lead. The added phase noise is generally independent of frequency over a very large range.

### IV. A. The New NBS Phase Noise Measurement Systems

The new NBS phase noise measurement systems are a combination of the traditional two oscillator approach shown in Figs. 1 and 2 plus the reference phase modulation technique mentioned in section IIIC. The complete block diagram is sketched in Fig. 13. This approach yields the widest possible bandwidth and the lowest phase noise of a single channel system. It does, however, require the use of two sources for oscillator measurements. From hardware considerations we generally use 3 different phase noise measurement systems. Test set A accepts carrier frequencies from 5 to 1300 MHz and can measure the phase noise from 1 Hz to about 10% of the carrier or a maximum of 100 MHz. Test set B accepts carrier frequencies from 1 GHz to 26 GHz and can measure the phase noise from 0.01 Hz to about 500 MHz from the carrier. Test set C accepts carrier frequencies from about 2 to 26 GHz and can measure the phase noise from 0.01 Hz to 1 GHz from the carrier. Test set D accepts carrier frequencies from 33 to 50 GHz in WR22 waveguide and can measure the phase noise from 0.01 Hz to about 1.3 GHz from the carrier.

The construction of the phase modulator between the reference source and the mixer will be described in

detail elsewhere [12]. The low pass filter section is used in order not to saturate the amplifiers with the carrier feedthrough signal from the mixer. One dc amplifier is used for phase noise measurements from dc to 100 kHz from the carrier. One ac amplifier is used for phase noise measurements from 50 kHz to 32 MHz. This range is well matched for one of our spectrum analyzers. The wideband ac amplifier has a bandwidth of 50 kHz to 1.3 GHz and is used when the desired measurement bandwidth exceeds 32 MHz. The wide bandwidth spectrum analyzer also provides a convenient way to observe the gross features of the output phase noise and to identify any major spurious outputs if present. In order to obtain the most accurate measurement of the phase noise it is, however, necessary to measure the amplitude of the first IF signal at about 21 MHz in order to avoid the variations in gains of the log amplifiers with various environmental factors.

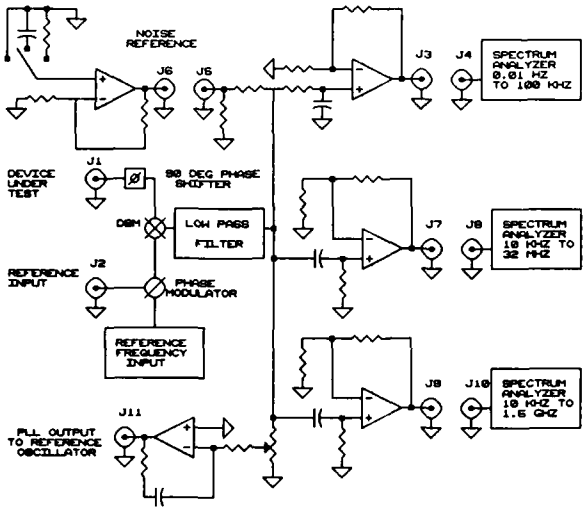



Fig. 13. Generalized block diagram of the new NBS phase noise measurement systems. The phase noise of carrier frequencies from 1 MHz to 100 GHz can be measured by varying the components in the phase shifters and mixers. Dedicated measurement systems covering 5 MHz to 50 GHz are described in the text.

## IV. B. Measurement Sequence

1) The output power of the two sources to be measured is typically set to between +5 and + 13 dBm at the mixer. This takes into account the insertion loss of the phase modulator. If the oscillators don't posses sufficient internal isolation to prevent unwanted frequency pulling, isolators (or isolation amplifiers) are generally inserted between the sources and the mixer.

2) The absolute sensitivity of the mixer and the dc amplifier for converting small changes in phase to voltage changes is determined in a way similar to the traditional method, namely by allowing the two oscillators to slowly beat. The output of the dc amplifier is recorded by the digitizer in the FFT connected to the dc amplifier in order to accurately determine the period of the beat. In test sets C and D an additional 50 MHz digitizer is used to average and record the beat frequency. The time scale of the digitizer is then expanded to approximately 10% of the beat period and pretrigerred at about -2V in order to accurately determine the slope of the output in volts per radian. This calculation, shown in eq. (23) below, is typically accurate ± 2% or 0.2 dB.

$$K = \text{Volts/Second})(\text{Period}/(2\pi) \qquad (23)$$

3) The two sources to be measured are phase locked together with sufficient bandwidth that the phase excursions at the mixer are less than 0.1 radian. The necessary phase lock gain is calculated using an estimate for the noise of the oscillators and the tuning rate for the reference oscillator. This is then verified by noting the peak to peak excursions of the dc amplifier and using the measured conversion sensitivity measured in step 2 above. If the peak phase excursions are in excess of 0.1 radian, then the phase lock loop bandwidth is increased (if possible) in order to satisfy this condition.

4) The modulator is driven by a reference frequency (typically at +7 to +10 dBm) which steps through the Fourier frequencies of interest and the detected RMS voltage recorded on the appropriate spectrum analyzer. This approach accurately yields the relative gains of each amplifier and its respective spectrum analyzer since it automatically accounts for the effect of the phase lock loop and residual frequency pulling as well as the termination of the mixer and the variations in gain of the various amplifiers with Fourier frequency. The amplitude of the phase modulation on the carrier is constant in amplitude to better than ± 1.5 dB (typically ±0.2dB for f < 500 MHz) for reference frequencies dc to about 10 % of the carrier frequency or a maximum of 1 GHz. Initial measurements of the prototype modulator are shown in Fig. 14.

This measurement is then combined with the measurement of the absolute mixer sensitivity multiplied by the gain of the dc amplifier described in step 2 above. The absolute gain of all the amplifiers shown in Fig. 13 can generally be determined to an accuracy of ± 0.3 dB (1.5 dB for Fourier frequencies from 500 MHz to 1 GHz) over the Fourier frequencies of interest.

5) Next the spectral density function of the FFT is verified. The level of the noise determined by the FFT for the input of the noise source amplifier sequentially shorted to ground, connected to ground through a 100 kΩ metal film resister, and connected to ground through 200 pF, is then recorded. From these data one can determine the inherent noise voltage and noise current of the noise source amplifier plus the FFT as well as the noise of the resistor to about ± 0.25 dB which is the stated accuracy of the FFT. This primary calibration of the FFT can be carried out from about 20 Hz to over 50 kHz. Above 50 kHz, the noise gain of the amplifier we used contributes a significant amount of noise. With some compensation the noise is flat to within ±0.2 dB from 20 Hz to 100 kHz. Next the noise source is switched into the output of the mixer and the relative noise spectrum of all the spectrum analyzers is calibrated by knowing their relative gains. This procedure verifies the voltage references and noise bandwidths of the various spectrum analyzers.

6) The noise voltage is recorded on the three spectrum analyzers over the Fourier frequencies of interest, generally the same one used in step 4

above. The measured noise voltages are scaled using the measured gains and the spectral density of phase noise calculated. The overlapping ranges of the various spectrum analyzers allows one the opportunity to compare the measurements on the three spectrum analyzers. Typically one can obtain $S_\varphi(f)$ of the oscillator pair to an accuracy of about ±0.6 dB at Fourier frequency above 100 Hz. The agreement with repeat measurements is often of order ±0.2 dB as shown in Fig. 15.

7) The noise floor of the system is determined by driving both sides of the measurement system with one oscillator having a similar power and impedance level as that used in these measurements. If the noise floor is within 13 dB of that measured in step 7 above, corrections are made to the measurement data to remove the bias generated by the noise floor.

Measurements of the amplitude accuracy of the phase modulation side bands generated by the prototype phase modulators are summarized in Fig. 15. The same modulator was used for carrier frequencies from 5 to 300 MHz. The error in the phase modulation amplitude is less than 0.5 dB for modulation frequencies from dc to 10% of the carrier. The 10 GHz modulator also maintains an accuracy of better than 0.5 dB out to 500 MHz from the carrier. At 1 GHz the modulation amplitude is 1.5 dB high. Once this is measured it can be taken into account in the calibration procedure. The 45 GHz modulator results shown in Fig. 15 should be attainable over the entire WR22 wavequide bandwidth.

The performance that can be obtained with this measurement technique is illustrated by actual phase noise data on oscillator pairs shown in Figs. 15 and 16. Typical accuracies are ±0.6 dB with a noise floor at about - 175 dB relative to 1 radian$^2$/Hz. The corrections applied to the raw data at 10 GHz are shown in Fig. 17. At low frequencies the effect of the phase-locked loop is apparent while at the higher frequencies the roll-off of the amplifiers are important. These effects have been emphasized here in order to examine the ability of the calibration process to correct for instrumentation gain variations.

## V. Conclusion

We have analyzed several traditional approaches to making phase measurements and found that they all lacked some element necessary for making phase noise measurements from essentially dc out to 10 % of the carrier frequency with good phase noise floors and an accuracy of order 1 dB. By combining several of the techniques and adding a phase modulator which is exceptionally flat from dc to about 10 % of the carrier frequency, we have been able to achieve excellent phase noise floors, bandwidths of at least 10 % of the carrier, and accuracies of order ±0.6 dB.

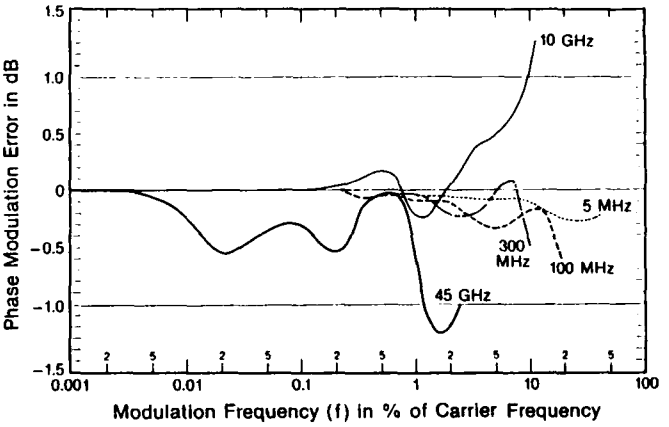



Fig. 14 Measurement of the amplitude error of modulation signal versus Fourier frequencies f, for these prototype phase modulators. Curves labeled 5, 100, and 300 MHz were obtained with the modulator used in 5 to 1,300 MHz test set. The curve labeled 10 GHz was obtained with the modulator for the 2 to 26 GHz test set. The curve labeled 45 GHz was obtained with the WR22 test set.

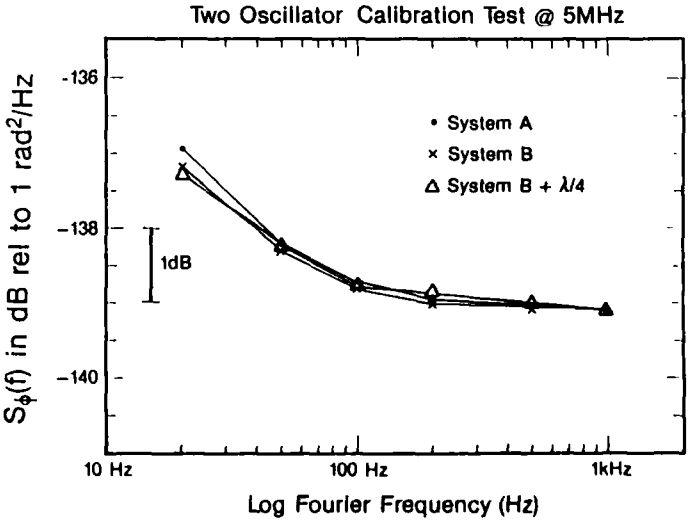



Fig. 15 Demonstration of calibration accuracy for two oscillator concept. The curve labeled System A shows the measured phase noise of a pair of 5 MHz oscillators using the test set shown in Fig. 13. The curve labeled System B shows the measured phase noise of the same pair of 5 MHz oscillators using a totally separate measurement system with the oscillators held in phase quadrature with the measurement test set of A. The curve labeled System B + λ/4 shows the phase noise of the same pair of oscillators using test set B with and extra cable length of λ/4 inserted into each signal path. The agreement between the three curves is in the worse case ± 0.15 dB

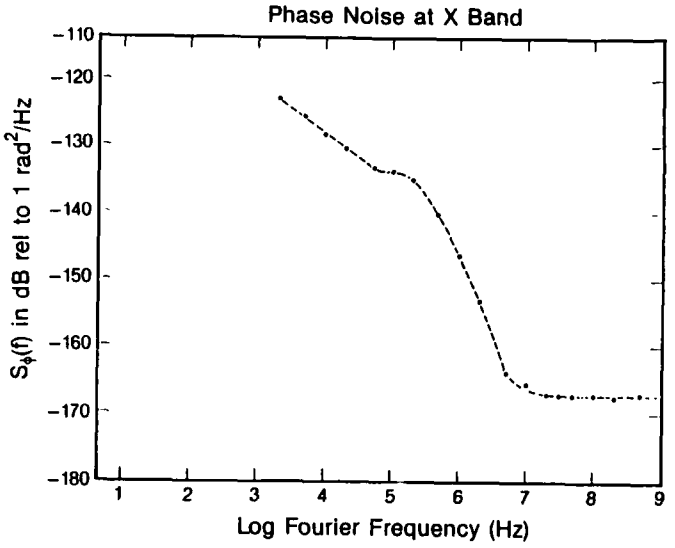



Fig. 16 Phase noise measurement on a pair of 10.6 GHz sources using the new NBS measurement technique.

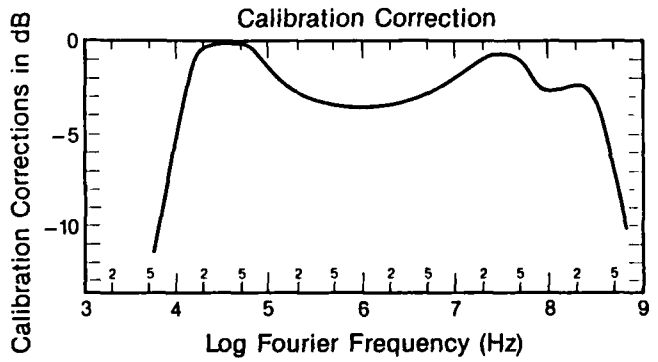



Fig. 17 Correction factor applied to the measurement data made to obtain the results of Fig. 16.


## Acknowledgements

The authors are grateful to many colleagues, especially David W. Allan, James C. Bergquist, Andrea DeMarchi, David J. Glaze, James E. Gray, David A. Howe, John P. Lowe, Samuel R. Stein and Charles Stone for many fruitful discussions on this topic and the Calibration Coordination Group for the funding to improve the accuracy and bandwidth of phase noise metrology.


## References


1. J. A. Barnes, A. R. Chi, L. S. Cutler, D. J. Healey, D. B. Leeson, T. E. McGunigal, J. A. Mullen, Jr., W. L. Smith R. L. Sydnor, R. F. C. Vessot, G. M. Winkler, Characterization of Frequency Stability, Proc. IEEE Trans. on I & M 20, 105-120 (1971).

2. J. H. Shoaf, D. Halford, and A. S. Risley, Frequency Stability Specifications and Measurement, NBS Technical Note 632, (1973). Document available from US Government printing office. Order SD at #C13.46:632.

3. D.W. Allan, H. Hellwig, P. Kartaschoff, J. Vanier, J. Vig, G.M.R. Winkler, and N.F. Yannoni, Standard Terminology for Fundamental Frequency and Time Metrology, to be published in the Proc. of the 42nd Symposium on Frequency Control, Baltimore, MD, June 1-4, 1988.

4. F. L. Walls, and S. R. Stein, Accurate Measurements of Spectral Density of Phase Noise in Devices, Proc. of 31st SFC, 335-343, (1977). (National Technical Information Service, Sills Building, 5825 Port Royal Road, Springfield, VA 22161).

5. F. L. Walls, S. R. Stein, J. E. Gray, and D. J. Glaze, Design Considerations in State-of-the-Art Signal Processing and Phase Noise Measurement Systems, Proc. 30th Ann. SFC, 269-274 (1976). (National Technical Information Service, Sills Building, 5285 Port Royal Road, Springfield, VA 22161).

6. R. L., Barger, M. S. Soren, and J. L. Hall, Frequency Stabilization of a cw Dye Laser, Appl. Phys. Lett. 22, 573 (1973).

7. 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 I & M 24, 15-20 (1975).

8. A. S. Risley, J. H. Shoaf, and J. R. Ashley, Frequency Stabilization of X-Band Sources for Use in Frequency Synthesis into the Infrared, IEEE Trans. on I & M, 23, 187-195 (1974).

9. J. R. Ashley, T. A. Barley, and G. J. Rast, The Measurement of Noise in Microwave Transmitters, IEEE Trans. on Microwave Theory and Techniques, Special Issue on Low Noise Technology, (1977).

10. A. L. Lance, W. D. Seal, F. G. Mendoza, and N. W. Hudson, Automating Phase Noise Measurements in the Frequency Domain, Proc. 31st Ann. Symp. on Freq. Control, 347-358 (1977).

11. A. L. Lance and W. D. Seal, Phase Noise and AM Noise Measurements in the Frequency Domain at Millimeter Wave Frequencies, from Infrared and Millimeter Waves, Ken Button Ed., Academic Press, NY 1985.

12. F. L. Walls and A. DeMarchi, RF Spectrum of a Signal After Frequency Multiplication Measurement and Comparison with a Simple Calculation, IEEE Trans. on I & M 24, 210-217 (1975).

13. F.L. Walls, A New Phase Modulator for Wideband Phase Noise Measurement Systems, to be submitted to IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control.

42nd Annual Frequency Control Symposium - 1988

# TECHNIQUE FOR MEASURING THE ACCELERATION SENSITIVITY OF SC-CUT QUARTZ RESONATORS

Milton H. Watts
Errol P. EerNisse
Roger W. Ward
Robert B. Wiggins

Quartztronics, Inc.
1020 Atherton Drive
Building C
Salt Lake City, UT 84123

## SUMMARY

We present a simple method for measuring the acceleration sensitivity of doubly-rotated quartz resonators. This method utilizes readily available electronic components and the voltage-frequency effect in doubly-rotated cuts to determine both the phase and magnitude information. The method is based on monitoring vibration-induced FM sidebands of an oscillator with a spectrum analyzer and using acceleration compensation for SC-cuts via the voltage frequency effect. An op amp with reversible polarity is used to apply a voltage, which is proportional to an accelerometer output, to the crystal. Shaker frequency and rf are separated by simple filter components. Since the goal is to observe a reduction or increase in FM sideband magnitude, the phase and magnitude requirements of the compensation circuit are less than that required for high level compensation.

Examples will be given of the method. Typical reductions or enhancements of 6 to 20db can be seen easily and are sufficient to establish the absolute phase (sign) of each gamma vector component. Data is presented to show that acceleration sensitivities of $1x10^{-10}$/g can be measured with confidence.

## INTRODUCTION

The acceleration sensitivity of quartz resonators arises from the stresses caused by the mass of the resonator blank reacting against the resonator mounting structure during acceleration. A common method of representing the acceleration sensitivity is by the gamma vector, made up of the frequency shifts observed for unit acceleration in three orthogonal axes[1]. Typical numbers for the components of the gamma vector range from low $10^{-9}$/g to mid $10^{-10}$/g. The acceleration-induced frequency response of a resonator is then the dot product of the gamma vector and the applied acceleration. In research and development, a measurement technique for the gamma vector is required for a wide range of frequencies, resonator configurations, and overtone operation. Often measurements must be done on unsealed devices: the Q will typically be lower, and the resistance higher, than under ideal, sealed conditions. Many situations require knowing both the magnitude and the direction of the gamma vector. These include anticipation of using various compensation techniques such as opposed-pair compensation, accelerometer feedback compensation, and spring/mass compensating systems[1]. Also, when applying theories regarding the acceleration performance of a device, a full description of the gamma vector is needed.

Measurement techniques commonly used have a number of disadvantages which make them less than ideal in R & D environments. A method commonly used is observing, with a spectrum analyzer, the FM sidebands generated by sinusoidal acceleration of the crystal under test[2]. With this method, one is able to obtain the magnitude, but not the phase, of the gamma vector. Another common method is FM demodulation[3]. This technique requires virtually noise free demodulation circuitry to detect sensitivities as low as $1x10^{-10}$/g. The voltage controlled oscillators used in such circuitry usually have a small frequency range over which they are useful. Retooling for crystals of varying frequencies can become time consuming and expensive. Another method which is used is the 2-g tip-over test. Because of the small frequency shifts involved, an extremely stable temperature environment and a high resolution counter is necessary to obtain useful results for crystals with low sensitivities. However, the 2-g tip-over test does provide the sign (phase) of the gamma vector. The reader is referred to Filler's tutorial on acceleration effects on crystals[1] for more details.

Errors can occur in the measurement of small gamma components in the low $10^{-10}$/g to high $10^{-11}$/g range. The errors are related to cable reactance modulation due to vibration of power cables and signal cables passing to and from the shaker table. The problem was observed in varying degrees for all three different methods tried in this work.

Presented here is a versatile method of determining both the magnitude and direction of the gamma vector for crystals with a

CH2588-2/88/0000-442 $1.00 C 1988 IEEE

broad range of frequencies. The equipment used is readily available laboratory equipment. Cable modulation effects are addressed.

## THEORETICAL BASIS

When a crystal is under vibration, its resonant frequency is shifted due to stresses induced by the interaction of the resonator blank and its mounting structure. The frequency shift $df_a$ is proportional to the acceleration. The frequency/voltage effect of an SC cut crystal (or other doubly rotated crystals) also results in a shift of the resonant frequency ($df_v$)[4]. If the applied voltage and the acceleration are correlated such that $df_v$ and $df_a$ are equal in magnitude but opposite in sign, an apparent decrease in the g-sensitivity of the device can be observed. This effect has been used previously by Rosatti and Filler as a compensation technique[5].

In the case of singly-rotated quartz cuts, such as the AT-cut, that do not have a voltage-frequency effect, one could use a voltage controlled oscillator circuit to pull the oscillator frequency to compensate the acceleration effect.

As discussed recently[6], the phase relationship between the acceleration and the applied voltage signal must be accurately controlled near 180° to achieve compensation. Figure 1 shows a theoretical plot of how much sideband suppression can be accomplished for phases approaching 180° as the ratio $df_v/df_a$ is varied. In order to observe a significant decrease in g-sensitivity, the amplitude and phase of the compensating signal must match the amplitude and phase of the acceleration effect of the device quite closely. Assuming precise amplitude control, the phase of the compensating signal must match the phase of the acceleration effect to within +/-5 degrees to achieve 20db sideband reduction. We have found that 10-20 db is adequate to identify the phase of $df_v$ that reduces the acceleration-induced $df_a$ sideband.

Logically, if the quartz crystal is turned upside down, the gamma vector component should be equal in magnitude and opposite in sign for the two orientations. This has never occured in any experiment we have tried when dealing with gamma responses below $2x10^{-10}$/g. We attribute this to the fact that the cables passing to and from the shaker, no matter what test configuration is used, are modulated by the vibration. The cable vibration sinusoidily detunes the oscillator, to varying degrees, depending upon the test configuration and the Q of the crystal or circuit under test. This creates an effect that adds vectorially to the gamma vector being studied, and is therefore a source of error. (In the configuration to be described this effect is repeatable, and on the order of $1x10^{-10}$/g at 50 MHz. The effect is much less significant at lower frequencies. When testing the same crystal either in the setup described herein or when shaking the complete oscillator circuit, some cable effects are observed.)

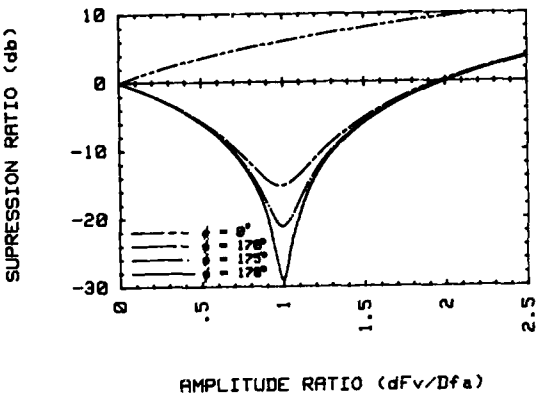



FIGURE 1. Maximum expected sideband suppression as a function of the amplitude ratio $df_v/df_a$ of two sinusoidal modulating functions at various phase angles. Sideband suppression occurs when the phase of the acceleration effect and the voltage effect approach 180°. If the phase is close to zero, the two effects add, and no sideband suppression occurs.

The present experimental equipment and procedures are based on holding cable effects constant and measuring the crystal both right side up and upside down. Since the cable effects add primarily in-phase or out-of-phase with the acceleration effect of the crystal, the algebraic average of the two readings is approximately equal to the crystal effect, with the cable effect cancelling. In other words, if $df_c$ is the cable effect, and assuming it is additive to the $df_a$ in the +x direction, it will be subtractive from the $df_a$ measured in the -x direction. The measured df is, thus

$$df(+x) = df_c + df_a \quad [1]$$

$$df(-x) = df_c - df_a \quad [2]$$

and the desired $df_a$ is given by

$$df_a = [df(+x) - df(-x)]/2. \quad [3]$$

## EQUIPMENT USED

A block diagram of the test system is shown in Figure 2. The test equipment used includes a LING Electronics Inc. model 203B shaker table driven by a variable frequency audio oscillator and an audio amplifier. Attached to the shaker is a fixture for mounting the crystal in any of 6 directions (+/- X,Y,Z).

Figure 3 shows the crystal fixture. Rigid wiring is used inside the cube to minimize motion during vibration. The cube can be turned over to test all six axes. The cables running to the shaker are carefully fixed so that their effects remain constant, allowing cancellation using Equation 3 above.

A Saunders & Associates 100HF CI meter was chosen as the oscillator circuit in

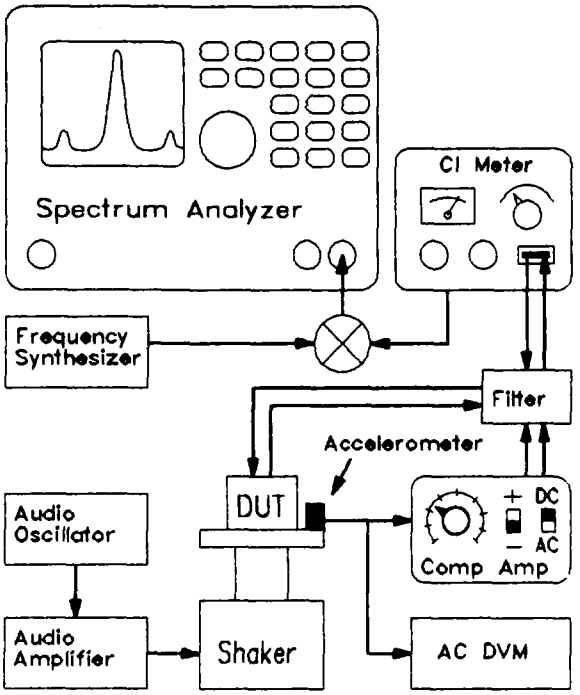



FIGURE 2. Block diagram of the test system used to determine the acceleration sensitivity of doubly-rotated quartz resonators. The spectrum analyzer is used to measure the magnitude of the sensitivity vector. Determining the polarity of compensation required for sideband suppression yields the absolute sign of the acceleration sensitivity vector.

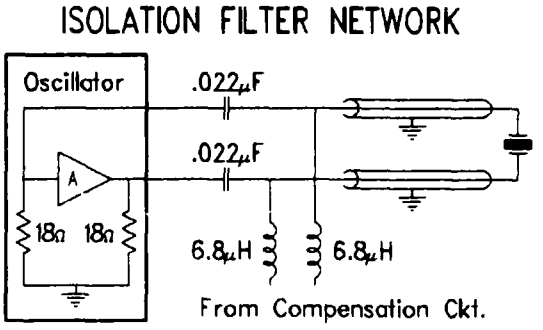



FIGURE 4. Typical filter network used to isolate the oscillator from the compensation circuitry. Actual component values will vary depending on the crystal frequency and oscillator circuit used.

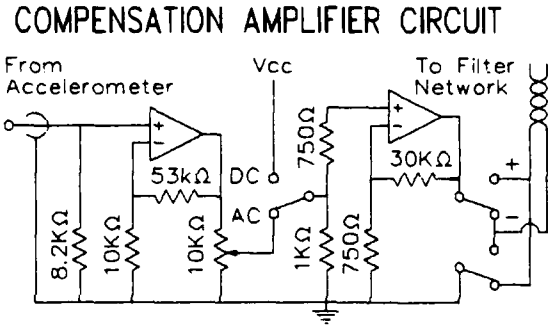



FIGURE 5. Simplified schematic of the compensation amplifier showing the buffer and gain stages, amplitude control, and switching schemes for phase reversal and D.C. voltage effect polarity checking.

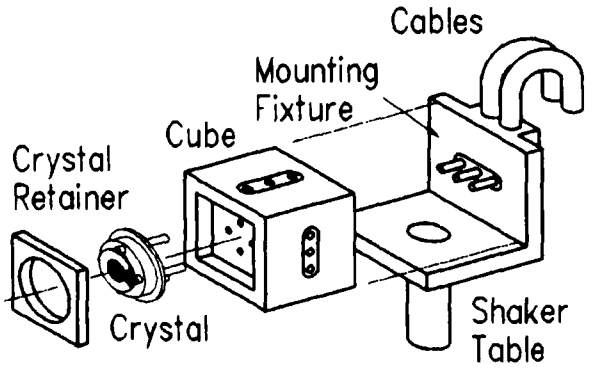



FIGURE 3. Fixture for measuring the acceleration sensitivity of quartz resonators. The crystal is secured in the cube, which can be rotated into +/- X,Y, and Z orientations. The cube is secured to the mounting fixture by a clamp (not shown). Cable effects are repeatable and do not change with cube orientation.

order to drive crystals with a broad range of frequencies and impedances. It is connected to the crystal through a pair of coaxial cables clamped at both ends. Figure 4 shows a block diagram of the oscillator and cable. The output of the oscillator is observed on an HP3585A spectrum analyzer. High frequency crystals (greater than 40 MHz) require a mixer with a relatively low noise reference oscillator (HP3325A Frequency Synthesizer) to shift the signal to a frequency within the operating range of the spectrum analyzer.

The compensation signal is obtained from an accelerometer mounted on the fixture that holds the crystal. (The shaker table driving signal may also be used for compensation as long as the phase requirements previously mentioned are satisfied.) The signal is fed through a variable gain op-amp circuit with switch selectable polarity and an option to put a dc voltage across the crystal. This op-amp circuit is shown in Figure 5. The output of this device is fed back to the crystal leads via a simple filter circuit as shown in Figures 2 and 4.

## MEASUREMENT PROCEDURE

The crystal under test is clamped in the cube, which provides for easy, repeatable alignment of the crystal in any of 6 directions. The polarity of the voltage/frequency effect of the crystal is determined by observing a slight shift in frequency when a dc voltage is applied to the crystal. The device is then shaken at approximately $3g_{rms}$ and the sideband amplitudes are measured using the spectrum analyzer.

The phase of the gamma vector in the axis of vibration is measured by slowly increasing the level of compensation from zero while observing the sideband amplitude. Referring to Figure 1, if the compensation signal is of the correct phase to cancel the crystal's acceleration effect, the sideband amplitude will first decrease, and then increase. The opposite phase will show only an increase in sideband amplitude as shown in Figure 1 for 0° phase. Switching the phase of the compensation without changing its amplitude will result in a fairly large difference in the amplitude of the sideband peaks. By correlating this information with the polarity of the accelerometer output relative to the crystal acceleration, the magnitude and direction of the gamma vector component can be uniquely determined.

The polarity of the voltage frequency effect reverses when the crystal is turned upside down in our fixture because the pin polarity of the crystal is reversed, while the polarity of the compensating circuit stays the same. It is important to keep track of this! By referencing everything to the dc shift test, one can keep the signs straight. If the cable capacitance effect is predominant, the phase comparator circuit requires a polarity reversal when the SC-cut crystal is inverted. In contrast, when the crystal acceleration effect predominates, no polarity reversal is needed because both the acceleration effect and voltage effect in the crystal change sign.

Measurements are made for opposing directions. The true answer (if cable effects are constant and in phase with the crystal acceleration effect) is found by using Equation 3.

## RESULTS

Several shaker frequencies were used to verify that no structural resonances occur in either the experimental fixturing or the crystal package. No structural resonances are observed from 30 Hz to 200 Hz in our fixturing.

When measuring small gamma vector components, one can see some variation with shaker frequency that we attribute to standing wave acoustic resonance in the cables. Figure 6 shows results for a crystal with a low Q (in air), where the cable modulation effects are larger than the acceleration effect in the crystal. Note that one would measure an erroneously high gamma value if only one orientation of the crystal was measured (either the open circles or filled circles of Figure 6). Also, one would erroneously conclude that there was a structural resonance. In fact, the true answer (x's in Figure 6) is approximately $-6x10^{-11}/g_{rms}$ ($-4x10^{-11}/g$), as seen by using Equation 3 with the two measurements at each shaker frequency.

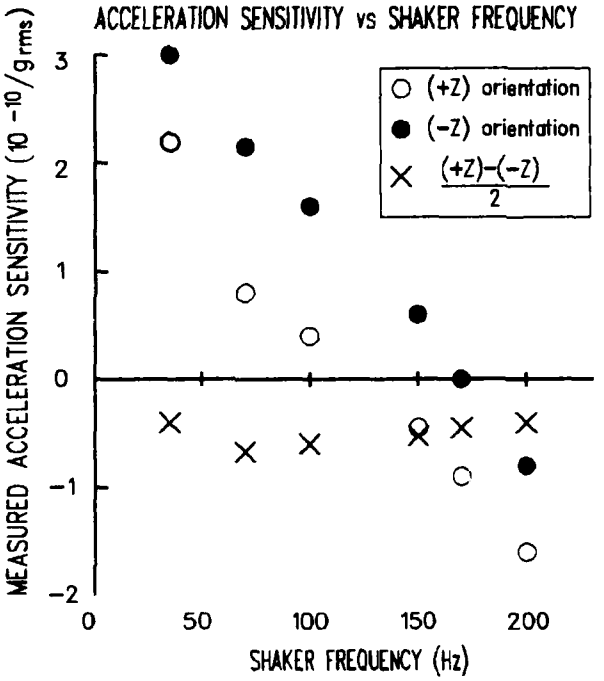



FIGURE 6. Measured acceleration sensitivity versus shaker table frequency for a 50 MHz, 4-point mount, SC-cut crystal. The effects of the cables are subtracted by measuring both + and - orientations for each axis. Note that once these cable effects are eliminated the measured acceleration sensitivity is independent of frequency.

Figure 7 shows a histogram of the gamma vector components measured for a group of 10 MHz, 3rd overtone, SC-cut, two-point mount crystals. Included in Figure 7 are theoretical results obtained using finite element modeling to calculate the stress in the quartz and the theoretical development of Tiersten[7] to calculate the frequency shift. The theoretical methods will be published later, but the comparison between the experimental average and the theory is excellent.

As a demonstration of the present experimental techniques, the magnitude of the full gamma vector is shown in Figure 8 before and after proprietary treatment of four of the resonators of Figure 7. Note that all four devices measure tightly around $2.7x10^{-10}/g_{rms}$ or ($1.9x10^{-10}/g$). This means that the technique described herein can consistently measure components on the order of $1x10^{-10}/g$.

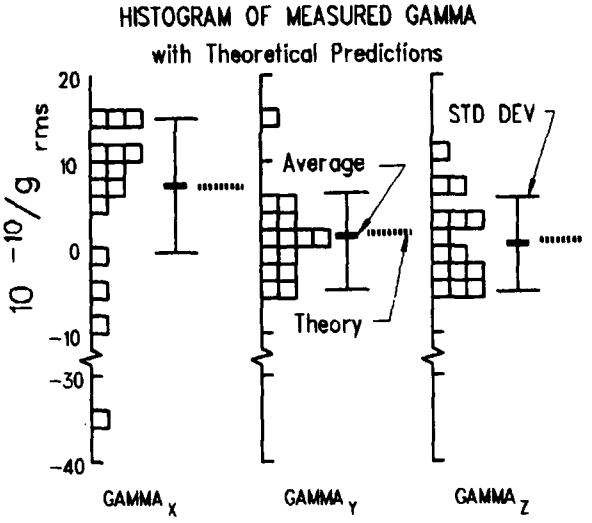



FIGURE 7. Histogram of gamma measured on each axis of 15 resonators (10 MHz, 3rd, SC-cut, 2-point mount.) Average and standard deviation of measured values are compared to analytically predicted values.

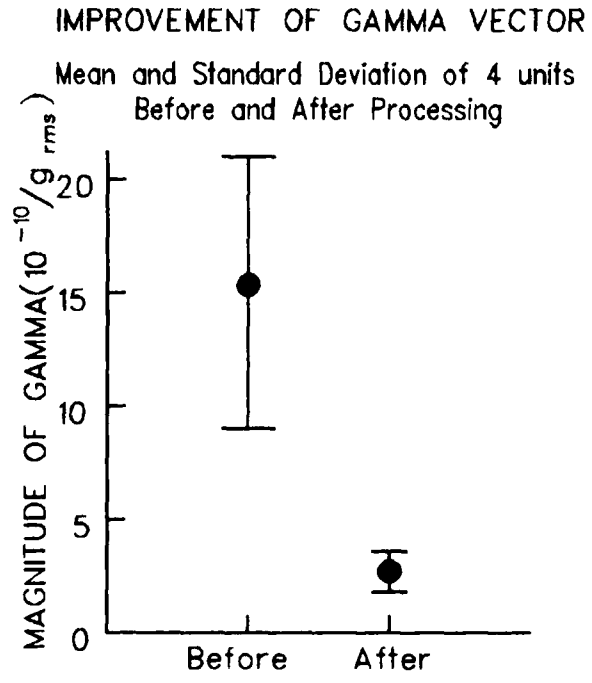



FIGURE 8. Improvement of gamma vector by proprietary processing. Mean and standard deviation of gamma vector magnitude of 4 units (10 MHz, 3rd, SC-cut, 2-point mount) before and after processing.

## CONCLUSIONS

The method described for measuring the gamma-vector on SC-cut crystals is quite practical in a research and development environment, and may be applicable to production facilities, as well. The method is able to detect both the magnitude and phase of g-sensitivities less than $1 \times 10^{-10}/g$ in crystals having a variety of frequencies and designs. The measurement procedure is rapid and uses equipment readily available in most laboratories.

## ACKNOWLEDGEMENTS

The authors wish to thank Mr. Kevin Kelly for making the measurements described herein.

## REFERENCES

[1] R.L. Filler, "The Acceleration Sensitivity of Quartz Crystal Oscillators: A Review," Proceedings 41st Annual Symposium on Frequency Control, pp. 398-408, 1987.

[2] A. Warner, B. Goldfrank, M. Meirs, and M. Rosenfeld, "Low 'g' Sensitivity Crystal Units and Their Testing," Proceedings 33rd Annual Symposium on Frequency Control, pp. 306-310, 1979.

[3] K.K. Tuladhor, et. al., "High Frequency Quartz Crystal Oscillators for Avionic Systems," Proceedings International Conference Frequency Control and Synthesis, University of Surrey, pp. 67-73, April 8-10, 1987.

[4] J. Kusters, "The Effect of Static Electric Fields on the Elastic Constants of Alpha Quartz," Proceedings 24th Annual Frequency Control Symposium, pp. 46-54, 1970.

[5] V.J. Rosatti and R.L. Filler, "Reduction of the Effects of Vibration on SC-Cut Quartz Crystal Oscillators," Proceedings 35th Annual Symposium on Frequency Control, pp. 117-121, 1981.

[6] V.J. Rosatti, "Supression of Vibration-Induced Phase Noise In Crystal Oscillators: An Update," Proceedings 41st Annual Symposium on Frequency Control, pp. 409-412, 1987.

[7] H.F. Tiersten, "Electroelastic Interactions and the Piezoelectric Equations," Journal Acoustics Society of America, 70, p. 1567, 1981.

42nd Annual Frequency Control Symposium - 1988

# Kalman Filter Analysis for Real Time Applications of Clocks and Oscillators

S. R. Stein*
Ball Communication Systems Division
P.O. Box 1235
Broomfield, CO 80020

R. L. Filler
U.S. Army Laboratory Command
Ft. Monmouth, NJ 07703-5000

## Introduction

This paper discusses the application of Kalman filtering techniques to the design of an intelligent timing system. One such system, the Modular Intelligent Frequency, Time and Time Interval (MIFTTI) Subsystem is currently under development. Our approach is oriented towards fully automated systems which perform real-time state estimation, forecasting and control. We assume that data availability is determined by factors outside of the control of the timing system. Under these conditions the Kalman filter approach has advantages compared to any other approach known to us. One of us has published several background papers which provide more information on alternative techniques, such as ARIMA modeling [1,2].General information may be found in the excellent books by Box and Jenkins [3] and Gelb [4].

The Kalman filter has many attractive features for this application; some are shared by other techniques. It is an optimum estimator in the minimum squared error sense. No other technique based on the same set of assumptions about system properties can provide better state estimates or forecasts, although other techniques may do as well. The Kalman filter applies to dynamic systems and, provided that it is properly initialized, has optimum transient response.

The particular needs of the real time control application are supported by several features of the Kalman filter. It is intrinsically robust. The filter is built on a model of the system dynamics which has the effect of reducing the deleterious effects of unreasonable data. Further, the Kalman filter provides correct internal estimates of the variances of each of the state estimates, making it easy to perform parametric outlier rejection. Kalman provided a simple recursive recipe for the computation of the filter to ease its application to any system for which one may write equations of motion.

The Kalman filter approach discussed herein was suggested for clock analysis by Dick Jones in 1984 [5]. This approach is unique in its ability to accept measurements at arbitrary intervals. Aspects of the filter have been discussed by Chaffee [6] and Bierman [7], but no one has published the details of a complete filter of this type for clock analysis. We have extended the Jones approach in several regards, adding the capability to accept both time and frequency measurements, to automatically initiate operation, and to reject outliers. This filter is fully capable of performing state estimation and forecasting for an oscillator or clock characterized by any combination of white phase noise, white frequency noise, random walk frequency noise, random walk frequency aging, and constant frequency aging.

*Work supported by U.S. Army Laboratory Command, Contract No. DAAL01-87-C-0717

In the future, it is expected that this technique will be applied to embedded time and frequency systems. It will reduce the hardware cost of meeting host system specifications by optimizing the performance of the system clock and by using all available sources of time and frequency information.

## General Procedure for Kalman State Estimation

In 1961 Kalman and Bucy published a methodology for recursive minimum squared-error state estimation. The technique has both continuous and discrete realizations, and we use an approach which is a mixture of the two. The general approach, known as a Kalman filter, is to describe the system dynamics as a set of linear ordinary difference equations and then to obtain the optimal estimate of a set of parameters that make up the "state vector." We most often see the Kalman filter equations written down for the case where the data occur at equal time intervals. This type of Kalman filter is not sufficient for many real-world situations where data may be collected at random intervals. It is possible to deal with this problem because the noise processes which we are modeling are not intrinsically discrete; rather, they result from sampling an underlying continuous time series. Thus, we use a differential equation to describe the continuous increase of the error covariance between measurements and a difference equation to describe the decrease in the error covariance which occurs at each measurement. The result is a discrete Kalman filter with equations of motion which depend explicitly on the time interval since the last measurement.

The power of the Kalman approach is that once the problem has been formulated according to the required prescription, then there is a general recursive solution to the minimum squared error estimation problem. Usually, no closed form solution will be found, but the recursion is easily programmed on a computer. The use of the actual system dynamics in the estimation process stabilizes the state estimates against occasional large measurement errors, and the Kalman filter automatically provides estimates of the errors of each component of the "state vector."

The $n$ parameters which we wish to estimate are formed into an n dimensional state vector $\vec{x}(k)$. The system evolves from the $k-1^{st}$ sample to the $k^{th}$ sample according to

$$\vec{x}(k) = \Phi(k,k-1)\vec{x}(k-1) + \vec{s}(k,k-1) \qquad (1)$$

where the $n \times n$ matrix $\Phi(k-1)$ embodies the system model and the $n$ dimensional vector $\vec{s}(k,k-1)$ contains the noise inputs to the system. Each element of $\vec{s}(k,k-1)$ is normally distributed with zero mean and uncorrelated in time. Equation 1 generates a random walk in the elements of the state vector.

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT 

The measurement vector $\vec{z}(k)$ is described by another matrix equation

$$\vec{z}(k) = \mathbf{H}(k)\vec{x}(k) + \vec{v}(k). \quad (2)$$

The $r$ measurements made at the $k$th sample time are related linearly to the $n$ elements of the state vector by the $r \times n$ dimensional measurement matrix $\mathbf{H}(k)$ and the $r$ dimensional white noise vector $\vec{v}(k)$. Kalman and Bucy defined a recursive procedure for estimating the next state, requiring that the mean squared error of the estimates from the true state is minimum.

We define the error in the estimate of the state vector after the $k^{th}$ measurement to be $\vec{x}(k|k) - \vec{x}(k-1)$. Then we define the error covariance matrix to be

$$\mathbf{P}(k|k) = E\left\{[\vec{x}(k|k) - \vec{x}(k)]\,[\vec{x}(k|k) - \vec{x}(k)]^T\right\}. \quad (3)$$

The diagonal elements of this $n \times n$ matrix are the variances of the estimates of the components of $\vec{x}(k)$ after the $k^{th}$ measurement. We also define the error covariance matrix just prior to the $k^{th}$ measurement

$$\mathbf{P}(k|k-1) = E\left\{[\vec{x}(k|k-1) - \vec{x}(k)]\,[\vec{x}(k|k-1) - \vec{x}(k)]^T\right\}. \quad (4)$$

Finally, we let $\mathbf{R}(k)$ be the covariance matrix of the measurement noise and $\mathbf{Q}(k)$ be the covariance matrix of the system noise

$$\mathbf{R}(k) = E\left[\vec{v}(k)\vec{v}(k)^T\right] \quad (5)$$

$$\mathbf{Q}(k) = E\left[\vec{s}(k)\vec{s}(k)^T\right] \quad (6)$$

The error covariance matrix evolves according to the system model

$$\mathbf{P}(k|k-1) = \mathbf{\Phi}(k-1)\mathbf{P}(k-1|k-1)\mathbf{\Phi}(k-1)^T + \mathbf{Q}(k-1). \quad (7)$$

The new estimate of the state vector depends on the previous estimate and the current measurement according to

$$\vec{x}(k|k) = \mathbf{\Phi}(k-1)\vec{x}(k-1|k-1) + \mathbf{K}(k)\,[\vec{z}(k) - \mathbf{H}(k)\mathbf{\Phi}(k-1)\vec{x}(k-1|k-1)] \quad (8)$$

where the Kalman gain matrix $\mathbf{K}(k)$ determines how heavily the new measurements are weighted. The Kalman gain is determined when we minimize the "square of the length of the error vector" that is the sum of the diagonal elements (the trace) of the error covariance matrix. This gives

$$\mathbf{K}(k) = \mathbf{P}(k|k-1)\mathbf{H}(k)^T\left[\mathbf{H}(k)\mathbf{P}(k|k-1)\mathbf{H}(k)^T + \mathbf{R}(k)\right]^{-1}. \quad (9)$$

Finally, the updated error covariance matrix is

$$\mathbf{P}(k|k) = [\mathbf{I} - \mathbf{K}(k)\mathbf{H}(k)]\,\mathbf{P}(k|k-1) \quad (10)$$

where $\mathbf{I}$ is the identity matrix. Equations 7 through 10 define the Kalman filter and so defined it is an optimal estimator in the minimum squared error sense.

## Evolution of the State Between Measurements

Here we wish to describe a clock with white phase noise, white frequency noise, random walk frequency noise and random walk or constant frequency aging. The state vector, at time $t$, contains three elements: the time, frequency and frequency aging of the clock:

$$x(t) = \begin{bmatrix} x(t) \\ y(t) \\ w(t) \end{bmatrix}. \quad (11)$$

We proceed with a model based on the assumption that there is an underlying continuous noise process and then integrate for a period equal to the next sampling interval. The continuous state equation is

$$\frac{d}{dt}\begin{bmatrix} x(t) \\ y(t) \\ w(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} x(t) \\ y(t) \\ w(t) \end{bmatrix} + \begin{bmatrix} \xi(t) \\ \mu(t) \\ \zeta(t) \end{bmatrix} \quad (12)$$

where the noise inputs, $\xi(t), \mu(t)$, and $\zeta(t)$ are continuous Gaussian 'white noise' processes with spectral densities $S_\xi, S_\mu$, and $S_\zeta$ respectively. The covariance matrix of the continuous noise processes is

$$\mathbf{Q}dt = \begin{bmatrix} S_\xi & 0 & 0 \\ 0 & S_\mu & 0 \\ 0 & 0 & S_\zeta \end{bmatrix}. \quad (13)$$

These noise inputs produce the white frequency noise, random walk frequency noise, and random walk frequency aging required by the physical model. The required white phase noise input will be accounted for later.

Integration of the continuous model for a time $\delta$ yields the state transition matrix and the error covariance matrix

$$\mathbf{\Phi}(\delta) = \begin{bmatrix} 1 & \delta & \delta^2/2 \\ 0 & 1 & \delta \\ 0 & 0 & 1 \end{bmatrix} \quad (14)$$

$$\mathbf{Q}(\delta) = \begin{bmatrix} S_\xi\delta + S_\mu\delta^3/3 + S_\zeta\delta^5/20 & S_\mu\delta^2/2 + S_\zeta\delta^4/8 & S_\zeta\delta^3/6 \\ S_\mu\delta^2/2 + S_\zeta\delta^4/8 & S_\mu\delta + S_\zeta\delta^3/3 & S_\zeta\delta^2/2 \\ S_\zeta\delta^3/6 & S_\zeta\delta^2/2 & S_\zeta\delta \end{bmatrix} \quad (15)$$

The discrete measurement model was written in matrix form in Equation 2. The filter may utilize several simultaneous time and/or frequency measurements. Although it would be possible to include all simultaneous measurements within the measurement matrix, we choose to process them serially. The final result is independent of the order of the computations. The phase measurements of the clock relative to the reference are described by

$$\mathbf{H}(k) = (\,1 \;\; 0 \;\; 0\,) \text{ and } \mathbf{R} = \sigma_{x_r}^2 \quad (16)$$

where the variance of the phase measurement process contains two contributions: the white phase noise of the oscillator and the white phase noise of the measurement equipment. The frequency measurements are described by

$$\mathbf{H}(k) = (\,0 \;\; 1 \;\; 0\,) \text{ and } \mathbf{R} = \sigma_{y_r}^2. \quad (17)$$

Each application of the Kalman recursion yields estimates of the time $x(t+\delta)$, frequency $y(t+\delta)$, and frequency aging $w(t+\delta)$ of the clock a time $\delta$ after the last measurement. The interval between measurements, $\delta$ varies from one measurement to the next.

## Values of the Parameters

The noise that directly perturbs the phase of the clocks results in a white frequency noise spectrum. Thus we may write the spectral density of this noise in terms of the standard oscillator noise coefficients:

$$S_\xi = h_0. \quad (18)$$

The noise that directly perturbs the frequency results in a random walk frequency spectrum. Thus we may write the spectral density of this noise:

$$S_\mu = (2\pi)^2 h_{-2}. \quad (19)$$

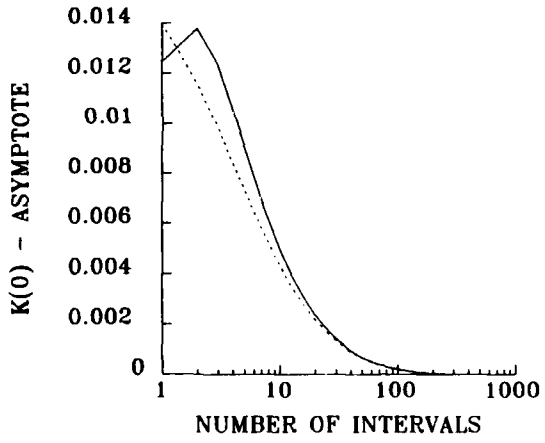



Figure 1: Kalman phase gains for different estimates of the initial covariance matrix.

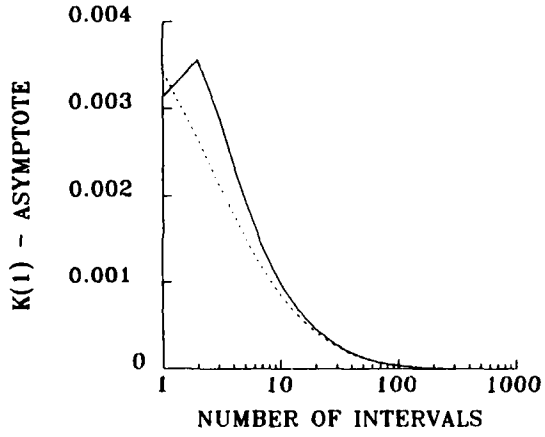



Figure 2: Kalman frequency gains for different estimates of the initial covariance matrix.

The noise that directly perturbs the frequency aging results in a spectral density of frequency which varies with Fourier frequency as $f^{-4}$. There is no standard nomenclature for the coefficient of this noise type.

# Startup Procedure

The most difficult startup conditions occur during cold start, that is when there is no prior knowledge of the time, frequency or frequency aging of the clock. This startup mode is required for a fully automated system. The startup procedure consists of determining initial values for the state vector and the covariance matrix.

## Initializing the State Vector

Since three states must be initialized, the procedure requires three measurements. One time measurement is needed to complete the initialization. If only frequency measurements are available, Kalman filter operation is initiated after two measurements and the time initialization is completed when the first time measurement is made.

Here we describe the case of cold start after three phase measurements to illustrate the method. Other measurement sequences may be treated in a similar fashion. A reasonable way

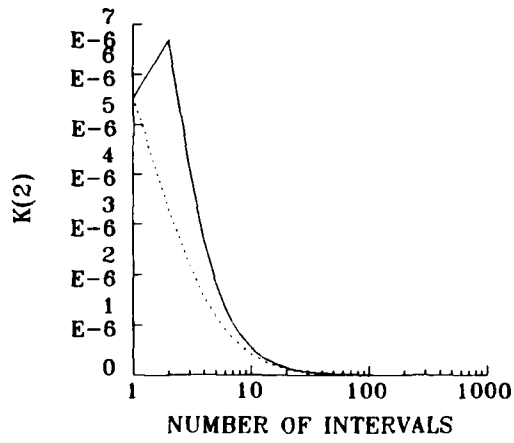



Figure 3: Kalman frequency aging gains for different estimates of the initial covariance matrix.

to start is to require that the time, frequency, and frequency aging just after the third measurement define a quadratic passing through the three measurements. After some algebra we obtain the state estimates

$$\begin{aligned} x(2|2) &= z(2) \\ y(2|2) &= \frac{\delta_1+\delta_2}{\delta_1}\left[\frac{z(2)-z(1)}{\delta_2}\right] - \frac{\delta_2}{\delta_1}\left[\frac{z(2)-z(0)}{\delta_1+\delta_2}\right] \\ w(2|2) &= \frac{2}{\delta_1+\delta_2}\left[\frac{z(2)-z(1)}{\delta_2} - \frac{z(1)-z(0)}{\delta_1}\right] \end{aligned} \tag{20}$$

where $\delta_1$ is the time interval between measurements 0 and 1 and $\delta_2$ is the time interval between measurements 1 and 2. All variable enumeration is done from base zero.

## Initializing the Covariance Matrix Elements

We calculate the covariance matrix, $\mathbf{P}^+(k)$ $\mathbf{P}(k|k)$ from the definition of Equation 3. In the case of startup after three phase measurements, we find

$$\begin{aligned} P_{00}^+(2) &= \sigma_{v_\epsilon}^2, \\ P_{11}^+(2) &= \frac{2\delta_1^4+8\delta_1^3\delta_2+10\delta_1^2\delta_2^2+4\delta_1\delta_2^3+2\delta_2^4}{\delta_1^2\delta_2^2(\delta_1+\delta_2)^2}\sigma_{v_\epsilon}^2 \\ &\quad + \frac{\delta_1^3+4\delta_1^2\delta_2+4\delta_1\delta_2^2+\delta_2^3}{\delta_1\delta_2(\delta_1+\delta_2)^2}S_\xi \\ &\quad + \frac{\delta_2}{3}S_\mu + \frac{\delta_2^2(3\delta_1^3+8\delta_1^2\delta_2+7\delta_1\delta_2^2+2\delta_2^3)}{60(\delta_1+\delta_2)^2}S_\zeta, \\ P_{22}^+(2) &= \frac{8(\delta_1^2+\delta_1\delta_2+\delta_2^2)}{\delta_1^2\delta_2^2(\delta_1+\delta_2)^2}\sigma_{v_\epsilon}^2 + \frac{1}{\delta_1\delta_2(\delta_1+\delta_2)}S_\xi \\ &= + \frac{4}{3(\delta_1+\delta_2)}S_\mu + \frac{3\delta_1^3+15\delta_1^2\delta_2+20\delta_1\delta_2^2+8\delta_2^3}{15(\delta_1+\delta_2)^2}S_\zeta, \\ P_{01}^+(2) &= \frac{\delta_1+2\delta_2}{\delta_2(\delta_1+\delta_2)}\sigma_{v_\epsilon}^2, \\ P_{02}^+(2) &= \frac{2}{\delta_1(\delta_1+\delta_2)}\sigma_{v_\epsilon}^2, \quad \text{and} \\ P_{12}^+(2) &= \frac{2(2\delta_1^3+5\delta_1^2\delta_2+3\delta_1\delta_2^2+2\delta_2^3)}{\delta_1^2\delta_2^2(\delta_1+\delta_2)^2}\sigma_{v_\epsilon}^2 \\ &\quad + \frac{2}{\delta_1\delta_2}S_\xi + \frac{\delta_2}{3(\delta_1+\delta_2)}S_\mu \\ &\quad + \frac{\delta_2(6\delta_1^3+20\delta_1^2\delta_2+21\delta_1\delta_2^2+7\delta_2^3)}{60(\delta_1+\delta_2)^2}S_\zeta. \end{aligned} \tag{21}$$

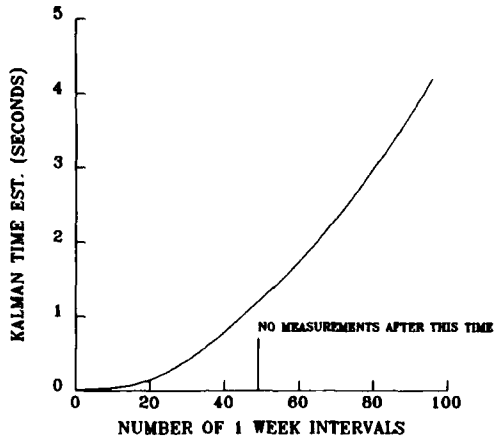



Figure 4: Kalman time estimate of a quartz oscillator.

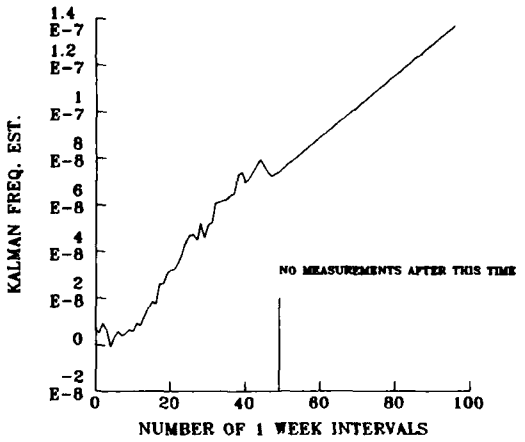



Figure 5: Kalman frequency estimate of a quartz oscillator.

The results of using the calculated startup values of the covariance matrix are shown in Figures 1, 2, and 3 where we plot the Kalman gains for the three elements of the state vector for the case of a time series of phase measurements. The dotted lines show the results of using the exact startup values of the covariance matrix. The solid lines show the results that were obtained with approximate values of the initial covariance matrix elements. We see that the use of calculated consistent values of all the covariance matrix elements results in Kalman gains that are smoother and smaller than the gains resulting from the approximated initial covariance. The filter that uses the calculated values yields improved state estimates.

We cannot stress too much that the startup procedure is arbitrary to the extent that it employs requirements not part of the Kalman methodology. The procedure that we have adopted functions well when there is no prior knowledge of the state of the oscillator, and is easily modified to deal with the case when the startup measurements yield state estimates small compared to the estimated standard deviations of those measurements. Although the startup procedure has to be adapted to the requirements of the application, it is always valuable to calculate the initial covariances if good transient performance is desired.

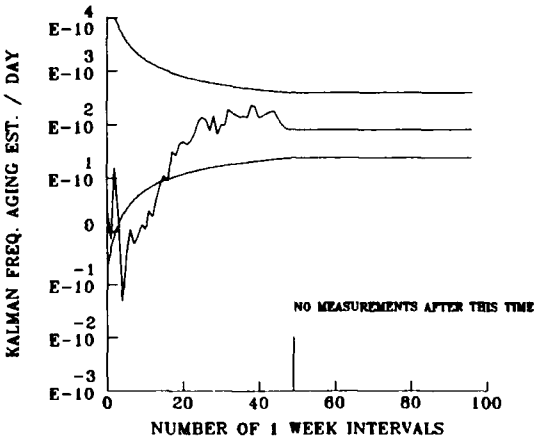



Figure 6: Kalman frequency aging estimate of a quartz oscillator.

## Results of a Simulation Experiment

A simulation experiment was performed to begin the process of evaluating the continuous Kalman filter. We simulated a quartz clock which was required to keep time for a two-week interval following calibration and initial synchronization. We compared two different calibration procedures:

1. Frequency calibration was obtained from a Rb oscillator which itself had been perfectly calibrated one year earlier.

2. Frequency and frequency aging calibration were obtained over a one-year period, using weekly frequency and time measurements relative to a GPS receiver.

The quartz oscillator had specifications similar to the Hewlett Packard 10811A. The frequency aging was $2 \times 10^{-10}$ per day and the frequency stability is shown in Table 1. Figures 4, 5, and 6 show the time, frequency, and frequency aging estimates of the quartz oscillator. The first year is the calibration interval. After calibration, the Kalman filter receives no further information,

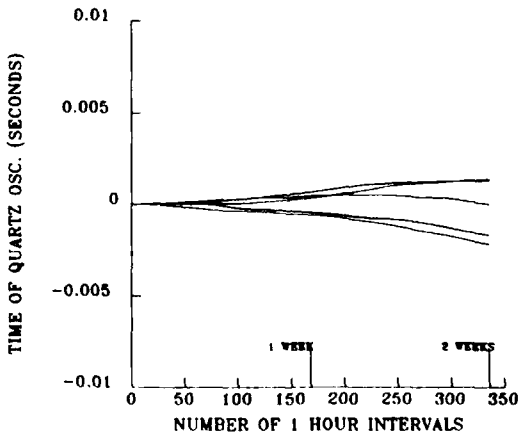



Figure 7: Time error of a quartz oscillator during the two weeks following calibration of the time, frequency and frequency aging using a Kalman filter.

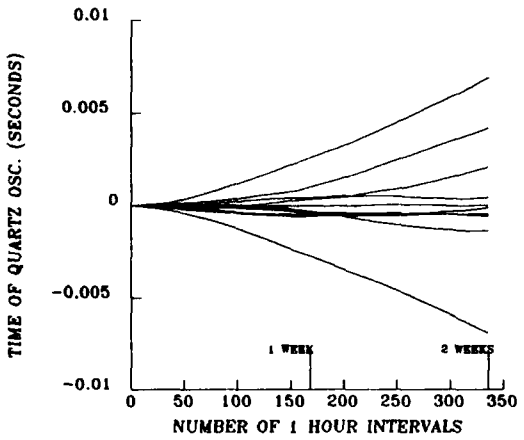



Figure 8: Time error of a quartz oscillator during the two weeks following calibration of the frequency with a rubidium oscillator.

Table 1: Quartz oscillator short term stability

| $\tau$ | $\sigma_y(\tau)$ |
|---|---|
| .01 s | $1.5 \times 10^{-11}$ |
| 1 s | $2.2 \times 10^{-12}$ |
| 10 s | $5.1 \times 10^{-12}$ |
| 100 s | $1.6 \times 10^{-11}$ |
| 86400 s | $4.7 \times 10^{-10}$ |

but continues to update the state estimates. The frequency aging estimate remains at its last value; the frequency estimate is propagated linearly in time using the final frequency aging estimate for the slope; and the time is propagated forward in time quadratically according to the final frequency aging estimate.

Figures 7 and 8 show several simulations of the time error of the quartz oscillator for two weeks following calibration. Figure 7 is with and Figure 8 without the Kalman filter. The Kalman filter reduces the effect of aging so the accumulated time error is dominated by the random walk of frequency. The simulations in Figure 7 show a worst case error of $\pm 2$ ms. Figure 8 shows a total timing error due to random walk of frequency, frequency syntonization error, and frequency aging of $\pm 6$ ms. This improvement will increase quadratically with time. For example, after 4 weeks the time error would be reduced by a factor of 12.

## Application to Phase and Frequency Lock Loops

Figure 9 illustrates a fairly general control loop for a disciplined oscillator using the information from the Kalman filter described above. This configuration can function as a frequency lock loop or a first or second order phase lock loop. The estimates or forecasts of the errors are used to control the VCO frequency in the traditional way.

In this example, the measurement noise is assumed to be negligible, so the Kalman estimate of the phase difference between the slave oscillator and the reference is equal to the measured phase. If this signal is fed back with some gain while no frequency information is utilized, then a first order phase lock loop results. On the other hand, if the gain of the phase feedback is zero and the frequency difference state is fed back with unit gain, then a frequency lock loop results. If some phase feedback is added to the frequency correction, then a second order phase lock loop results. The detailed performance analysis of this control loop has been published previously [8].

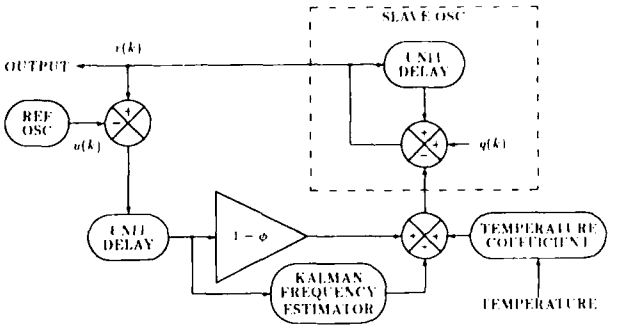



Figure 9: Phase/frequency lock loop with feedforward

## Conclusions

An "optimum" algorithm has been developed for estimating the time, frequency, and frequency aging of clocks and oscillators. The algorithm is sufficiently general to be used with all types of quartz oscillators and atomic clocks yet the approach is appropriate for use with an embedded controller in automated systems intended for field applications. The algorithm is based on Kalman filtering techniques and is able to process either time or frequency calibration data from various sources of different qualities. Data are accepted whenever available and are not required at fixed multiples of a fixed sampling time.

This algorithm may be used to analyze data already available in many systems in order to improve system timing with no additional hardware. Alternatively, it may be possible to maintain system timing performance while reducing the cost of the clocks. These goals are an important part of the Modular Intelligent Frequency, Time and Time Interval Subsystem currently under development.

## References

[1] A Disciplined Rubidium Oscillator, A. MacIntyre and S. R. Stein, Proc. 40th Annual Symposium on Frequency Control, 1986, p. 465-469.

[2] Control of the Frequency and Phase of an Atomic Clock by an External Reference, S. R. Stein, A. MacIntyre and G. Huebner, International Conference on Frequency Control and Synthesis, 1987.

[3] Box, G.E.P. and Jenkins, G.M. "Time Series Analysis, Forecasting and Control." Holden Day, San Francisco, 1970.

[4] Gelb, A. ed. "Applied Optimal Estimation." M.I.T. Press,Cambridge, 1974.

[5] Continuous Time Series Models for Unequally Spaced Data Applied to Modeling Atomic Clocks, R.H. Jones and P.V. Tryon, SIAM J. Sci. Stat. Comput., vol. 8, 1987.

[6] Chaffee, J., OAO Corporation Internal Memorandum, 1982.

[7] Private Communication, 1986.

[8] Application of Kalman Filters and ARIMA Models to Digital Frequency and Phsae Lock Loops, J.A. Barnes and S.R. Stein, Proc. of the 19th Annual Precise Time and Time Interval Applications and Planning Meeting, 1987.

42nd Annual Frequency Control Symposium - 1988

# RADAR MEASUREMENT APPLICATIONS OF FIBER OPTIC LINKS

I.L. Newberg, C.M. Gee, G.D. Thurmond, and H.W. Yen

Hughes Research Laboratories, Malibu, CA 90265

## Summary

Novel applications of fiber optic delay lines to improve the capability and performance of radar repeater and phase noise test sets are described. Using a 2.25-km-long fiber optic link, we demonstrated generation of an ideal target for a radar repeater test set, and measurement of phase noise closer than 100 Hz to a 9.6 GHz radar carrier signal using a delay line discriminator phase noise test set. FM and AM signal-to-noise measurements were made to determine the performance capability of the 10 GHz modulated fiber optic links for the radar applications.

## Introduction

Fiber optic links are potentially important for future radar and avionic systems. Fiber optics offer many advantages for microwave signal transmission, including wide bandwidth, low loss, light weight, small size, and EMI resistance. Furthermore, fiber optics can provide delays much longer than are practical with coaxial cable or waveguide at X-band (10 GHz) frequencies. In this paper, we describe how analog fiber optic links used as delay lines can dramatically increase the capability and flexibility of radar repeater and phase noise test sets. We present FM and AM signal-to-noise ratio (SNR) measurements to determine the performance capability of the links for these applications. This is the first known implementation of fiber optics for a radar system at X-band frequencies. The test set applications demonstrate that fiber optics can meet the stringent SNR requirements of radar systems.

## Phase Noise Test Set

In Doppler radar applications, the velocity of a target is determined by small frequency shifts of the return signal compared with the transmitted signal. The target return signal must be distinguishable from environmental clutter and from the radar equipment noise. Phase noise is random noise arising from thermal, shot, and flicker noise from rf signal generators. For Doppler radar application, the radar exciter phase noise must be well-characterized and very small so it does not mask the desired return target signal. This requirement is especially important in airborne radar applications, where the ground return generates a large frequency-shifted echo.

The absolute phase noise of an rf exciter in a Doppler radar can be measured with a phase noise test set. One type of phase noise test set uses a delay line frequency discriminator technique, and is illustrated in Figure 1. The delay line forms a discriminator that decorrelates the phase noise of the rf signal source and converts the noise to a baseband signal amplitude that can be measured on a spectrum analyzer. Unlike the phase-detector method that uses a low noise rf source to mix with the device under test (DUT), the discriminator technique needs no reference source. Thus, this method is simple and especially useful for measuring noise of sources that are difficult to phase lock.

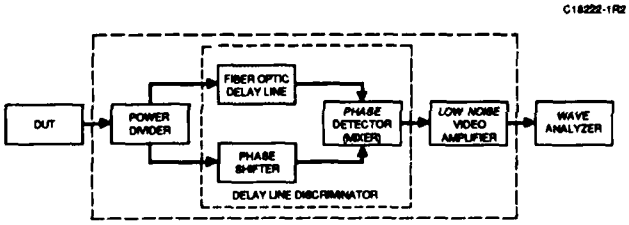



Figure 1. Delay line frequency discriminator technique for measuring phase noise of the device under test (DUT). Fiber optics provides a long delay line.

A delay line type of phase noise test set works as follows.[1] The reference frequency arrives at the double-balanced mixer at a particular phase. As the frequency of the DUT varies slightly, the phase shift incurred in the fixed delay time will change proportionally. If one sets the phase shifter so the two input signals to the mixer are 90° out of phase (phase quadrature), the mixer becomes a phase detector and the output voltage is proportional to the input phase fluctuations. The voltage fluctuations can then be measured with a spectrum analyzer and converted to phase-noise units relative to the signal carrier in a given bandwidth (dBc/Hz). A known level and frequency of FM sideband energy can be added to the test set signal input for calibration.

The sensitivity of the noise test set in measuring noise close to the carrier directly improves with longer delay lines. This enhancement is important for Doppler radar applications where the target return signal is close to the carrier signal. The sensitivity of the phase noise test set is given by the following equation:

$$\Delta V(f_m) = \kappa 2\pi\tau\,[\sin(\pi f_m \tau)/\pi f_m \tau]\,\Delta f(f_m) \quad , \qquad (1)$$

where $\Delta V$ is the voltage fluctuation, $\Delta f$ is the noise frequency fluctuation, $\kappa$ is the phase to voltage translation constant, $\tau$ is the delay provided by the delay line, and $f_m$ is the frequency offset from the carrier at which the phase noise is being measured. Eq. (1) shows that the sensitivity improves with large delay $\tau$. However, large $\tau$ also decreases the $f_m$ at which measurements can be made because the system sensitivity becomes poor near the nulls of Eq. (1). Measurements are generally made at $f_m$ much less than $1/\tau$. Then Eq. (1) simplifies to:

$$\Delta V(f_m) = \kappa 2\pi\tau\Delta f\,(f_m) \quad , \qquad (2)$$

where the output voltage is directly proportional to the delay $\tau$. The length of the delay line is selected to maximize the test set sensitivity for a specific frequency range. Long delay lines are used for close-in noise, while shorter delay lines are for measuring noise at frequency offsets far from the carrier.

Waveguide delay lines presently used in these test sets are limited to ≡0.25 μs because of their size, weight, and loss at X-band frequencies. Therefore, the test set sensitivity is reduced considerably for phase noise measurements closer than 10 kHz from the carrier signal. On the other hand, using a 2.25-km-long fiber optic link as a delay line, we have demonstrated a test set with 11 μs delay that is sensitive enough to measure noise closer than 100 Hz to the rf carrier. This is a significant capability for determining the performance of Doppler radars.

## Radar Repeater Test Set

Another important and novel application of high-speed fiber optic links is in a radar repeater test set, shown colocated with a radar system in Figure 2. The radar system includes an exciter unit that generates a low noise coherent cw rf signal that is amplified in the transmitter and pulsed at a pulse repetition frequency (PRF). The duplexer directs the transmitted energy out to the antenna and the received energy into the receiver while providing isolation. During the transmit phase, the receiver is blanked to prevent damage to the receiver by the transmitter output.

Current repeater test targets are typically located several miles from the radar to achieve sufficient delay for the radar to receive its transmitted pulse. The delay in this case is obtained by the time required for the radar transmitted pulse to reach the repeater and be retransmitted as a target return to the radar. Because it is transmitted through the air, the return radar pulse must compete with environmental noise and clutter. In addition, covert radars that

CH2588-2/88/0000-453 $1.00 C 1988 IEEE

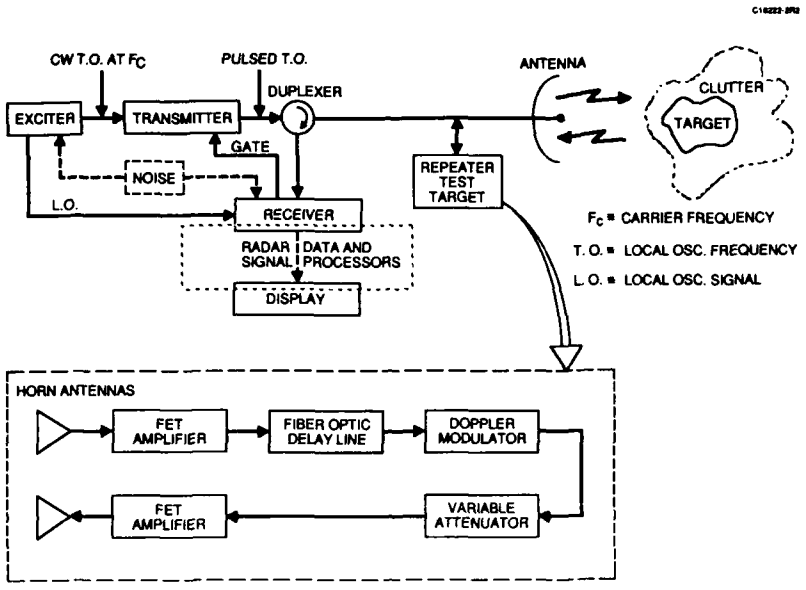



Figure 2. A radar system with a colocated repeater test set to provide the delay generally achieved with a test target located many miles from the radar. A fiber optic delay line is incorporated in the repeater.

cannot transmit outside a closed laboratory cannot be characterized using this type of radar repeater.

Figure 2 illustrates how fiber optics can replace a test target located many miles away from the radar transmission site. In the radar repeater application, a long fiber optic delay line provides a delayed coherent replica of the input signal, allowing a radar to receive and process its own transmitted pulsed rf signal as if it were an actual target return. The long fiber optic delay enables a method of complete end-to-end radar system evaluation within a single test facility. Thus, an ideal test target signal is provided, without environmental effects. This facility permits determination of radar performance parameters under the influence of all added noise and spurious signals caused by the radar exciter, transmitter, waveform modulation, and signal processing.

We have demonstrated this concept with a 4-km-long fiber optic link. A pulsed rf X-band signal was transmitted through a fiber optic link, coupled back into the radar system with negligible degradation, and then displayed as a delayed target signal.

## Fiber Optic Link Characterization

For these applications, we have characterized close-in AM and FM noise for both externally and directly modulated fiber optic links. These links are illustrated in Figure 3. The external modulation link transmitter includes a 1.3 μm cw InGaAsP laser, a GRIN lens, a series of two optical isolators, and a traveling-wave $LiNbO_3$ electro-optic modulator.[2] Incoming rf signals modulate the external electro-optic device. The direct modulation link transmitter is an rf-modulated high-speed semiconductor laser. The receiver, consisting of a high-speed InGaAs photodetector, is linked to the transmitter by single-mode optical fiber, which also provides the signal delay. The direct modulation link is simple and more compact, but the external modulation technique offers better broadband performance. For both methods, incoming rf signals to the link amplitude-modulate the optical carrier emitted by the laser. The modulated light travels along the fiber and is detected at the other end with a high-speed photodetector. The photodetector demodulates the optical signal to produce a microwave output that coherently replicates the rf input to the link. This device detects the AM modulation in a manner similar to a microwave diode detector, in that it responds to the amplitude only and not to the phase of the light.

The relative SNR corruption that the fiber optic link adds to the input rf carrier determines whether or not fiber optics would degrade the radar test set performance. The AM and FM relative SNRs of the link were measured and Figure 4 summarizes the performance.

C17720-2R1

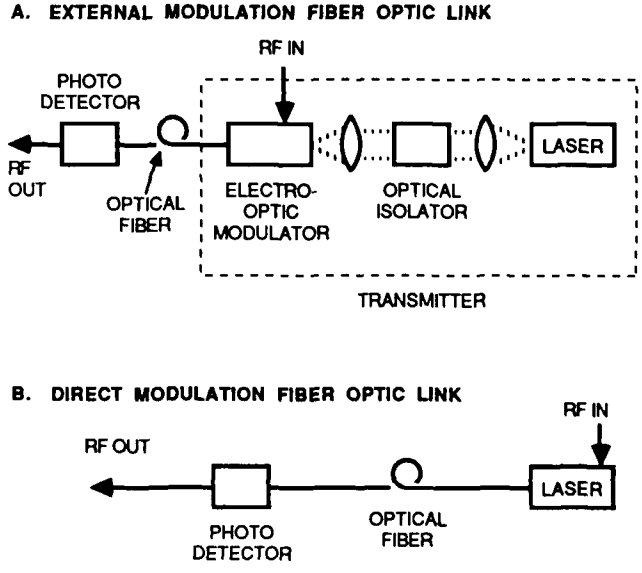



Figure 3. External modulation and direct modulation fiber optic links.

| Unit Under Test | FM Noise | AM Noise |
|---|---|---|
| Good radar signal source | -125 dBc/Hz | -145 dBc/Hz |
| (4m)* External mod FO link | -138 dBc/Hz | -138 dBc/Hz |
| (100 m) External mod FO link | -135 dBc/Hz | -135 dBc/Hz |
| (2.25 km) External mod FO link | -130 dBc/Hz | -133 dBc/Hz |
| (100 m) Direct mod FO link | -130 dBc/Hz | -129 dBc/Hz |

FM and AM noise levels at 10 kHz offset frequency in a 1 Hz bandwidth relative to the 9.6 GHz rf carrier signal

*( ) Length of single mode fiber between FO 1.3 μm wavelength transmitter and receiver

Figure 4. Performance of various fiber optic links for radar applications. The noise measurements were made at 10 kHz offset frequency from a 9.6 GHz carrier signal.

The best SNR at 10 kHz offset frequency from a 9.6 GHz carrier was -138 dBc/Hz, achieved with a 4-m-long external modulation fiber optic link. In that case, the link output noise level was about equal to the link photodetector and output first stage rf amplifier noise. The 100-m direct modulation link noise level was higher, with an SNR of -130 dBc/Hz. This SNR is poorer because the frequency response of the laser is lower than that of the external modulator, and the noise level of the laser is higher under direct laser modulation. It was readily apparent that reflections back to the laser increased the noise level and caused instability. The SNR of the 2.25 km fiber optic link was lower and less stable. Degraded performance could be due to increased fiber loss, reflections from Rayleigh scattering in the fiber, increased noise level because of fiber dispersion, and perturbations over the long fiber length. When reflections are low, the link is stable and the SNR for the fiber optic links meets the -130 dBc/Hz needed for most radar test set applications.

## Conclusions

We have demonstrated that fiber optics can improve the capability of radar test set systems, providing delays longer than are practical with coaxial cables or waveguides. This capability will allow performance measurements to be made on modern Doppler radars. The results show that fiber optic links will meet the SNR requirements of modern radars, and will potentially have many radar and electronic warfare applications in the near future.

## Acknowledgment

The authors acknowledge the support and useful discussions of M. Radant, L. Seeberger, and G. Wagner of the Hughes Radar Systems Group (RSG), and A. Popa of the Hughes Research Laboratories. We also acknowledge Ortel Corporation for providing a prototype fiber optic link.

## References

1. "Phase Noise and its effect on microwave systems," MSN Microwave Syst. News & Commun. Technol. 16, No. 7, 433-447 (July 1986).

2. C.M. Gee, G.D. Thurmond, and H.W Yen, "17-GHz bandwidth electro-optic modulator," Appl. Phys. Lett. 43, 998-1000 (Dec. 1983).

42nd Annual Frequency Control Symposium - 1988

# PRECISE MEASUREMENTS OF QUARTZ CRYSTAL UNITS BY NETWORK ANALYZER TECHNIQUE APPLIED TC TWO DIFFERENT TYPES OF TEST JIG

Vera P. Popović-Milovanović, Bojan J. Dobnikar and Vojislav Lj. Popović

Institute "Mihajlo Pupin", Volgina 15, 11000 Belgrade, Yugoslavia

## Summary

The authors investigated the accuracy and repeatability of the quartz crystal parameter measurement method using network analyzer techniques applied to two different types of test jigs: coaxial and Pi-network, under controlled temperature conditions.

The accuracy and repeatability as a function of frequency, for the mentioned types of test jigs, is presented. Results will be very useful for the activities in the IEC TC 49.

## Introduction

The measurement of frequency and dynamic parameters is a process closely tied to crystal unit development and production. Since the required tolerances are becoming narrower and the demand for high precision, as well as high frequency units is rising, the development of measurement methods is receiving considerable attention.

The crystal unit parameter measurement system, i.e. the measurement method employed must ensure:

- high accuracy and repeatability of measurement results
- satisfaction of the crystal unit manufacturer's needs as regards measurement speed, flexibility of the system (handling various types of crystal units), simple physical connection of the load capacitance $C_L$ and especially measurement automatisation, enabling operation by less experienced personnel.

The high accuracy and repeatibility of the final results is obtained by satisfying the basic metrological principles:

a) High repeatibility of measurement results is achieved by employing instruments with a stable signal source, connecting instruments to a frequency standard, using high resolution instruments, introduction of a settling time before system response measurement and controlling ambiental and working conditions of the crystal unit (temperature, rate of temerature change and drive level).

b) Elimination of the systematic error.

c) Estimation of the random error.

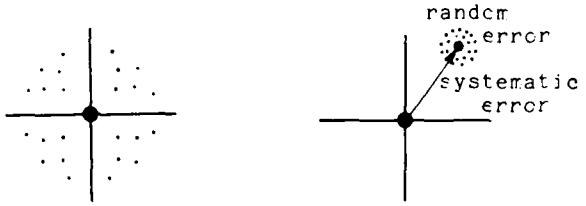



Fig.1a. Accurate but not precise

Fig.1b. Precise but not accurate

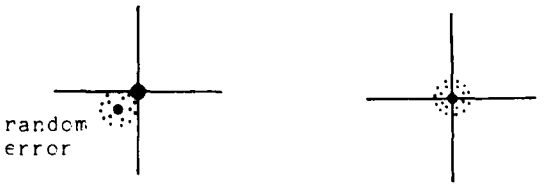



Fig.1c. Precise and accurate

Fig.1d. Theoretical limit:best accuracy by random error estimation

Fig. 1. Possible distributions of measurement results and target - hitting analogy.

Assuming that the measurement result distribution about the true value, when neither of the aforementioned principles is applied, takes the form shown in Fig. 1a, application of principle "a" gives a distribution having high measurement result precision and considerable systematic error ( Fig. 1b ), and application of principle "b" and application of statistical methods and fitting procedure techniques causes a distribution having high measurement result precision and practically negligible systematic error (Fig. 1c).

Numerous methods and measurement systems have been developed and realized, achieving the

CH2588-2/88/0000-456 $1.00 C 1988 IEEE

aforementioned requirements more or less successfully.

The reference method of crystal unit parameter measurement being applicable to all frequencies and satisfying the basic metrological principles has an advantage over other methods. On the other hand, having in mind that now and in the near future, due to practical reasons, a large portion of the units produced can be measured by using the IEC 444 method, the fact that the Pi-network fits various types of crystal holders, that connection of the load capacitance $C_L$ is simple and so is measurement of temperature characteristics, the efforts to keep the IEC Pi-network as a test jig are justified.

## The reference method of quartz unit parameter measurement using network analyzer techniques

The scattering parameter technique was developed for the measurement of devices for microwave application, but became available in everyday use with the new generation of modern HP network analyzers, allowing separation of the incident, reflected and transmitted signals.

The most important fact is that the network analyzer technique allows separation of the measurement system up to the reference plane, from the crystal unit under test. Existing systematic imperfections of the measurement system are modelable and can be subtracted from the measurement of a crystal unit and thus "corrected" scattering parameter values can be obtained.

Error model parameters are determined from the s-parameter measurement, when known impedances such as a short circuit, an open circuit, a 50 ohm matched termination for one-port (3-term error model), and in addition a thru-line connection for two - port measurements (12 term error model) are connected to the reference plane.

To avoid the connection of calibration standard impedancies each time a measurement is to be made, the error model parameters being all a function of frequency are measured at a selected number of points. For frequencies between these points the cubic interpolation technique is applied, and a set of third-order polynomial function coefficients for each single interval is calculated. Knowing the cubic spline interpolation coefficients one can calculate all the error parameters at any frequency within the frequency range.

Knowing the values of error parameters the corrected S-parameter values are obtained, and using the familiar relationship between S and Y parameters, the true $Y_{21}$ transadmittance for the corresponding frequencies are calculated. The data reduction technique, proposed by EIA-512 can then be employed.

It is well known that the $Y_{21}$ transadmittance function is a parametric equation of a circle in the admittance plane with a radius of $(1/2\ R_1)$, and whose center coordinates have values of $(1/2\ R_1)+Gc;Bo$ (they can be seen in fig. 2):

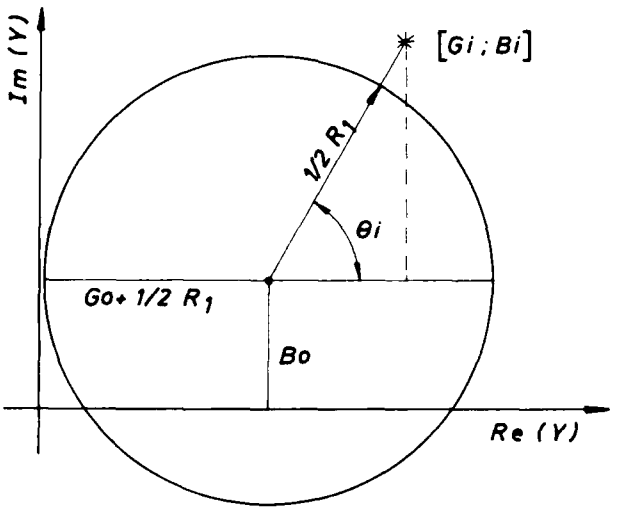



Fig. 2. Locus of a crystal unit transadmittance function in the Y-plane.

$$(G-(1/2\ R_1)-Gc)^2+(B-Bo)^2=(1/2\ R_1)^2 \quad (1)$$

The measured $Y_{21}$ (f) values are fitted to a circle. The points are then translated along the radii to lie on the fit circle, and finally the mathematical removal of Co influence is achieved by subtraction of Bo from the susceptance component of the $Y_{21}$ (f).

The fitted smoothed and shifted data are then converted to impedance.

If we denote Qi (f) as the angular position of the measured point as viewed from the center of the circle, then:

$$Qi(f) = \arctan\left(\frac{Bi(f)-Bo}{Gi(f)-(1/2\ R_1)-Gc}\right) \quad (2)$$

and the reactance function is:

$$Xi(f) = -Bi(f)/(Gi(f)^2+Bi(f)^2) \quad (3)$$

where Gi(f) and Bi(f) are expressed as:

$$Gi(f)=(Gc+1/2\ R_1)+(1/2\ R_1)\cdot\cos(Qi(f)) \quad (4)$$

$$Bi(f) = (1/2\ R_1)\cdot\sin(Qi(f)) \quad (5)$$

Applying the fitting procedure to a third order polynomial:

$$dF(Xi)=Cc+C_1Xi+C_2Xi^2+C_3Xi^3 \qquad (6)$$

one can calculate the coefficients Cc, $C_1$, $C_2$ and $C_3$. fs is calculated as the value for which the reactance is zero (Xi=0) and $L_1$ using the expression for the slope of X(f) at the point fs is:

$$L_1=(1/4\cdot\pi^2)\left.\frac{dX}{df}\right|_{f=fs} \qquad (7)$$

$$C_1=1/(4\cdot\pi^2\cdot fs^2\cdot L_1) \qquad (8)$$

The admittance points should be measured within the 3dB bandwidth of the resonance (the aforementioned relations are valid for the narrow-band aproximation), which corresponds to the right half of the admittance circle. This can be accomplished by examining the real part Gi(f) of the admittance and taking into account the range at whose boundaries the Gi(f) drops to half of its maximum value. If this condition is not satisfied, then, depending at which boundary the required condition is not fulfilled, the lower or upper limit if the frequency range is expanded and new measurements at the new points are performed. The software iteratively searches for the required measurement range so the measurement is completely operator independent.

The reference method requires that the test jig must provide a nominally constant impedance 50 Ω connection from the network analyzer to the reference plane at which the crystal unit is to be connected.

## Methods of quartz crystal unit parameter measurement using a network analyzer and a Pi-network as a test jig

In recent year several methods, basically similar to the reference method, but using a Pi-network as a test jig have been developed.

The idea in common with the reference method is in calculating the quartz crystal unit parameters from the $Y_{21}$ transadmittance. The transadmittance values are obtained from the voltage transfer function of the Pi-network. The Pi-network shown in figure 3. can be depicted by the equivalent circuit shown in figure 4.

The general relation for the voltage Vb at the output of the Pi-network is:

$$Vb = \frac{Z_2}{Z_1+Z_2+Zm}\cdot kVa \qquad (9)$$

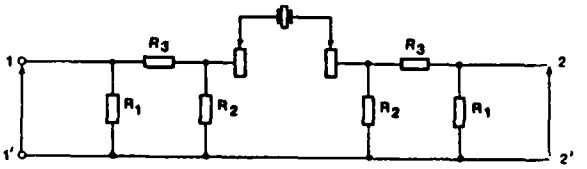



Fig.3. Crystal adapter in accordance with IEC 444. ($R_1$=159 , $R_2$=14.2 , $R_3$=66.2 )

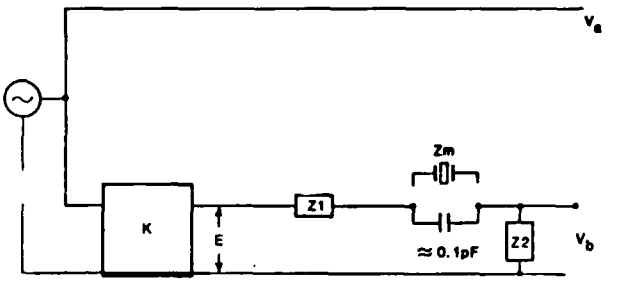



Fig.4. Equivalent circuit of jig

With the shorting blank inserted into the Pi-network expression (9) become:

$$Vbcal = \frac{Z_2}{Z_1+Z_2}\cdot kVacal \qquad (10)$$

$(Z_1+Z_2)$ is the equivalent impedance of the jig seen by the crystal unit and can be expressed as $Zj=\sqrt{Rj^2+Xj^2}\cdot e^{j\varphi}$. Equalizing "k" in (9) and (10) one can obtain:

$$|(Vb/Va)m|\cdot e^{j\theta m}(Zj+Zm)=\sqrt{(Rj^2+Xj^2)}\cdot e^{j\varphi}|(Vb/Va)cal| \qquad (11)$$

From (11) the impedance of the crystal can be derived:

$$R=|Va/Vb)m|\cdot\cos(\theta m-\varphi)\cdot|(Vb/Va)cal|\sqrt{(Rj^2+Xj^2)}-Rj \qquad (12)$$

$$X=|Va/Vb)m|\cdot\sin(\theta m-\varphi)\cdot|(Vb/Va)cal|\sqrt{(Rj^2+Xj^2)}-Xj \qquad (13)$$

and finally, we obtain the admittance of the crystal:

$$G=R/(R^2+X^2);B=-X/(R^2+X^2) \qquad (14)$$

In order to calculate the quartz crystal unit parameters the same procedure as in the reference method can be applied.

In this method the impedance of the Pi-network, seen by the crystal unit under test is subtracted from the actual measured values

and the corrected values of the $Y_{21}$ transadmittance function are thus obtained. This is similar to obtaining the corrected S-parameters values in the reference method, which is accomplished by means of subtraction of the error parameters from the measured ones.

Xj can be found from the relation (12) if we adopt Rj=0. In practise this can be accomplished by measuring any good quality capacitor.

The second method is very similar to the first one. The main difference is in neglecting the reactive part of the Pi-network impedance seen by the crystal side. If we assume that Xj=0, and that the voltage transfer, when a shorting blank is inserted is $(\overline{Vb}/\overline{Va})cal = 1e^{j0}$, than relations (12) and (13) become much simpler:

$$R = |(Va/Vb)m| \cdot \cos(\theta m) \cdot Rj - Rj \qquad (15)$$

$$X = -|(Va/Vb)m| \cdot \sin(\theta m) \cdot Rj \qquad (16)$$

The reciprocal of the real part $(1/Re(Y_{21}))$ of the transadmittance function $Y_{21}(f)$ is:

$$G^{-1} = \frac{1}{G} = \frac{R^2+X^2}{R} = Rj\frac{1+(|(Va/Vb)m|)^2-2(|(Va/Vb)m|)\cos\theta m}{|(Va/Vb)m| \cdot \cos(\theta m)-1} \qquad (17)$$

On the other hand, using the narrow-band approximation:

$$\frac{1}{G} = R_1 + \frac{16 \cdot \pi^2 \cdot L_1^2}{R_1} \cdot \Delta f^2 = Co + C_1 \cdot \Delta f + C_2 \cdot \Delta f^2 \qquad (18)$$

After fitting the values of $G^{-1}(f)$ to a second order polynomial, fs is calculated as:

$$fs = -C_1/2C_2 + fmin \qquad (19)$$

$R_1$ is calculated as the value of the function at the point fs, and $L_1$ as:

$$L_1 = \sqrt{R_1 C_3}/4 \cdot \pi \qquad (20)$$

This method of crystal parameter measurement can be applied up to 200 MHz approximatelly.

The third method, also using a Pi-network as a test jig is a method which can unfortunately be applied only up to 80 MHz, because it does not incorporate Cc influence cancellation, but gives very high accuracy and repeatability.

It is known that for frequencies up to 80MHz, physical or mathematical Cc influence cancellation in most cases can be neglected, so the IEC 444 zero-phase method can be utilized for crystal unit parameter measurement.

The zero-phase transmission line set-up using a Pi-network according to the IEC 444 has been used at the "Mihajlo Pupin" Institute for almost 15 years in both production and quality control processes.

Software has also been developed, evolving from an initial go/no go measurement version to a program for very precise measurement of aging or crystal unit parameters.

The program, let us denote it as "IMP" embodies our experience which has shown that in order to achieve maximum accuracy and repeatability of measurement results instead of strict adherence to IEC 444 (one-point measurement), multi-point measurements at several frequencies are necessary so that the result in each point is the mean value of several measurements. The linear phase characteristic fit technique is then applied to the point set. By applying this technique to various portions of the phase characteristic and testing the correlation between the theoretical (calculated) and measured phase values we have come to the conclusion that the best ($> 0.99999$) correlation coefficient is obtained for the pairs $\pm 6^\circ$, $\pm 5^\circ$, $\pm 4^\circ$. Several measurements are made in these points and the mean value is taken to be the result. The linear fit technique is applied to these three pairs of points using the function:

$$\Delta f = Co + C_1 \cdot \varphi = fi - fmin \qquad (21)$$

and $fr = Co + fmin$ calculated for $\varphi = 0$

If a second order polynomial fit is applied to the amplitude characteristic and the fr value inserted, the resistance Rr of the crystal unit will be obtained.

In order to obtain the $C_1$ value it is necessary to insert $\varphi = \pm 5^\circ$ into the function $\Delta f = Co + C_1 \cdot \varphi$, so as to calculate the frequencies $f-\varphi$ and $f+\varphi$ at which the phase offset is equal in magnitude and opposite in sign. Then we have:

$$C_1 = \frac{(f-\varphi - f+\varphi)}{2 \cdot \pi \cdot f+\varphi \cdot f-\varphi \cdot (Rr+25) \cdot \tan(|\varphi|)} \qquad (22)$$

## Measurement results

Five crystal units in the 1-200 MHz range, for which it can be said that the aging process has reached saturation, so that aging does not influence the measurement results, have been selected for testing the accuracy and repeatability of the parameter measurement results obtained using the reference method and a coaxial test jig, as well the aforementioned methods and the Pi-network as a test jig.

The measurements were made under controlled temperature conditions at $t^\circ = 25^\circ \pm 0.1^\circ C$ with a stabilization period of 30'. The HP 3577A network analyzer was connected to the

"SULZER" frequency standard (long-term stability $10^{-10}$).

All the crystal units were first measured by the reflection type reference method using a coaxial test jig. Crystal units having fn = 2457.510 kHz, fn = 4996.130 kHz and fn = 24997.000 kHz were then, measured using the IMF method, and crystal units having fn=140987.000 kHz and fn=161850.000 kHz were measured using the so-called Williamson method both methods utilizing the Pi-network as a test jig.

The measurements were carried out by performing 100 parameter measurements in succession, subsequent to calibration.

Figures 5-9 depict the cumulative distributions of the absolute value of the measurement results difference from the calculated mean value expressed in [1.E-6] for the resonance frequency and in [%] for resistance and inductance, obtained using a coaxial test jig, and the Pi-network, for every single crystal unit.

Results obtained using a coaxial test jig are marked "S" and the ones obtained using the Pi-network "P".

The difference between the calculated mean values of the measurement results of the crystal unit resonance frequency, resistance and inductance, obtained using a coaxial test jig and the Pi-network are presented in Table 1.

As one can conclude from Table 1, the differences between the mean values are in very reasonable tolerancies, and so are the values, except for the crystal unit having fn = 161 850.000 kHz (0.5 ppm for resonance frequency, 4% difference in $R_1$ and 2% in $L_1$). This is believed to be due to incorrect measurement of Pi-network impedance.

Table 2. shows the difference between the results of measurements made by the transmission type reference method[2] and the measurement results obtained by the IMF reflection type reference method.

Very close correlation between the results obtained by the reflection method (IMF) and the transmission reference method[2] is observed. The differences between the frequency measurements are believed to be due mostly to the different temperatures at which the measurements were carried out.

## Conclusion

First, the close agreement between the results of crystal parameter measurements made by using the reference method 1,2 and the methods utilizing the Pi-network as a test fixture is verified.

It can also be concluded bearing in mind the remarks in the Introduction concerning the satisfaction of basic metrological principles, that the current measurement methods are in the domain of fig. 1c. In order to pass on to fig. 1d, i.e. achieve the maximum accuracy about the true value it is, in our opinion, necessary to manufacture standard crystal units which would serve to determine the magnitude of the random errors which are due to the physics of crystal unit oscillations in time as well as to the imperfection of the measuring equipment. The standard crystal units should operate under conditions similar to those of crystal units in quartz frequency standards, meaning high-stability units in precise thermostats whose aging exhibits a tendency towards saturation.

It is necessary to explore the effect of drive level repeatability in frequency since we believe that in the $10^{-10}$ range this also exerts an influence.


## Acknowledgement

The authors wish to thank W.L.Smith for his frendly support and most helpful advices over many years.


## References


1 W.L.Smith, "An overview of a proposed standard method for measurement of the equivalent electrical parameters of quartz crystal resonators up to 1 GHz", Proc. 6th Quartz Devices Conference, 1984

2 R.C.Peach and S.E.Morris "A system for precision parameter measurements on quartz crystal resonators and bipoles" Proc. AFCS 1985.

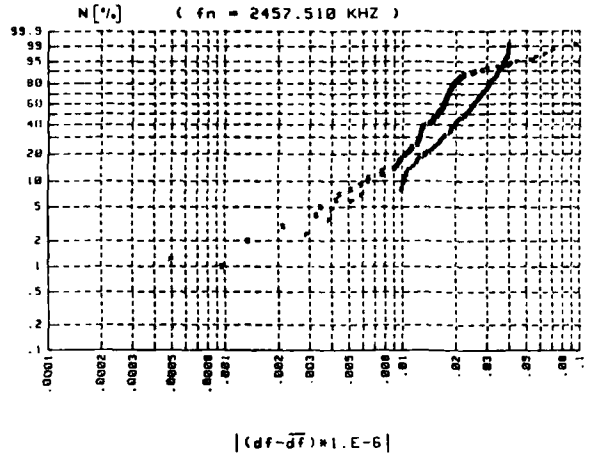
3 "Reference method of measurement for quartz crystal resonator units, using automatic network analyzer techniques for the determination of equivalent electrical parameters between 1 kHz and 1 GHz", 49 (U.S.A.) 86.

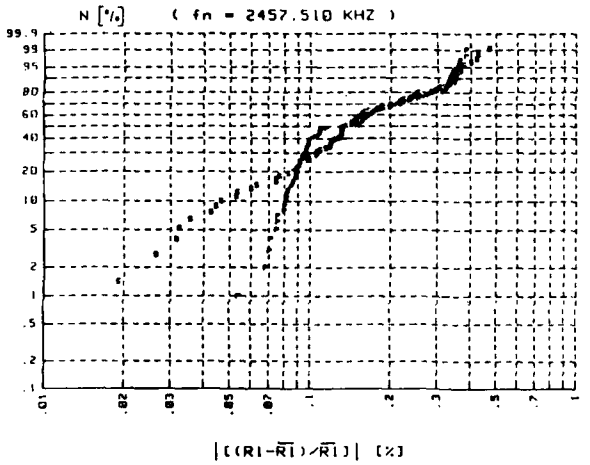
4 W.L.Smith "Electronic Industries Association Standard 512: Some Further Discussion and Comment", 7 th Quartz Devices Conference, pp 43-51, 1985.

5 R.J. Williamson "An Improved method, for measuring the parameters of a quartz resonator", Frequency Control and Synthesis, pp 115-121, 1987.

6 J.P.Aubry, E. Gerard and S. Lechopier "S.Y. Parameters method for accurate measurements of bulk wave crystal resonators of frequencies up to 2 GHz", Proc. 37 th ASFC, pp 306-316, 1983.

7 V. Popović, J. Dabić, V. Popović-Milovanović "Precise measurements of quartz crystal in a Pi-network" 9 th Quartz Devices Conference, pp 21-29, 1987.



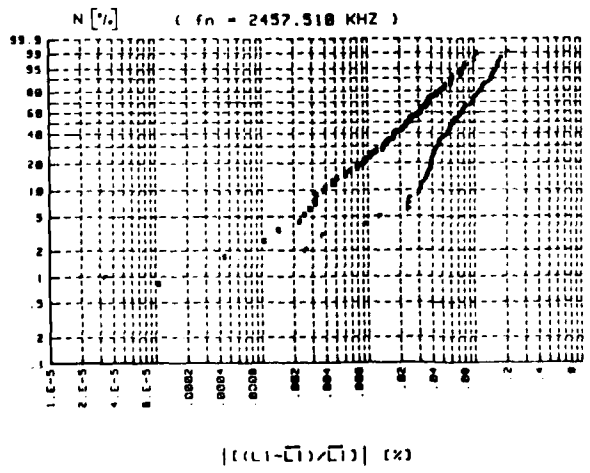
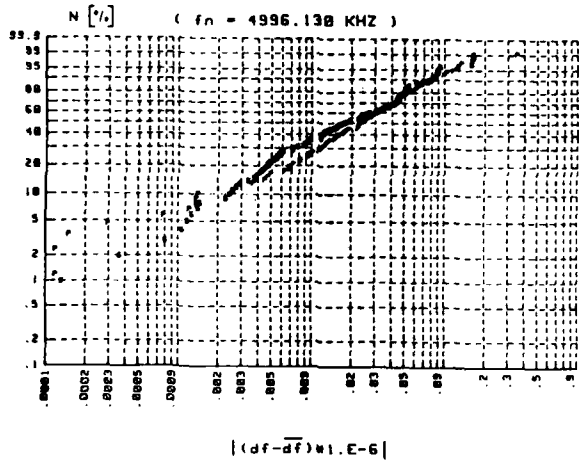
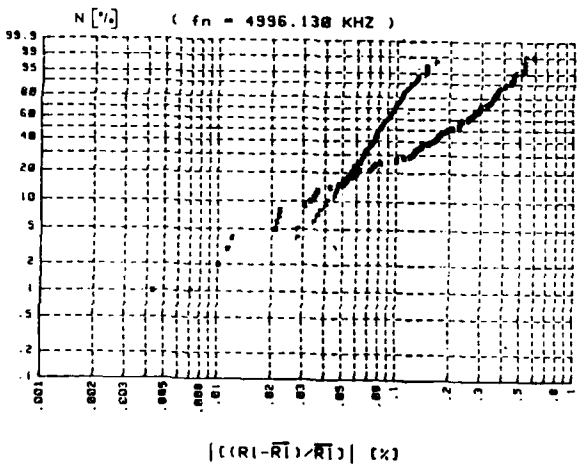




Fig.5. Cumulative distributions of the absolute value of the measurement result difference from the calculated mean value for fs, $R_1$ and $L_1$.
("S" - coaxial test jig, "P" - Pi-network)



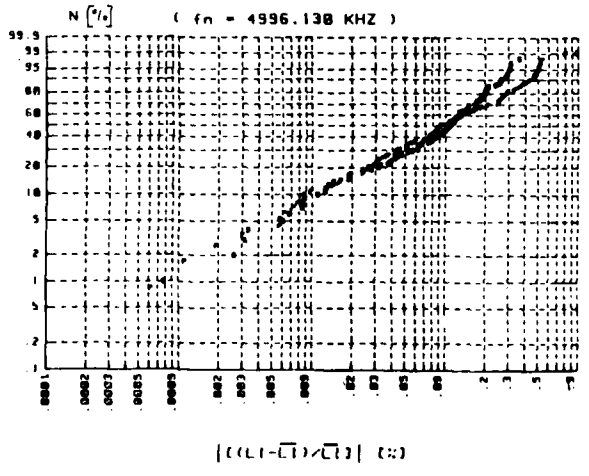
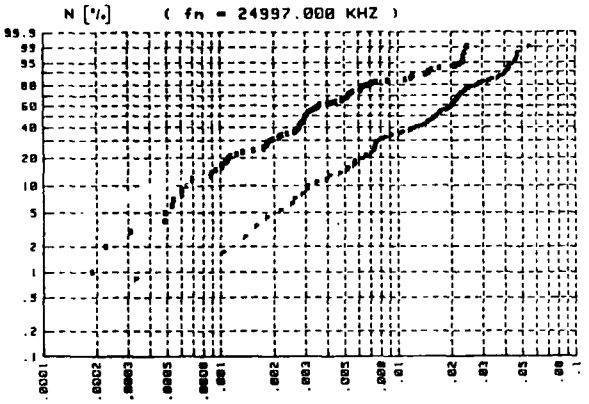
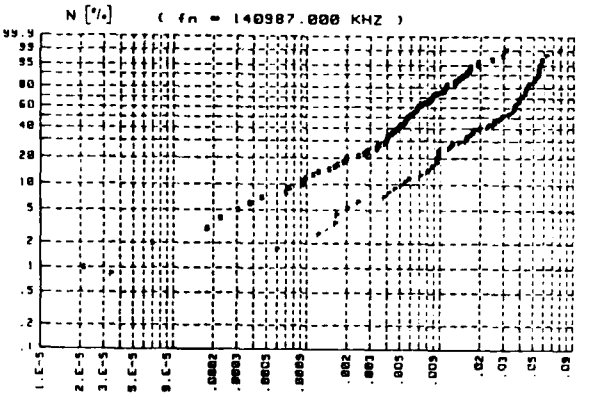
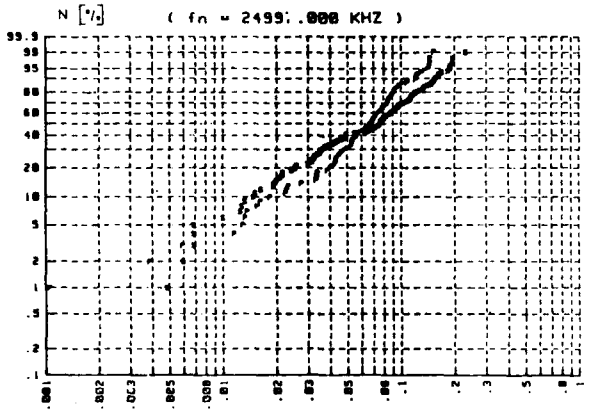
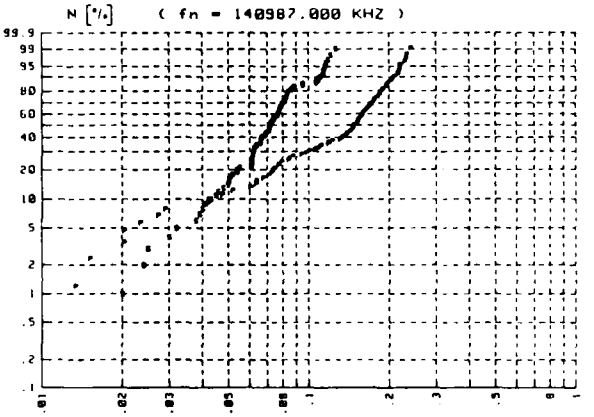



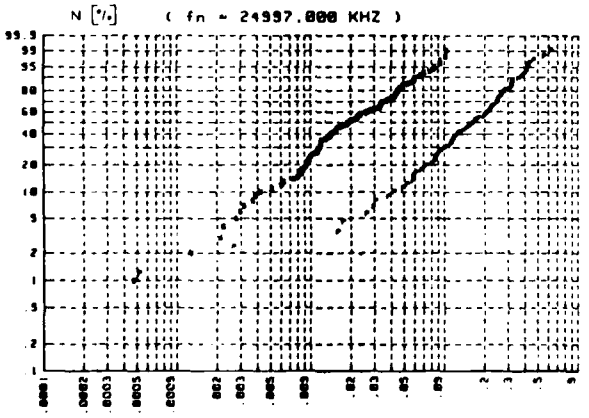

Fig.6. Cumulative distributions of the absolute value of the measurement result difference from the calculated mean value for fs, $R_1$ and $L_1$.
("S" - coaxial test jig, "P" - Pi-network)



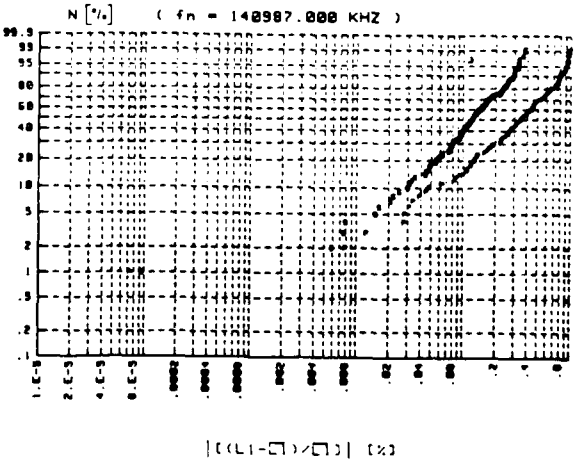
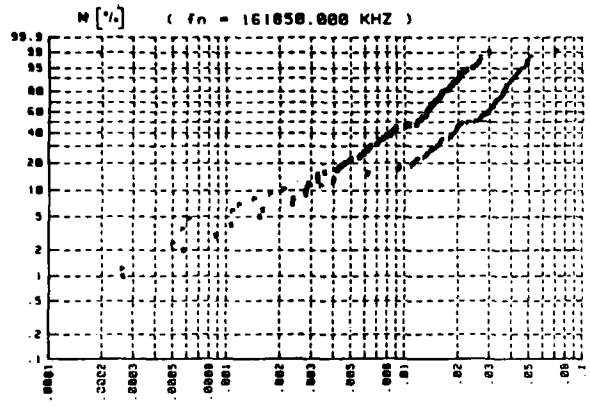
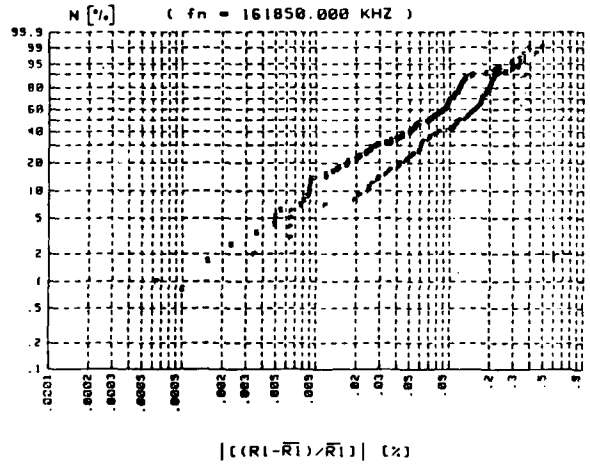




Fig.7. Cumulative distrubutions of the absolute value of the measurement result difference from the calculated mean value for fs, $R_1$ and $L_1$.
("S" - coaxial test jig, "P" - Pi-network)

Fig.8. Cumulative distrubutions of the absolute value of the measurement result difference from the calculated mean value for fs, $R_1$ and $L_1$.
("S" - coaxial test jig, "P" - Pi-network)



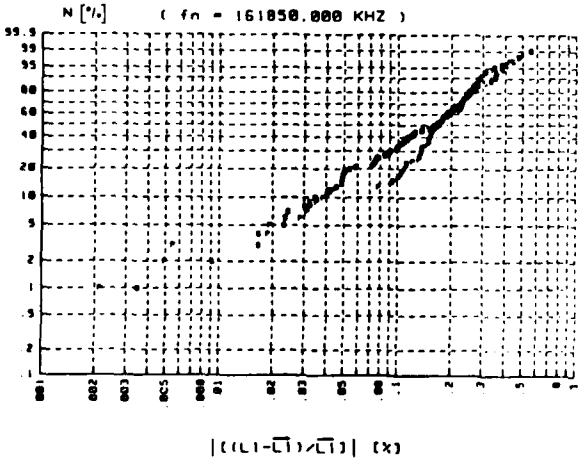




Fig.9. Cumulative distrututions of the absolute value of the measurement result difference from the calculated mean value for fs, $R_1$ and $L_1$.
("S" - ccaxial test jig, "P" - Pi-network)

| fn [KHz] | jig | $(\overline{\Delta f/f} \pm \sigma$ [*1.E-6] | $\overline{R}$(ohm)$\pm\sigma$[%] | $\overline{L}$(mh)$\pm\sigma$[%] |
|---|---|---|---|---|
| 2 457.510 | coax | 0.321 ± 0.010 | 66.372± 0.10 | 532.731±0.04 |
| | Pi-net. | 0.305 ± 0.027 | 65.930± 0.10 | 531.915±0.00 |
| | difference | 0.016*1.E-6 | 0.65 % | 0.15 % |
| 4 996.130 | coax | -6.636±0.035 | 15.952± 0.09 | 116.316±0.1 |
| | Pi-net. | -6.627±0.035 | 15.710± 0.27 | 115.422±0.1 |
| | difference | 0.009*1.E-6 | 1.5 % | 0.7 % |
| 24 997.000 | coax | 0.6 ±0.005 | 6.350 ± 0.09 | 2.538 ±0.035 |
| | Pi-net. | 0.730 ±0.021 | 6.401 ± 0.00 | 2.565 ± 0.21 |
| | difference | 0.130*1.E-6 | 0.6 % | 1.06 % |
| 140 987.000 | coax | 10.940 ±0.00 | 90.583± 0.00 | 9.025 ± 0.16 |
| | Pi-net. | 10.785±0.020 | 91.582± 0.10 | 8.041 ± 0.43 |
| | difference | 0.155*1.E-6 | 1.1 % | 0.19 % |
| 161 850.000 | coax | 2.531 ± 0.014 | 105.475±0.1 | 6.394 ± 0.22 |
| | Pi-net. | 2.362 ± 0.031 | 109.711±0.10 | 6.526 ± 0.22 |
| | difference | 0.169*1.E-6 | 4 % | 2.06 % |

Table 1.

Differences between the calculated mean values of the measurement results using a coaxial test jig and the Pi - network.

| fn [KHz] | method | Δf/f*1.E-6 | R(ohm) | L(mh) |
|---|---|---|---|---|
| 2 457.510 | transm.[2] | 1.05 | 65.49 | 528.4 |
| | reflec(IMP) | 0.321 | 66.372 | 532.915 |
| | difference | 0.729 | 1.3 % | 0.85 % |
| 4 996.130 | transm.[2] | -3.622 | 16.09 | 114.8 |
| | reflec(IMP) | -6.636 | 15.952 | 116.316 |
| | difference | 3.014 | 0.85 % | 1.3 % |
| 24 997.000 | transm.[2] | 1.04 | 6.345 | 2.500 |
| | reflec(IMP) | 0.6 | 6.350 | 2.538 |
| | difference | 0.59 | 0.2 % | 1.6 % |
| 140 987.000 | transm.[2] | 11.114 | 91.23 | 8.897 |
| | reflec(IMP) | 10.940 | 90.583 | 9.025 |
| | difference | 0.174 | 0.714 % | 0.897 % |
| 161 850.000 | transm.[2] | 2.074 | 107.4 | 6.474 |
| | reflec(IMP) | 2.531 | 105.475 | 6.394 |
| | difference | 0.457 | 1.8 % | 1.25 % |

Table 2.

Comparision of measurment results made by the reference (transmission) method[2] and the reference (reflection) method (IMP).

42nd Annual Frequency Control Symposium - 1988

# ENSEMBLE TIME AND FREQUENCY STABILITY OF GPS SATELLITE CLOCKS

by

David W. Allan and Trudi K. Peppler
Time and Frequency Division
National Bureau of Standards
Boulder, Colorado 80303

## Abstract

In anticipation of cross-link ranging in the Global Positioning System (GPS) and the establishing of GPS ensemble time, we have investigated the frequency stability characteristics of the clocks in the constellation. Once those characteristics were ascertained we then optimally combined the clock readings using the NBS time scale algorithm software to generate a GPS space clock ensemble.

Frequency drifts of some parts in $10^{16}$ per day were observed in the GPS cesium beam frequency standards on board the space vehicles. The white noise frequency modulation (FM) in the cesium clocks gave rise to a daily time prediction error of only 8 to 10 ns. The long-term random deviations in the cesium clocks were nominally modeled by a flicker noise FM at a few parts in $10^{14}$ for $\sigma_y(\tau)$. The rubidium gas-cell frequency standards with no temperature control on their base plate showed significant frequency perturbations during the solar eclipse season. Because of this, these clocks were not included in the GPS space clock ensemble. Their short-term stability is better than the cesium clocks, but their long-term stability was characterized by a random-walk FM at about 1 part in $10^{13}$ at one day on a $\sigma_y(\tau)$ diagram. The rubidium clock with a temperature controlled base plate had a comparable long-term performance to the cesium clocks except for a much larger linear frequency drift and a rate of change of the frequency drift.

The GPS space clock ensemble frequency stability reached about 2 parts in $10^{14}$. Not only is this comparable to some of the best clock ensembles at major timing centers throughout the world, but the calculated time dispersion for autonomous GPS performance has significant implications. The reference for this analysis was the NBS time scale.

## Introduction and Procedure

The list of uses of the GPS has become very impressive and continues to increase. Aside from navigation, for which it was designed, the other main areas benefitting from the availability of GPS have been the geodesy and timing communities. The timing community has used GPS in two basic modes: 1) direct access to the GPS or UTC(USNO MC) time, which provides an accuracy of better than 100 ns [1], and 2) use of GPS in the common-view mode -- providing a time and frequency measurement capability between remote sites to about 10 ns in time and 1 part in $10^{14}$ for integration times longer than a few days [2,3]. It is the purpose of this paper to demonstrate a new and powerful usage of GPS by combining the readings of the free running and independent space vehicle (SV) clocks to generate a computed time scale better than the best clock in the system.

The motivation for this work has been two-fold: First, the GPS Joint Program Office (JPO) has plans to generate an ensemble time -- both for the master control station and for the GPS [4]. This paper demonstrates the quality available from the ensemble concept with real data from the SV clocks. The future quality should be significantly better with the full constellation of SV clocks. Second, it has been known for some time that some of the individual SV clock performances have been better in the space environment than when they were pre-launch tested on the ground. To have a calculated clock time better than the best one of these SV clocks in both the short and long term would be quite outstanding and would provide an independent time and frequency metrology tool for accessing the performance of earth bound clocks [5,6]. Of particular concern at the present time are annual variations in cesium clocks.

The performance of the individual SV clocks was characterized for both random and systematic behaviors against the free running NBS(AT1). All of the clocks had frequency drifts and these were subtracted. Table 1 shows the values of the frequency drifts removed. The short-term instabilities for the GPS cesium and rubidium clocks are well modeled by white noise FM, and the long-term instabilities by either flicker-noise or random-walk FM.

Once the random and systematic behavior of the clocks were characterized, the NBS time scale algorithm for generating NBS(AT1) and UTC(NBS) was used to generate a GPS SV space clock ensemble time [7-9]. The NBS time scale algorithm deals with both the short and the long term processes. Individual clock sigmas are estimated, from which adaptive weights are calculated. By optimizing the integration time constant for the adaptive weights, we can de-weight the clocks automatically during an abnormal frequency deviation. A constant is entered for each clock for the frequency updates, $m_i$, which is a function of both the white FM and of the flicker or random-walk FM models for that clock. The entrance and exit of each SV clock into or out of the ensemble is a function of the operational realities for each. The NBS algorithm is designed to deal with these entrances and exits with minimum perturbation to the ensemble time output.

Contribution of the U.S. Government, not subject to copyright.

## GPS Clock Characteristics

Figures 1a and 1b are plots of the frequency and the frequency stability of NAVSTAR or SVN 3 (PRN 6). The eclipse periods are evident at 1/2 year spacings. This rubidium clock is well characterized by random walk FM in the long term.

Figures 2a and 2b are corresponding plots for NAVSTAR or SVN 4 (PRN 8). Again we see a very similar behavior as for SVN 3 for this rubidium clock.

Figures 3a and 3b are corresponding plots for NAVSTAR or SVN 6 (PRN 9). As a Block 1 rubidium clock, its performance is also very similar to SVN 3 and 4.

All of the above rubidium clocks had base plate temperature variations of a few degrees. The effect of the eclipse season on each of the clocks is clearly visible, but of a non-statistical nature. These clocks had adverse effects on the ensemble time output, and were therefore not included in the GPS space clock ensemble. Their long-term weighting factors were nearly negligible also; hence, their exclusion had no importance to the long-term stability of GPS ensemble time, which is one of the main points of this paper.

Figures 4a and 4b are corresponding plots for NAVSTAR or SVN 8 (PRN 11). Though this also was a rubidium clock, several upgrades were made on it over the previous rubidium clocks. One of the most important of which was to improve the temperature control for the clock's base plate from a few degrees to about 0.1 degree. These upgrades improved the long-term performance considerably, and reduced the effect of the solar eclipse seasons. The Block II SV rubidium clocks will have a similar base plate temperature control.

The time residuals for this clock were more nearly modeled by a cubic (drift of the frequency drift). Because the cubic makes long-term predictions difficult, we deliberately set the long-term time-constant for the frequency update low. We were then able to include this clock in the GPS space clock ensemble. The frequency drift was estimated from the minimum squared error of a quadratic fit to the frequency. This fitting process acts like a high pass filter, so the stability values for large $\tau$ values are biased low. The frequency drift on all of the other clocks was estimated using a second difference routine [11]. The long-term stability values are also biased low in this case, but not as much as for the removal of a cubic.

Figures 5a and 5b are corresponding plots for SVN 8 also, but for the cesium clock after switching from the rubidium clock. The rubidium clock was still demonstrating excellent frequency stability at the time of the change, but the switch to cesium occurred because of a failure in another part of the clock system.

Figures 6a and 6b are corresponding plots for NAVSTAR or SVN 9 (PRN 13).

Figures 7a and 7b are corresponding plots for NAVSTAR or SVN 10 (PRN 12).

Figures 8a, 8b, 8c, and 8d are corresponding plots for NAVSTAR or SVN 11 (PRN 3). Figures 8a and 8b are before the satellite was moved to a new position in the GPS constellation, and figures 8c and 8d are after the move.

Table 2 shows the starting parameters that were entered for each clock.

## Results and Conclusions

Figures 9a and 9b are the frequency deviation and the frequency stability plots of the GPS SV ensemble time respectively with NBS(AT1) being the reference. The short term stability was modeled by a white-noise FM at $\sigma_y(\tau = 2 \text{ days}) = 3.7 \times 10^{-14}$, which corresponds to 2.6 ns ( < 1 meter) rms GPS SV ensemble error contribution at 8 hours (the current nominal upload interval). Concerning 180 day autonomous operation performance capability, we estimate the GPS SV ensemble's rms error of prediction to be about 600 ns (180 m).

The primary cesium clock for the Federal Republic of Germany (PTB CS 1) was used along with NBS(AT1) to perform a three-cornered hat experiment in order to evaluate the independent stability of the GPS SV ensemble in its long-term performance. Table 3 gives the results.

We conclude that an optimum weighted combination of the times and frequencies from the GPS SV clocks can yield an ensemble output with outstanding performance both in the short-term as well as in the long-term integration times or prediction intervals. Some caution should be expressed, however, because even with about 800 days of data being analyzed, the confidence of the estimate for $\tau$ values of the order of a year is quite poor. Still it would appear that the use of such an approach in the future should be very beneficial to the system both in terms of performance and in terms of reliability and redundancy. Some caution also needs to be exercised in how the clocks are configured; for example, frequency steps in a clock can have some adverse effect on the ensemble performance. Research is being conducted to deal with this problem.

## References


[1] W.J. Klepczynski, H.F. Fliegel, and D.W. Allan, "GPS Time Steering," Proc. of 18th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Washington, D.C., Dec. 2-4, 1986, pp. 237-249.

[2] D.W. Allan, D.D. Davis, M.A. Weiss, A. Clements, B. Guinot, M. Granveaud, K. Dorenwendt, B. Fischer, P. Hetzel, S. Aoki, M.-K. Fujimoto, L. Charron, and N. Ashby, "Accuracy of International Time and Frequency Comparisons Via Global Positioning System Satellites in Common-View," IEEE Transactions on Instrumentation and Measurement, IM-34, 118-125 (1985).

[3] M.A. Weiss and D.W. Allan, "An NBS Calibration Procedure for Providing Time and Frequency at a Remote Site by Weighting and Smoothing of

GPS Common View Data," IEEE Trans. on I & M, IM-36, 572-578 (1987).

[4] F. Varnum, D.R. Brown, D.W. Allan, and T.K. Peppler, "Comparison of Time Scales Generated with the NBS Ensembling Algorithm," Proc. of the 19th Precise Time and Time Interval (PTTI) Planning Meeting, 1987.

[5] E. Bava, F. Cordara, V. Pettiti, and P. Tavella, "Analysis of the Seasonal Effects on a Cesium Clock to Improve the Long-Term Stability of a Times Scale," Proc. of the Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, Redondo Beach, CA, Dec. 1-3, 1987, pp. 185-202.

[6] J.E. Gray, H.E. Machlan, and D.W. Allan, "The Effect of Humidity on Commercial Cesium Beam Atomic Clocks," Proc. of the 42nd Annual Symposium on Frequency Control, Baltimore, MD, June 1-3, 1988, to be published.

[7] D.W. Allan and M.A. Weiss, "The NBS Time Scale Algorithm, AT1," NBS Technical Note 1316.

[8] M.A. Weiss, D.W. Allan, and T.K. Peppler, "A Study of the NBS Time Scale Algorithm," submitted for publication in the Proc. of CPEM'88. Digest version published in 1988 Conf. on Precision Electromagnetic Measurements, Tsukuba Science City, Japan, June 7-10, 1988.

[9] D.W. Allan, M.A. Weiss, and T.K. Peppler, "In Search of the Best Clock," submitted for publication in the Proc. of CPEM'88. Digest version published in 1988 Conf. on Precision Electromagnetic Measurements (CPEM), Tsukuba Science City, Japan, June 7-10, 1988.

[10] D.A. Howe, D.A. Allan, and J.A. Barnes, "Properties of Signal Sources and Measurement Methods," Proc. of the 35th Annual Symposium on Frequency Control, 1981.

[11] D.W. Allan, "Time and Frequency (Time-Domain) Characterization, Estimation, and Prediction of Precision Clocks and Oscillators," IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control, UFFC-34, 647-654 (1987).

TABLE 1

(Parts in $10^{14}$)

| SVN (PRN) | Mean Freq. | Drift/day | Drift/day/day |
|---|---|---|---|
| 3 (6) | -1285.7 | -0.65 | |
| 4 (8) | 4370.9 | -3.27 | |
| 6 (9) | 10861.0 | -7.53 | |
| 6 (9) | 1177.4 | -5.07 | |
| 8 (11) | -931.6 | -13.51 | +0.0194 |
| 9 (13) | 219.0 | -0.02 | |
| 10 (12) | 572.9 | -0.07 | |
| 11 (3) | -54.4 | -0.03* | |
| 11 (3) | -69.9 | -0.06 | |

*The frequency drift for this segment was sufficiently small that it was not subtracted before using the data in the GPS ensemble.

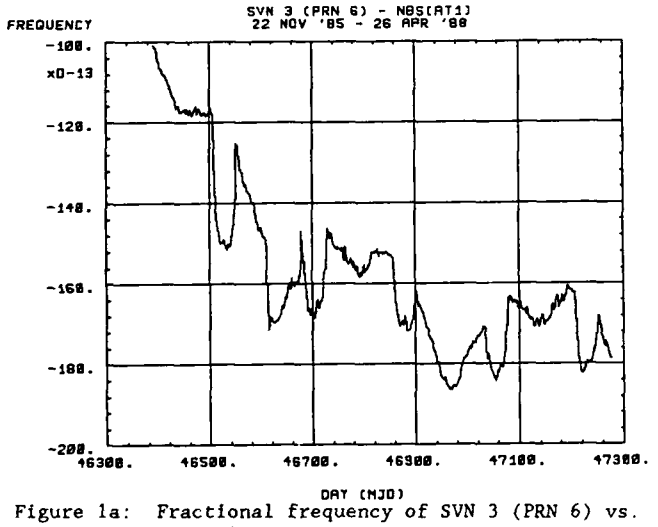



Figure 1a: Fractional frequency of SVN 3 (PRN 6) vs. NBS(AT1).

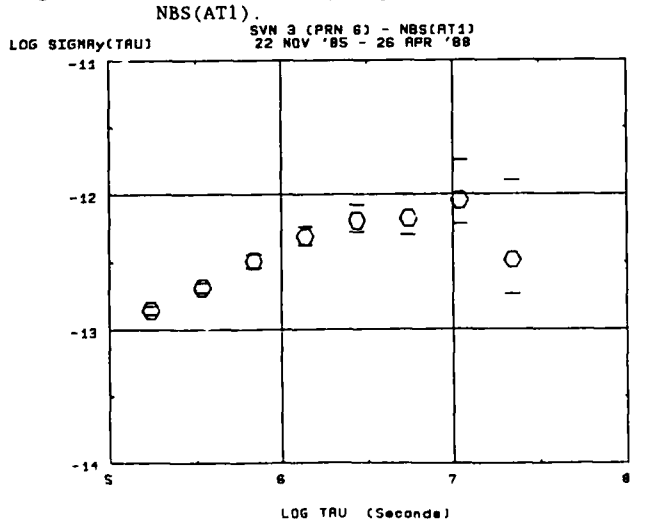



Figure 1b: Frequency stability of SVN 3 (PRN 6) after the drift was removed vs. NBS(AT1).

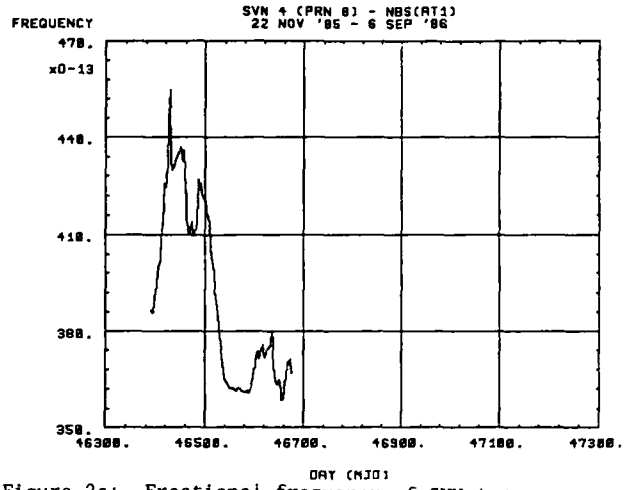



Figure 2a: Fractional frequency of SVN 4 (PRN 8) vs. NBS(AT1).

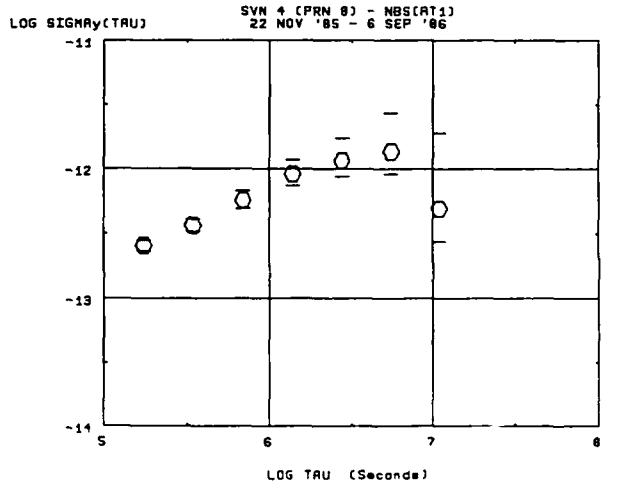



Figure 2b: Frequency stability of SVN 4 (PRN 8) after the drift was removed vs. NBS(AT1).

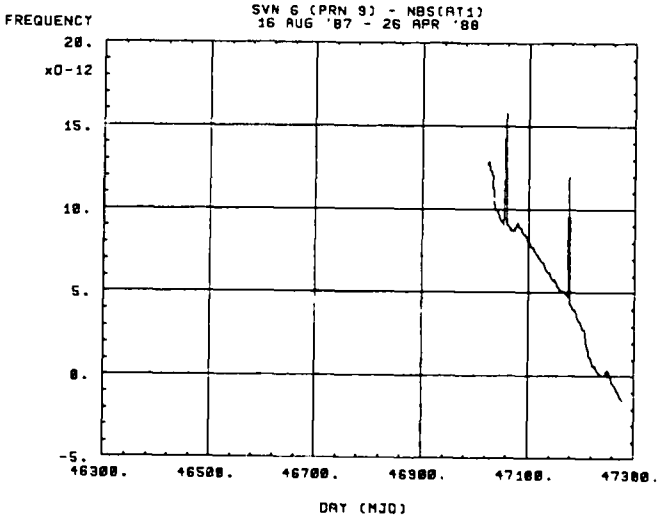



Figure 3a: Fractional frequency of SVN 6 (PRN 9) vs. NBS(AT1).

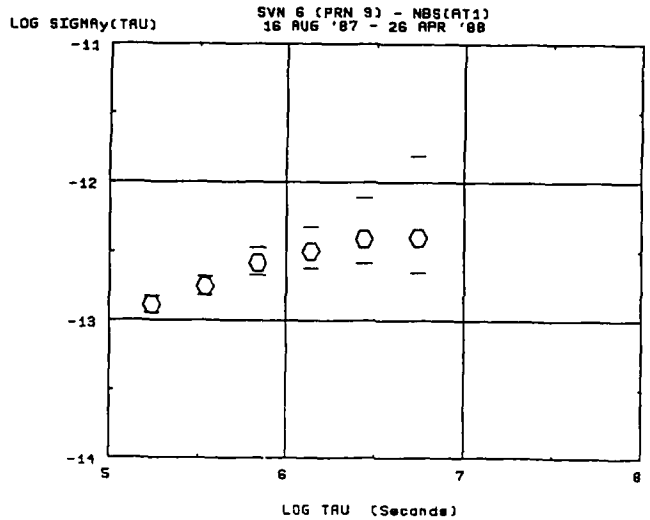



Figure 3b: Frequency stability of SVN 6 (PRN9) after the drift was removed vs. NBS(AT1).

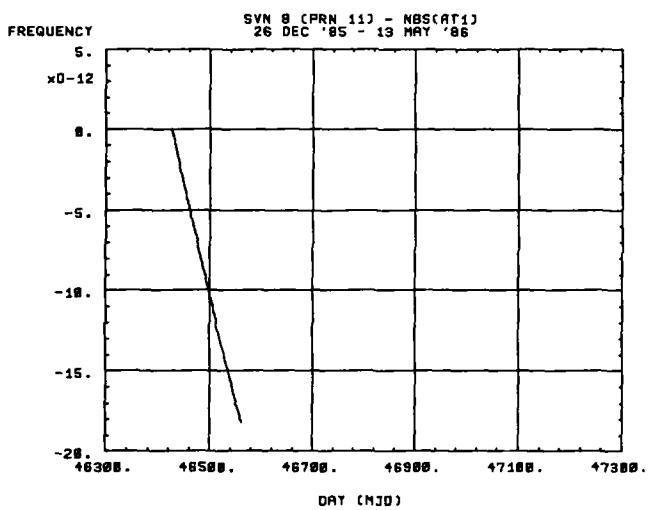



Figure 4a: Fractional frequency of SVN 8 (PRN 11) rubidium clock vs. NBS(AT1).

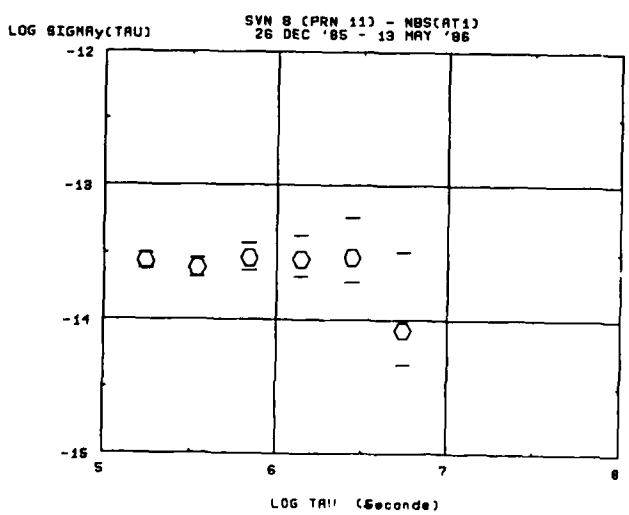



Figure 4b: Frequency stability of SVN 8 (PRN 11) rubidium clock after quadratic fit to frequency was removed vs. NBS(AT1).

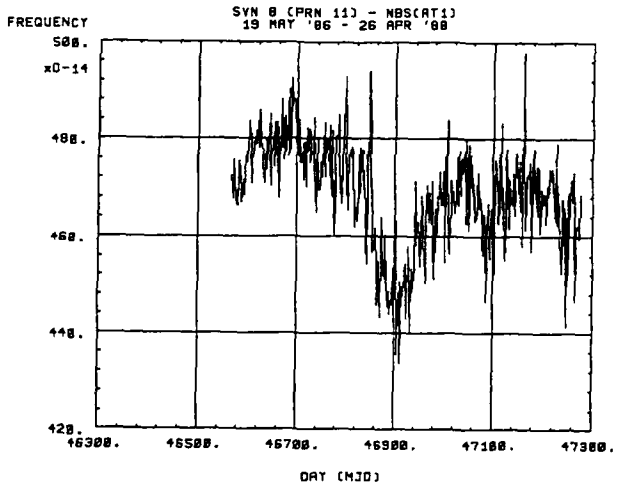



Figure 5a: Fractional frequency of SVN 8 (PRN 11) after switching to the cesium clock vs. NBS(AT1).

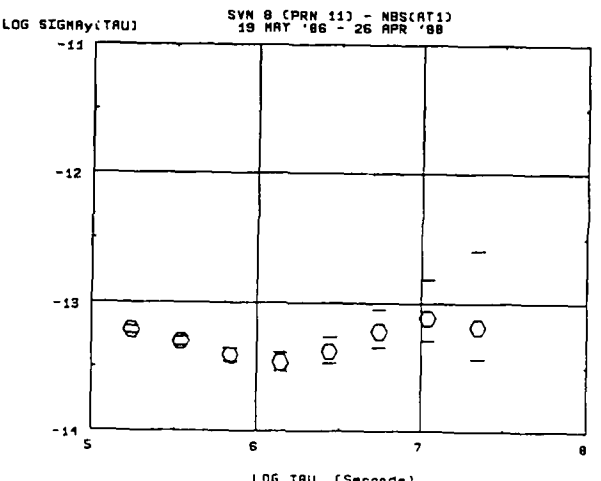



Figure 5b: Frequency stability of SVN 8 (PRN 11) after switching to the cesium clock and after the drift was removed vs. NBS(AT1).

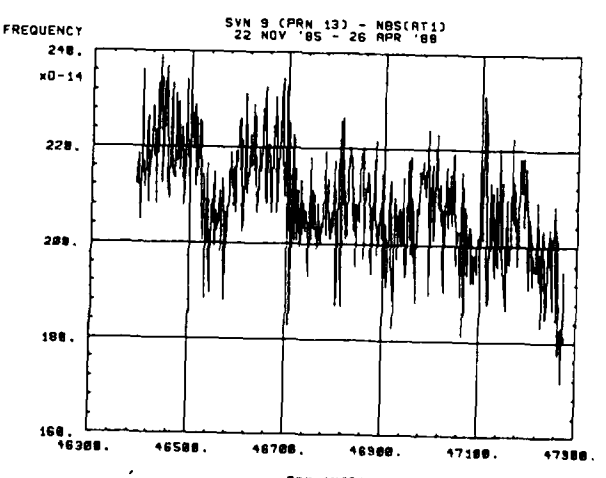



Figure 6a: Fractional frequency of SVN 9 (PRN 13) vs. NBS(AT1).

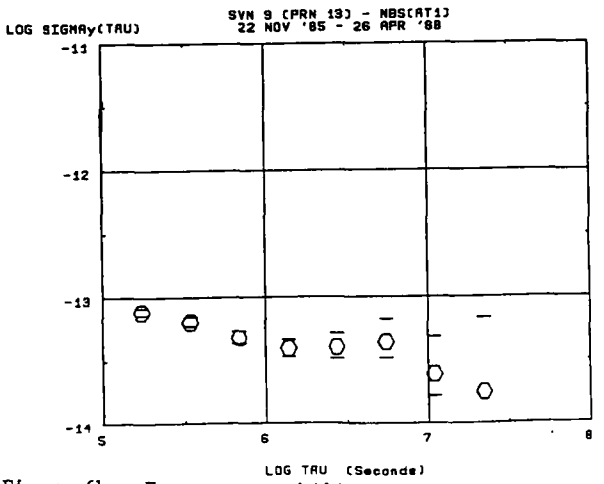



Figure 6b: Frequency stability of SVN 9 (PRN 13) after the drift was removed vs. NBS(AT1).

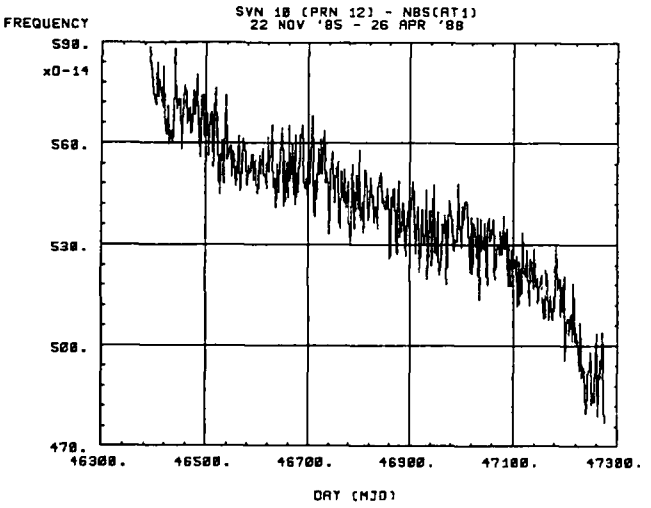



Figure 7a: Fractional frequency of SVN 10 (PRN 12) vs. NBS(AT1).

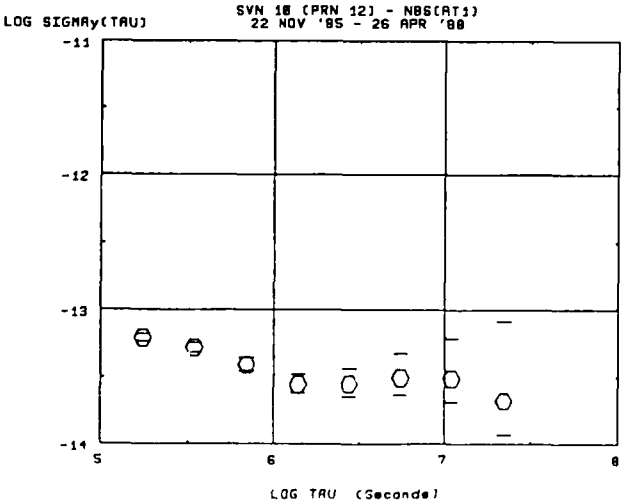



Figure 7b: Frequency stability of SVN 10 (PRN 12) after the drift was removed vs. NBS(AT1).

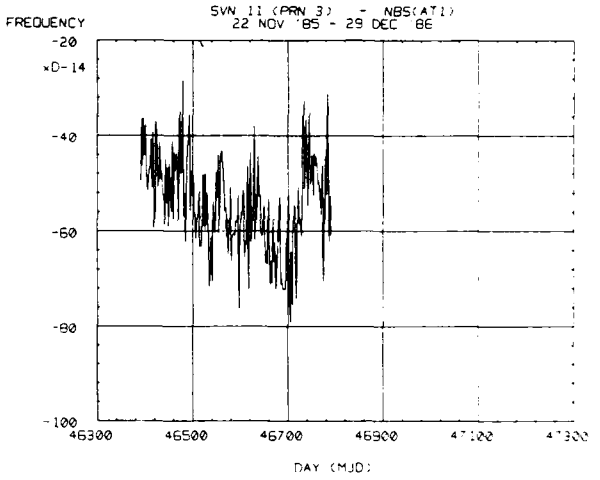



Figure 8a: Fractional frequency of SVN 11 (PRN 3) before the space vehicle was moved vs. NBS(AT1).

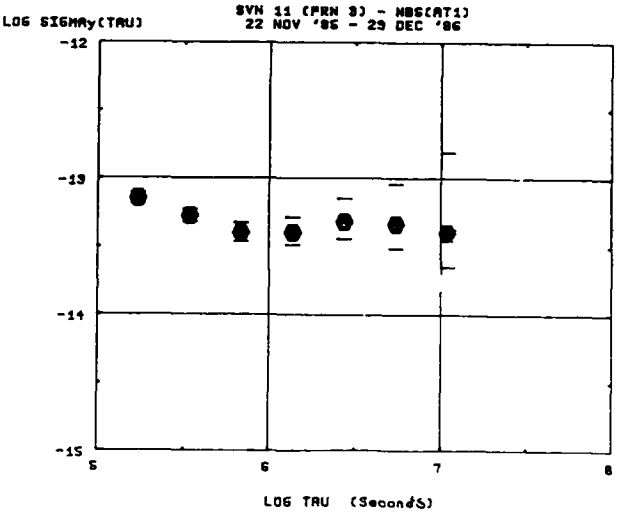



Figure 8b: Frequency stability of SVN 11 (PRN 3) with no drift removed and before the space vehicle was moved vs. NBS(AT1).

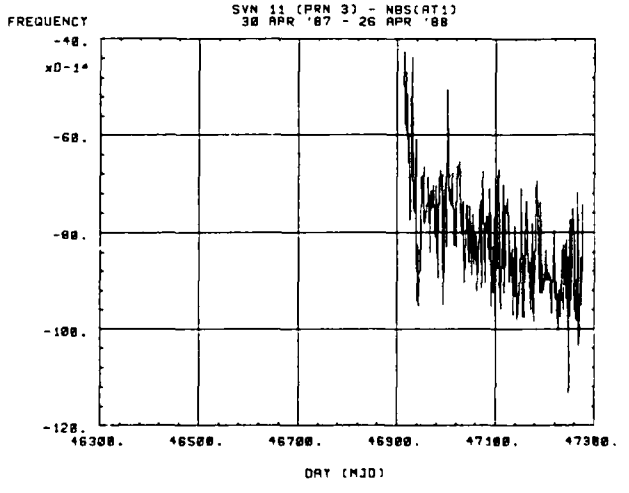



Figure 8c: Fractional frequency of SVN 11 (PRN 3) after the space vehicle was moved vs. NBS(AT1).

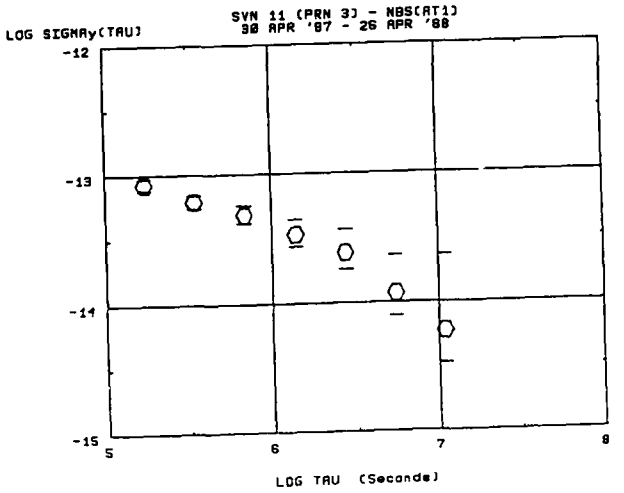



Figure 8d: Frequency stability of NBS 11 (PRN 3) after the drift was removed and after the space vehicle was moved vs. NBS(AT1).

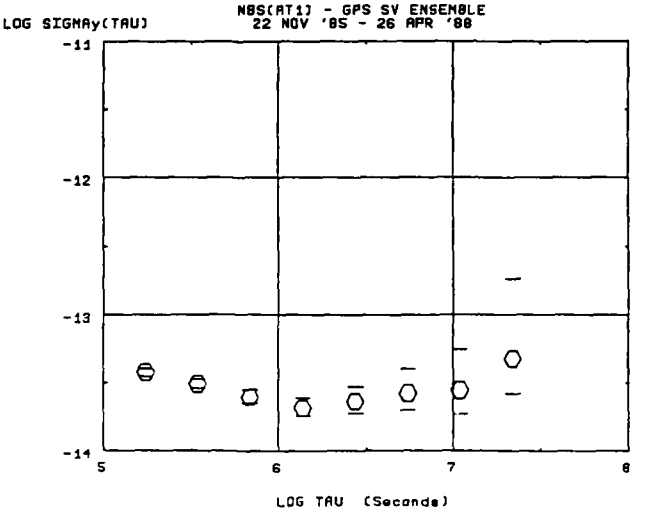



Figure 9b: Frequency stability of the GPS SV space clock ensemble vs. NBS(AT1).

TABLE 2

| SVN (PRN) | Initial Clock Sigma at two-day prediction interval (ns) | Freq. update constant ($m_i$) |
|---|---|---|
| 6 (9) | 20.0 | 0.2 |
| 8 (11) | 8.0 | 1.0 |
| 8 (11) | 11.0 | 1.5 |
| 9 (13) | 15.0 | 4.0 |
| 10 (12) | 12.0 | 4.0 |
| 11 (3) | 14.0 | 4.0 |
| 11 (3) | 14.0 | 5.0 |

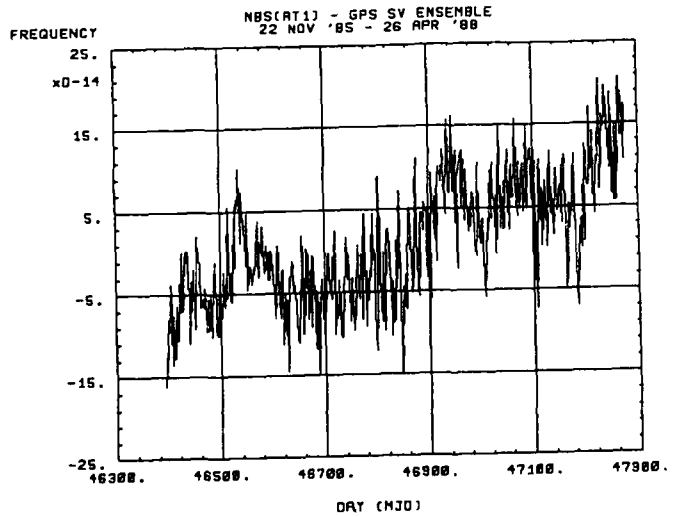



Figure 9a: Fractional frequency deviation of the GPS SV space clock ensemble vs. NBS(AT1).

TABLE 3

Frequency Stability, $\sigma_y(\tau)$ (Parts in $10^{14}$)

| $\tau$(days) | PTB(CS 1) | GPS SV ENSEMBLE | NBS(AT1) |
|---|---|---|---|
| 10 | 2.0 | 2.1 | 0.9 |
| 20 | 1.6 | 2.0 | 0.3 |
| 40 | 1.9 | 2.3 | ** |
| 80 | 2.5 | 2.9 | ** |
| 160 | 2.8 | 3.7 | ** |
| 320 | 4.0 | 6.3 | ** |

**NBS(AT1) had a negative variance for these days as a result of the finite data length and the confidence of the estimates [10]. In a probablistic sense the negative variance implies that the instability of NBS(AT1) was less than either of the other two for these sample times, $\tau$.

42nd Annual Frequency Control Symposium - 1988

# PRELIMINARY COMPARISON BETWEEN GPS AND TWO-WAY SATELLITE TIME TRANSFER

W. J. Klepczynski, P. J. Wheeler, W. Powell,
J. Jeffries, A. Myers, R. T. Clarke
U.S. Naval Observatory, Washington, D.C. 20392

W. Hanson, J. Jespersen, D. Howe
National Bureau of Standards, Boulder, Colorado 80302

## Abstract

The time scales of the U.S. Naval Observatory in Washington, D.C. and the National Bureau of Standards in Boulder, Colorado were compared during the same time periods by the common view mode using the satellites of the Global Positioning System (GPS) and by the two-way mode using a U.S. domestic communications satellite. Data collected over a 3-month period showed residuals from a linear regression of 10 nanoseconds for the two-way technique and 30 nanoseconds for common view GPS. The two-way technique achieved better than 500 ps of precision with less than 2 minutes of data. Absolute calibration of either technique was not attempted during this measurement period.

## Introduction

Satellites have been used for the exchange of timing information as early as 1962. In this first experiment, clocks between England and the U.S. were synchronized by the simultaneous exchange of timing information through the Telstar satellite (Steele et al, 1964). The use of this two-way technique to synchronize remotely located clocks had the advantage of being independent of the motion of the satellite during the course of the performance of the experiment and the geographic positions of the stations participating in the experiment. Since the signals transmitted by both participating stations travel through the same atmosphere, the timing data are only affected by the differential variations that occur in the path delay during the transmission time of the signals (usually about 1/4 second) and those that might arise if different transmit and receive frequencies are used by the stations. If the satellite does not simply reflect the transmitted signals, but retransmits them using some type of transponder, then differential path delay through the satellite must be taken into account.

It is also possible to do one-way time transfers using satellites. In this case either the satellite directly transmits time signals if it has a clock on board or relays signals from some station either by reflection or retransmission. One-way time transfers are used to either distribute time (Klepczynski, 1983) or to synchronize clocks through the use of common view techniques (Allan and Weiss, 1980). In the former case, the user must first account for his distance from the satellite and propagation path delays which might affect the signals. In the latter case many of the effects cancel because both participants are observing the same signal at the same time. This latter technique more closely approaches the precision of the two-way technique. It is the purpose of this study to compare the precisions attainable with the two-way and common view techniques.

## Hardware Configuration

At each site, there is a complete satellite Earth station, one at National Bureau of Standards (NBS) in Boulder and another at the U.S. Naval Observatory (USNO) in Washington. Each Earth station is located in close proximity to the primary time scales of the laboratory. These Earth stations are basically identical except for the size of the antennas. The antenna diameters are 6.1 and 4.5 meters for NBS and the USNO respectively. These rather large antennas are not necessary to achieve results typical of that reported in this paper. They were purchased during a period when new FCC rules relating to the Fixed Satellite Service (FSS) were being established. The rules regarding antenna side lobes were severe enough that only the larger sized antennas qualified at that time. Today, antennas as small as 1.8 meter in diameter qualify for use in the FSS and will work satisfactorily for two-way time transfers with very modest power outputs, a couple of watts being typical. The Earth stations transmit at 14.307 GHz and receive at 12.007 GHz, frequencies assigned by the operator of the satellite transponder. Low noise preamplifiers with noise temperatures less than 250 K are employed. Dual conversion up and down converters are also used. One Hertz pulses from each time scale are converted to pseudo-noise sequences that bi-phase modulate a 70 MHz carrier in a commercially available modem. The modem uses a delay locked loop in the demodulator to recover a 1 PPS signal (Hartl et al, 1983). A commercial U.S. domestic communications satellite provides the channel for the two-way time transfer. The satellite, located in geostationary orbit at 95 degrees West longitude, carries 10 transponders each with 43 MHz of usable bandwidth. The two-way system uses approximately 5 MHz bandwidth of one transponder and 2% of its total available power. This channel has been used one-half hour every Monday, Wednesday and Friday since August of last year. Data are in the form of time interval readings generated using a high precision time interval counter. The data are collected by computer and archived at the USNO where they are available for retrieval by the participating laboratories over telephone line. Final results require that data taken at each end be subtracted on a point-by-point basis. Collocated with each of the satellite Earth stations were GPS timing receivers. Each receiver was connected to a computer for the collection of measurements with retrieval and archiving at the USNO made possible through connection with telephone modems. In all cases, measurements were made with reference to the primary time scales maintained at each laboratory.

## Two-Way Time Transfer Data

The fundamental datum obtained during the course of a two-way time transfer experiment is a time interval counter reading and the instant of time at which the reading was made. The time interval counter at Station A is started by the one pulse per second (1 PPS) which is being transmitted by Station A and stopped by the 1 PPS which is received from Station B. The reading is a measure of the difference between the clocks at Station A and Station B plus the total travel time of the signal from Station B to Station A. Similarly, Station B records a time interval counter reading and the time of the measurement. The reading of the time interval counter at Station B is a measure of the difference between the clocks at Station B and Station A plus the total travel time of the signal from Station A to Station B. In order to obtain the difference between the clocks at Station A and Station B, it is necessary to first bring the data from the two stations together. Then one simply divides the difference between the two readings by two. The readings must first be adjusted by any calibrations that have been made at each site. It is also assumed that the propagation path delays that each signal undergoes are approximately equal. Figure 1, adopted from Veenstra et al, 1981, graphically exhibits the process involved.

One pair of readings, i.e., one second worth of data, is not sufficient to allow one to judge the processes affecting the technique and the data. In order to be sure that one is observing over a sufficiently long enough period, there are several

CH2588-2/88/0000-472 $1.00 C 1988 IEEE

tests which can be done. The most significant is to test the data for white noise (Box and Jenkins, 1980) by doing a periodogram analysis. Another simple test which can be done to assure that a sufficiently large span of data has been taken is to look at the dispersion in the value as a function of the sampling interval. Veenstra et al (1981), shows that fifteen minutes of data was found sufficient to get a good average value for the time difference between the two stations. Because new hardware has been developed since those experiments, it is worthwhile to investigate how much of a span of data is now required to get a good value for time transfers. For the day which was closely scrutinized, it was found that the standard deviation of the value using 100 seconds of data was 456 ps. and that using 900 seconds of data was 480 ps. At these levels of precision, it would be ridiculous to think that there is any significance in the difference between these two values. It is obvious that there is no significant improvement in the mean value by using more than 100 seconds of data. Table 1 summarizes the results obtained.

Table - 1
Standard Deviation of Time Transfer Value on 19 February 1988 as function of number of data points used.

| Number of Points | Standard Deviation (picoseconds) |
|---|---|
| 900 | 480 |
| 800 | 485 |
| 700 | 485 |
| 600 | 494 |
| 500 | 500 |
| 400 | 502 |
| 300 | 506 |
| 200 | 467 |
| 100 | 456 |

In addition, a periodogram analysis was done in order to determine that no significant periodicities are in the data. This test assures that the data are white (random). A second order polynomial was fit to three different data sets: the data obtained at NBS and USNO and the time differences between NBS and USNO derived from this data. As an example, Figure 2 shows that there was a significant period in the data obtained at NBS on one day. Figure 3, which is typical of the time difference data, shows that there are no significant periodicities.

Figure 4 shows a 100 second span of data obtained on 16 March 1988. The residuals with regard to a linear regression are plotted as a function of fraction of a day. This was done in order to remove any slope due to the difference in frequency between the clocks being compared. From the plot of the data, it is difficult to see any structure in the residuals. It is also easy to see that the data are intrinsically sub-nanosecond type of data. None of the residuals are greater than 900 ps.

## GPS Time Transfer Hardware and Time Transfer Data

As mentioned earlier, GPS provides the most precise and accurate worldwide time synchronization service presently available. The common view technique takes full advantage of the system, minimizing the effects of certain systematic errors. The precision and accuracy achievable is a function of the amount of processing that is done to the data.

The two sites used in this experiment both have GPS Time Transfer Units collocated with the Earth Stations. Unfortunately, the receivers at each site are not identical nor are they similar in their mode of operation. The receiver at the USNO is a STel (formerly STI) 502 TTU.* It is the GPS TTU which is designated as the USNO's primary GPS TTU. Therefore, it adheres to a rather inflexible observing schedule. The receiver at the NBS Earth station is a Trimble 5000 A.* It is programmed to operate in the automatic mode. This means that it selects GPS satellites to observe for time transfer according to an internal algorithm which is weighted by satellite altitude. Thus, the two receivers do not exactly observe the same satellites at the exact same times. The only thing that the two receivers have in common is that they both observe a single satellite for 13 minutes (780 seconds).

Common view values for UTC(USNO)-UTC(NBS) were obtained for a 90-day period by combining observations of the same satellite which occurred within 10 minutes of each other at the two sites. An inspection of Figure 5 reveals a surprisingly large spread in these values. The dispersion is probably due to two factors. One arises from errors in the ionospheric model used by the single frequency GPS TTU's. Miranian (1988) reports that he observes a difference of about 30-40 ns. between the ionospheric correction measured by a dual frequency GPS TTU and that predicted by the model contained within the single frequency TTU's. Another factor probably arises from the Kalman filter parameters which are used to optimize the navigation function of the GPS System. It was also discovered that there was an error in the receiver's location at NBS. The error was approximately 0.0280 minutes of latitude.

Because of the large dispersion in the GPS data, two different approaches in reduction techniques have evolved in order to maximize the precision obtained from GPS time transfer data. One can either average the GPS data obtained from all the satellites over a span of several days, or one can average the data from one satellite obtained at the same time over several months. In both cases, calibration by some other technique is also needed for the ultimate in accuracy. In the study presented here the data are averaged over all satellites for one day. This rather simplistic approach was taken in this case because the Earth Station data was also treated simplistically and accuracy was not of concern in this preliminary study. Calibration techniques will be a part of a future study. Figure 6 shows the daily difference between UTC(USNO)-UTC(NBS) obtained by averaging the Common View GPS values over one day. The peak-to-peak spread is about 30 ns.

## Comparison and Discussion of the Two Techniques

During the period of time covered by the GPS data, estimates for the difference between UTC(USNO)-UTC(NBS) were also obtained from the two-way Earth station time transfers. Figure 7 shows the values, each point represents an average of only 100 seconds of data. Unfortunately, some data were inadvertently lost through a programming error. These data are not irretrievably lost. They can be recovered through additional processing of the archived data. But, it was not possible to reconstruct it before the deadlines of publication.

In this study, only the deviations from linear regressions are being investigated, not the actual values themselves. Here, we are only concerned with the difference in precision between the two techniques, not with their difference in actual values (accuracy). This will be the subject of a later investigation.

A linear regression was done to the data shown in Figures 6 and 7. The residuals to these linear regressions are exhibited in Figures 8 and 9. In comparing Figures 8 and 9, it should be noted that each GPS value for UTC(USNO)-UTC(NBS) is a daily average of about 12-18 points with each point being thirteen minutes of observation, while each Earth Station data point represents only 100 seconds of observation. During the interval covered by the observations, the USNO introduced a change in frequency in its time scale. Both the GPS and Earth Station observations were corrected for this change.

## Conclusions

The two-way Earth Station Time Transfer data seems to exhibit smaller fluctuations over the interval covered by the data used in this study than the GPS data. Because the signals in the case of the Earth Station data are reciprocal and travel through the same atmosphere, we expect that the majority of the propagation delays will cancel each other except for the differential effects due to the use of different frequencies on

the up and down links. For the GPS data, the signals from the satellite travel through different atmospheres to the observing stations, therefore the delays are not reciprocal. Furthermore, the GPS timing units are single frequency receivers. They assume a model for the atmospheric propagation path delays. Other evidence indicates that there are some deficiencies in the assumed model. Therefore,it is thought that the Earth Station data should be inherently smoother than the GPS data in showing the difference between the two time scales. These preliminary data tend to support this thesis.

While the two-way time transfer technique using Earth Stations seems to be more precise than Common View GPS, it should be pointed out that the former technique is not universally practical. The two-way technique requires that the practitioners be able to receive and transmit. The latter requires a license which may be difficult to obtain. In addition, the costs involved with operating an Earth Station are considerably greater than in operating a GPS TTU. Therefore, we expect that the two-way technique will only be used by a few laboratories which require the highest precision attainable for the comparison of state-of-the-art oscillators and time scale algorithms.

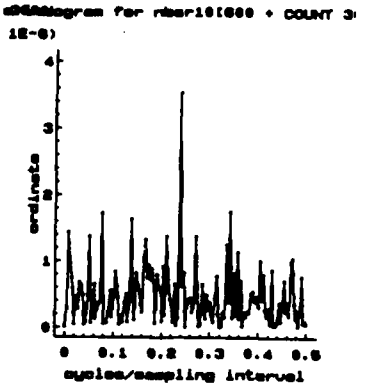



Figure 2 - Periodogram Analysis of NBS data on 19 February 1988.

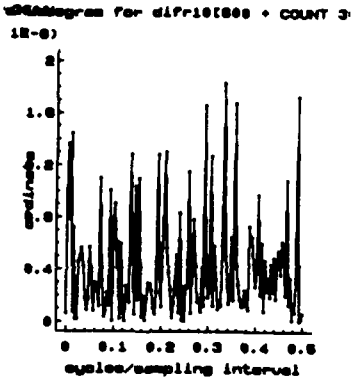



Figure 3 - Periodogram Analysis of UTC(NBS)-UTC(NBS) on 19 February 1988.

## References


Allan, D. W.. and Weiss, M. A., May, 1985, "Accurate Time and Frequency Transfer During Common View of a GPS Satellite", Proceedings of the 34th Annual Frequency Control Symposium.

Box, G. E. and Jenkins, G. M., "Time Series Analysis; Forecasting and Control", Holden-Day, San Francisco, 1970

Hartl, P., Veenstra, L., Gieschen, N., Wende, D. M., Klepczynski, W. J., Nau, H., and Stoiber, R., "Spread Spectrum Time Transfer Experiment via INTELSAT", Proeedings of 15th Annual Precise Time and Time Interval Applications and Planning Meeting, December 6-8, 1983, Washington, D.C.

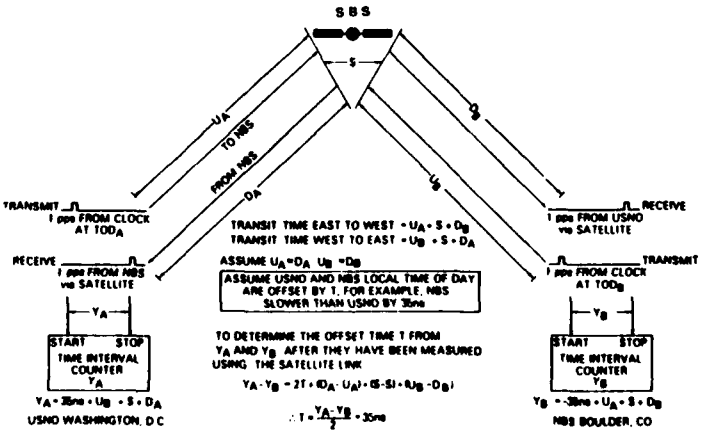
Klepczynski, W. J., October, 1983, "Modern Navigation Systems and Their Relation to Timekeeping", Proceedings of the IEEE, Vol. 71, No. 10, 1193-1195.

Miranian, M., 1988, Private Communication

Steele, J. McA., Markowitz, W., and Lidback, C. A., December 1964, Telestar time synchronization, IEEE Trans. Instr. and Meas., Vol. IM-13, 4, 164-170.

Veenstra, L., Kaiser, J., Costain, C., Klepczynski, W. J. and Allan, D., "Frequency and Time Coordination via Satellite", Comsat Technical Review, Vol 11, No. 2, Fall 1981, 369-402.


* "The use of trade names does not imply endorsement."




**Time Transfer via Satellite**

Figure 1 - Satellite and Earth Station Geometry for Time Transfer

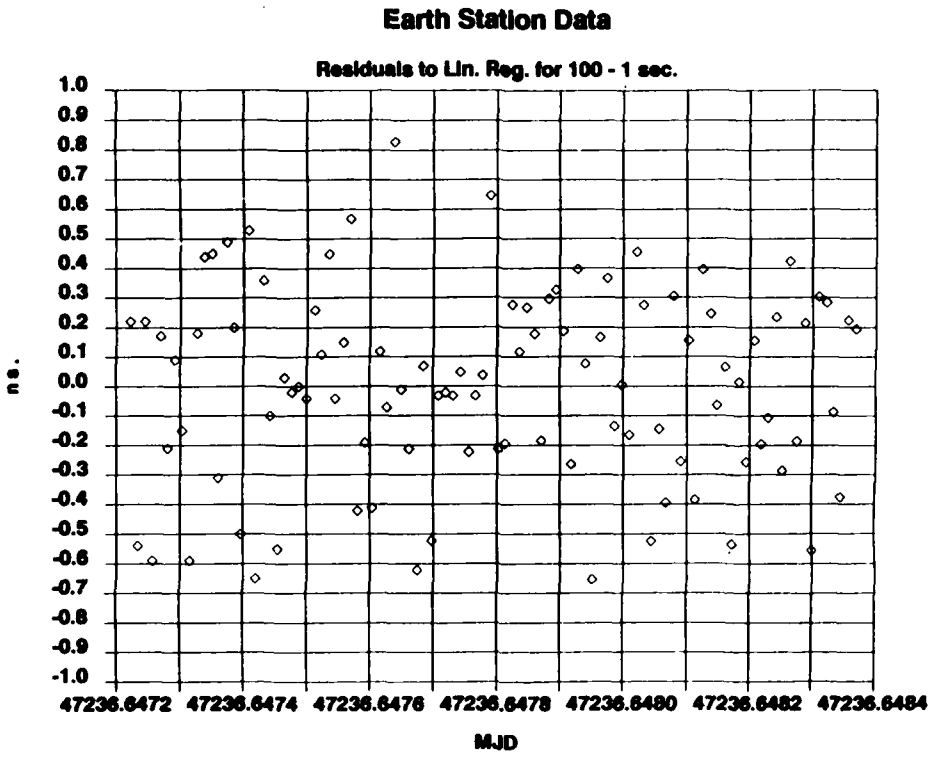



Figure 4 - 100 second span of data on 16 March 1988.

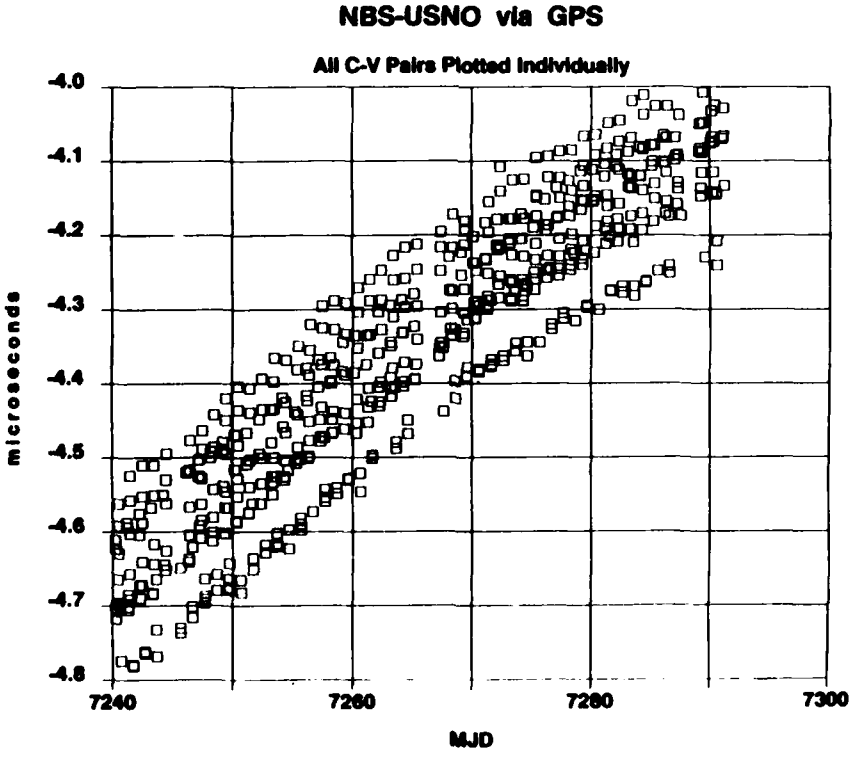



Figure 5 - UTC(USNO)-UTC(NBS) *via* GPS Common View.

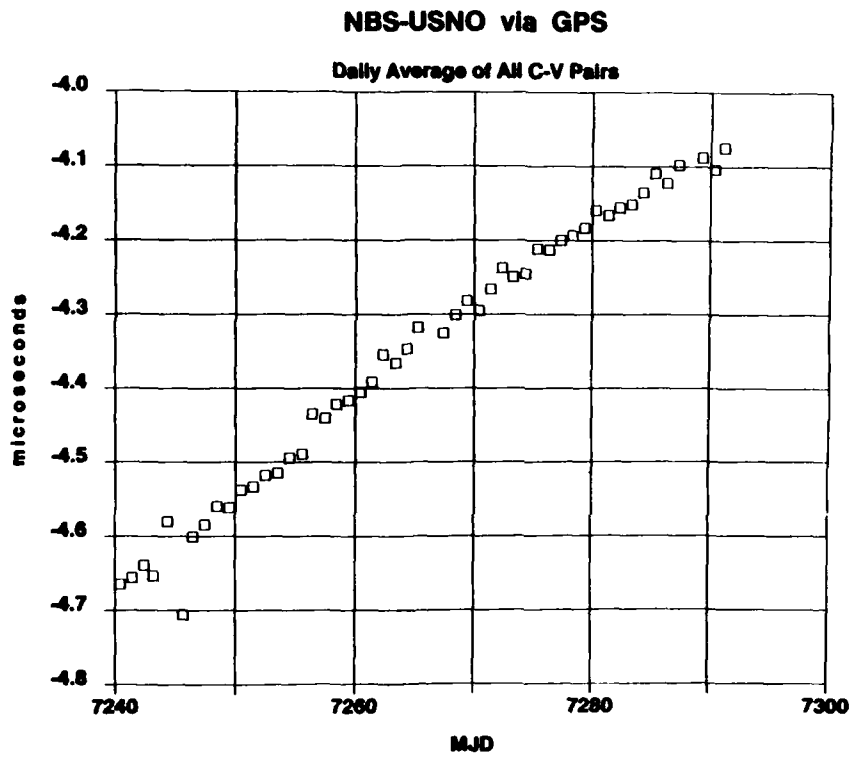



Figure 6 - Averaged Values of UTC(USNO)-UTC(NBS) obtained via GPS Common View.

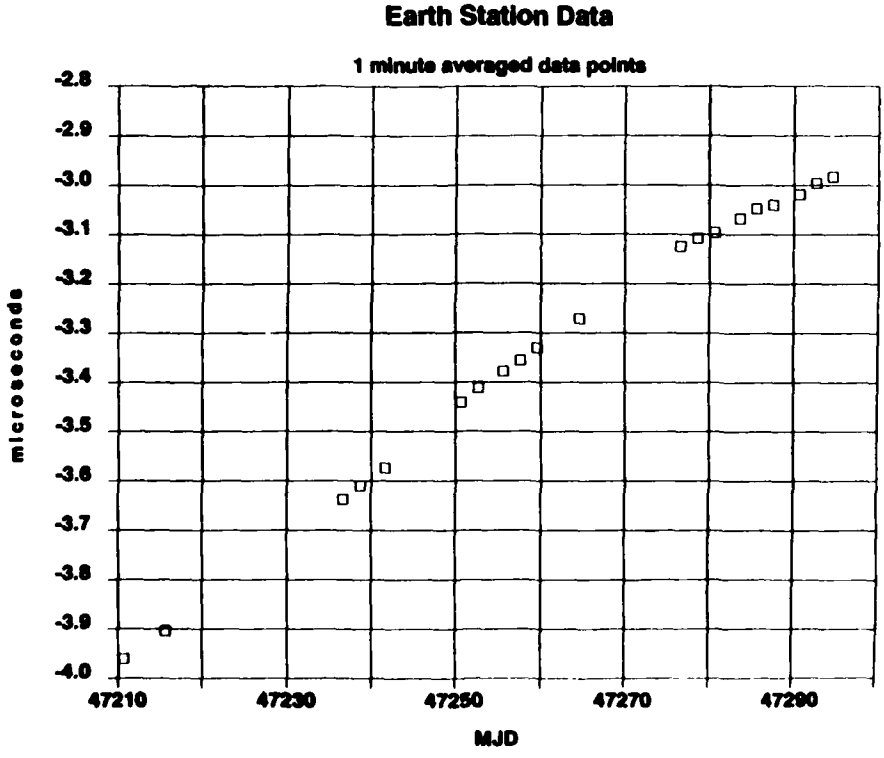



Figure 7 - UTC(USNO)-UTC(NBS) obtained via Two-Way Earth Station Time Transfers.

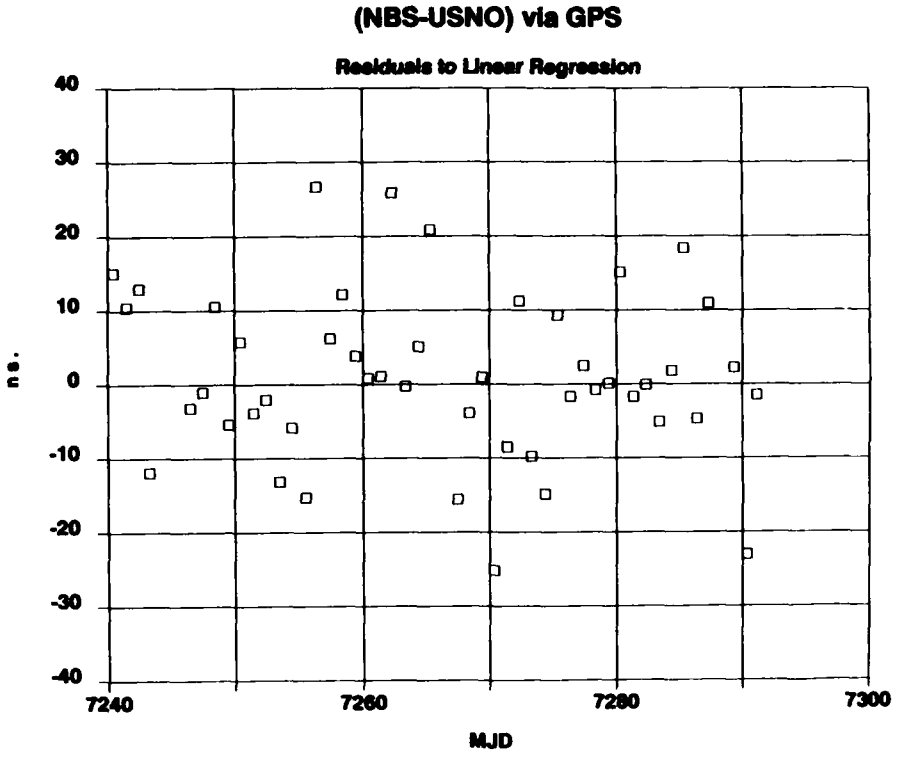



Figure 8 - Residuals to a Linear Regression of the Averaged Values of UTC(USNO)-UTC(NBS) obtained via GPS Common View.

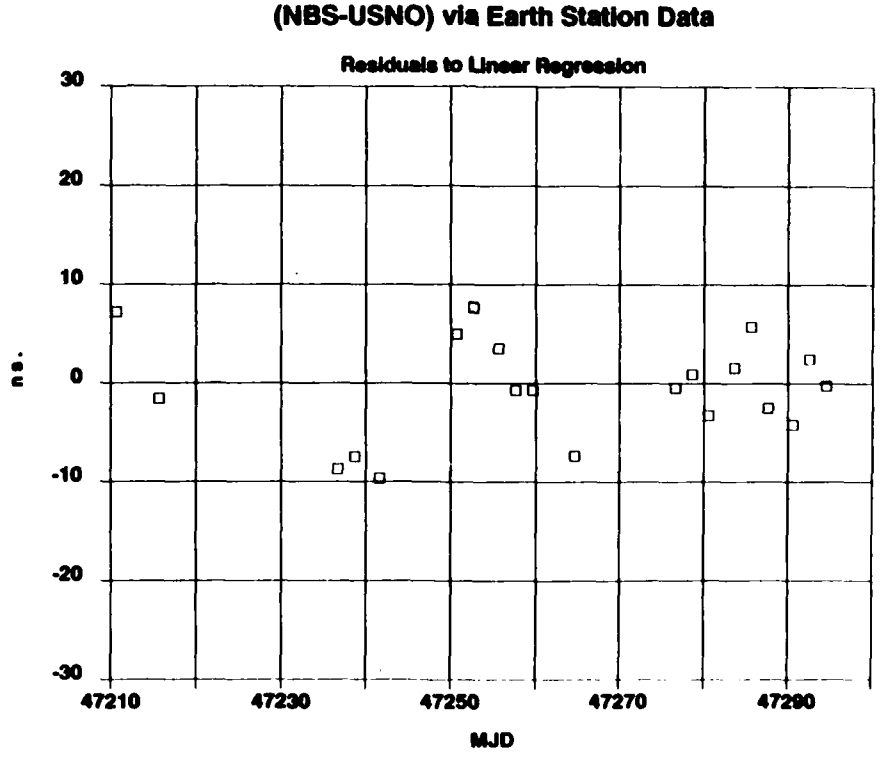



Figure 9 - Residuals to a Linear Regression of the values of UTC(USNO)-UTC(NBS) obtained via Two-Way Earth Station Time Transfers.

42nd Annual Frequency Control Symposium - 1988

# FIBER OPTIC FREQUENCY TRANSFER LINK

L. Primas, G. Lutes and R. Sydnor
Jet Propulsion Laboratory, Pasadena, Ca.

## Abstract

A stabilized fiber optic reference frequency distribution system has been fabricated and tested. The distribution system is designed to transmit a 100 MHz reference signal, which is generated by a hydrogen maser frequency standard, over a distance of 22 km. The goal is to maintain a stability of 1 part in $10^{17}$ for 1000 seconds averaging times. The stabilizer circuitry reduces delay variations that result from environmental changes on the fiber such as temperature. The stabilizer does the phase correction from one end of the link and maintains a constant phase relationship at the far end of the link. The distribution system was designed using a computer simlation program that allows the delays, bandwidths, gains and damping factors to be varied for optimum performance. A bread board version of the stabilizer was fabricated and tested in an environmentally controlled test chamber. In preliminary tests, the stabilizer reduced phase variatons caused by temperature changes of 20 $^{o}$ C by as much as 45 times.

## Introduction

State-of-the-art frequency standards, such as hydrogen masers, generate stable reference frequencies in support of precision measurements made in the NASA/JPL Deep Space Network (DSN). DSN applications of frequency standards include support of unmanned space projects, flight radio science, radio and radar astronomy, very long baseline interferometry, geodynamic measurements and gravitation wave detection[1].

The distribution of reference frequencies to multiple remote users in the DSN is accomplished through a frequency distribution system. A frequency distribution system is required to distribute the reference frequencies over distances as great as 30 km from the frequency standard to the subsystems that use them. The distribution system must also have minimal impact on the stability of the transmitted reference frequency[2].

The stability of present frequency standards has an Allan variance on the order of $8 \times 10^{-16}$ for 1000 seconds averaging times. Researchers expect future frequency standards to be improved by an order of magnitude over this value.

The degradation of the distributed reference frequency is primarily due to the variations in the group delay in the transmission medium[2]. A constant rate of change in group delay (D) adds a constant frequency offset $\Delta f$, but does not degrade the frequency stability. If the rate of change of group delay is not constant the frequency stability is degraded by,

$$\frac{d(\Delta f)}{dt} = f\frac{d^2D}{dt^2} \qquad (1)$$

where $f$ is the transmitted frequency. Temperature change is the primary cause of group delay variations in a cable. For a step change in temperature ($\Delta T$), the change in frequency offset $\Delta f$ is related to the temperature coefficient of delay ($\alpha$) and time constant of the fiber ($\tau$) by,

$$\frac{d(\Delta f)}{dt} = \frac{-\alpha \Delta T e^{-\frac{t}{\tau}}}{\tau^2} \qquad (2)$$

From equation (2) we see that the rate of change of frequency offset is decreased by decreasing the temperature change or increasing the time constant of the cable.

Because of the small temperature coefficient of delay and the low loss of optical fibers, they are the most practical medium for transmitting reference frequencies over distances longer than a few meters[3]. Optical cables used in the DSN are buried underground to decrease the temperature changes on the cable and to increase the time constant of the cable. Electronic control systems are used to further reduce group delay variations.

There is a continuing program at the Jet Propulsion Laboratory to develop fiber optic frequency distribution systems for stable reference frequency distribution. The subject of this paper is an electronic control system used to reduce group delay variations in a fiber optic cable.

## Method of Stabilization By Phase Conjugation

The stabilized fiber optic distribution system described in this paper uses the phase conjugation method of stabilization. This method is used because it does not require a variable delay device in the two way signal path, as required by other methods. The variable delay device must have a range of delays equal to the group delay variation to be reduced and must have exactly the same phase delay in both directions at all times. Devices meeting these requirements are difficult to implement, particularly in optical fiber.

In the conjugation method (Figure 1), as in most cable stabilization methods, the signal propagates through the optical fiber in both directions. Therefore, the midpoint of the round trip signal path is at the far end of the cable. The system maintains a conjugate (antisymetric) relationship to the reference between the forward signal and the reverse signal at the input to the fiber.

The phase modulo $2\pi(\theta_o)$ at the far end of the cable is,

$$\theta_m = (\theta_o - \theta_1) + \frac{(\theta_o + \theta_1) - (\theta_o - \theta_1)}{2} = \theta_o \qquad (3)$$

CH2588-2/88/0000-478 $1.00 C 1988 IEEE

where:

$\theta_o$ = the reference phase modulo $2\pi$, and

$\theta_1$ = the delay phase modulo $2\pi$.

With the conjugate relationship satisfied, the phase at the far end of the cable is the same as the reference phase and the two way link is stabilized.

## System Description

The stabilized fiber optic distribution system is electronically controlled and uses the conjugation method to maintain frequency stability. The distribution system consists of a reference unit, located at the reference frequency source, and a remote unit, located at the site receiving the reference frequency (Figure 2).

The reference unit consists of the phase conjugator, a fiber optic transmitter, a fiber optic receiver, a phase lock loop (PLL) and a fiber optic coupler. The phase conjugator compares the phase at the transmitter and receiver of the reference unit and uses a voltage controlled oscillator (VCO) to maintain a constant phase at the remote unit. The conjugator requires a 100 MHz reference signal and a 20 MHz auxillary signal. A previous phase conjugator design used a single 100 MHz reference signal, but required two precisely matched phase detectors and tightly controlled signal levels[4]. The latest design is much easier to implement because a single phase detector is used to measure the phase error.

Figure 3 shows a block diagram of the stabilized fiber optic distribution system. The phase conjugator contains three mixers (M1,M2, and M3), two RF power splitters (S1,S2), two band pass filters (BPF), a phase detector (PD), a voltage controlled oscillator (VCO), and an inner loop filter (ILF).

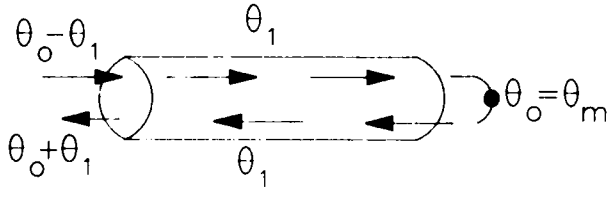



Figure 1. Stabilization by Phase Conjugation

A synthesizer supplies a 100 MHz signal and a 20 MHz signal to the first mixer (M1) which multiplies the two signals together to produce 80 MHz and 120 MHz signals. A power splitter (S1) separates the signal out of the mixer (M1) into two signal paths. Filters in each signal path pass only one frequency so that 120 MHz is the output from one signal path and 80 MHz is the output from the other signal path.

A second mixer (M2) multiplies the 80 MHz signal and the 100 MHz signal from the VCO to produce a 20 MHz intermediate (IF) signal. The 20 MHz IF signal contains the instantaneous phase difference between the VCO signal and the 80 MHz reference.

Similarly, a mixer (M3) multiplies the 120 MHz signal and the 100 MHz signal coming from the remote unit to produce another 20 MHz IF. This 20 MHz IF signal contains the instantaneous phase difference between the return signal (100 MHz) and the 120 MHz reference signal.

The phase detector (PD) receives the two 20 MHz IF signals and produces a voltage that is proportional to the phase difference between them. The voltage is applied to the VCO error input through the inner loop filter (ILF). Delay changes in the fiber optic cable result in changes in the control voltage. This voltage thus controls the phase of the VCO relative to the original 100 MHz reference signal

An RF power splitter (S2) divides the output of the VCO into two signals. Mixer M2 receives one of the signals while the other signal modulates the optical carrier emitted from the laser transmitter.

The modulated optical signal passes through a two way optical coupler into the fiber. A 50/50 mirror at the remote unit reflects half of the optical signal back toward the reference unit. The other half of the optical signal passes through the mirror to an optical receiver. The receiver demodulates the optical signal and amplifies the resulting 100 MHz RF signal. A PLL filters the signal which is used as a remote reference frequency. The reflected optical signal returns to the reference unit where it passes through the optical coupler and is detected by another optical receiver and filtered by another PLL. Mixer M3 receives the resulting 100 MHz signal. With the signal back at the reference unit, the system loop is closed.

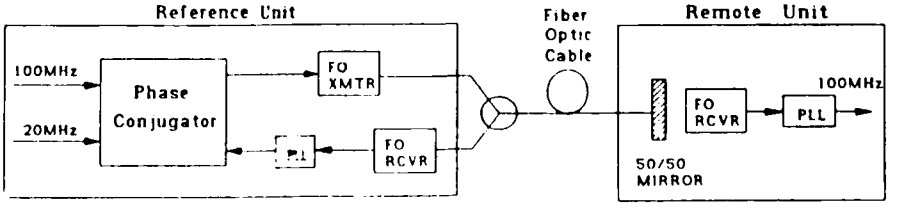



Figure 2. Stabilized Fiber Optic Distribution System

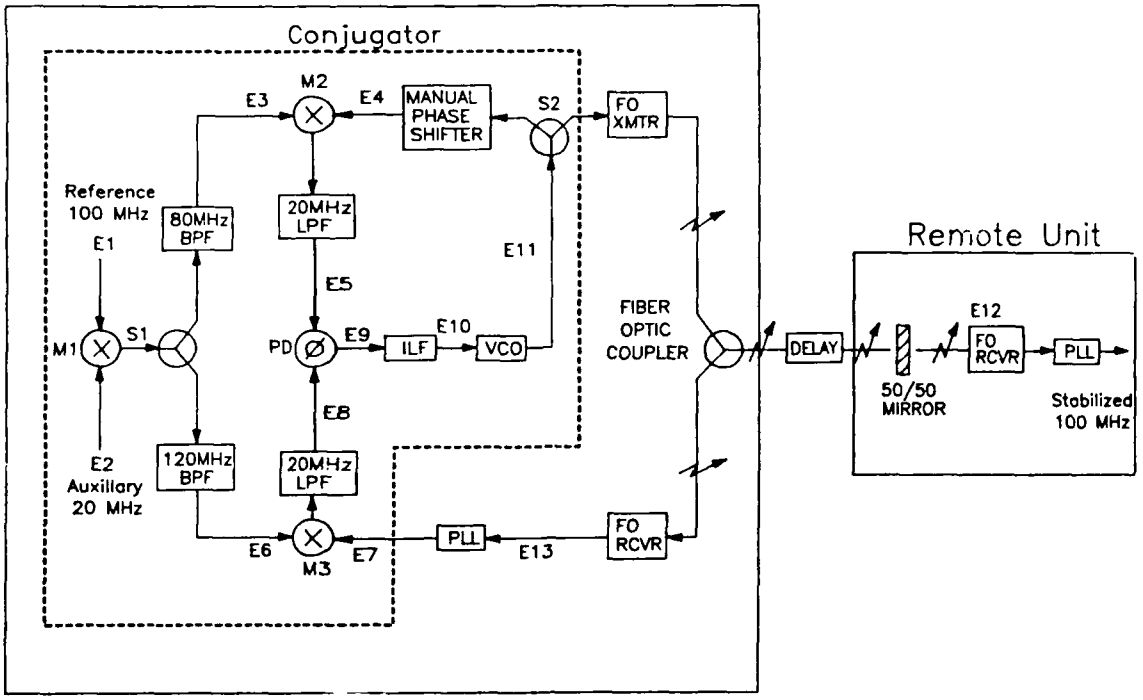



Figure 3. Block Diagram of Stabilized Fiber Optic Distribution System

## Static Analysis

The steady state operation of the stabilized distribution system was analyzed by determining the phase at various points in the system (Figure 3). All signals were derived from the 100 MHz reference signal. The result of the analysis shows that the remote phase is independent of the delay in the fiber optic cable.
Let,

$$E_1 = A_1 \sin(\omega t) = \text{the 100 MHz reference, and} \qquad (4)$$

$$E_2 = A_2 \sin\left(\frac{\omega}{5} t + \theta_{20}\right) = \text{the 20 MHz reference.} \qquad (5)$$

Mixer M1 multiplies $E_1$ and $E_2$. The result after band pass filtering is,

$$E_3 = A_3 \cos\left(\frac{4}{5}\omega t - \theta_{20}\right), \text{ and} \qquad (6)$$

$$E_6 = -A_6 \cos\left(\frac{6}{5}\omega t + \theta_{20}\right). \qquad (7)$$

Let,

$$E_4 = A_4 \sin(\omega t + \theta_v) \qquad (8)$$

where, ignoring for the moment the manual phase shifter,

$A_4 \sin(\omega t)$ = the 100 MHz signal from the VCO, and

$\theta_v$ = the correction phase added by the VCO.

Mixer M2 multiplies $E_3$ and $E_4$, which after low pass filtering results in,

$$E_5 = A_5 \sin\left(\frac{\omega}{5} t + \theta_{20} + \theta_v\right). \qquad (9)$$

The 100 MHz signal returning from the remote unit is,

$$E_7 = A_7 \sin(\omega t + \theta_v + 2\theta_d), \qquad (10)$$

where,

$\theta_d$ = the one way phase delay in the fiber.

Assuming identical delays in both directions, the total phase delay is $2\theta_d$. Mixer M3 multiplies $E_6$ and $E_7$, which after low pass filtering results in,

$$E_8 = A_8 \sin\left(\frac{\omega}{5} t - 2\theta_d - \theta_v + \theta_{20}\right) \qquad (11)$$

The phase detector (PD) compares the phases of $E_5$ and $E_8$ resulting in,

$$E_9 = A_9 \cos(2\theta_d + 2\theta_v). \qquad (12)$$

The signal at the remote unit is,

$$E_{12} = A_{12} \sin(\omega t + \theta_v + \theta_d). \qquad (13)$$

When the control voltage on the VCO is zero (steady state condition), the phases at the input of mixers M2 and M3 are conjugate and the phase at the remote unit is independent of delay in the two way signal path. In equation (12), the voltage on the VCO is zero when,

$$2\theta_d + 2\theta_v = \frac{\pi}{2}. \qquad (14)$$

In order to meet the condition in (14), the VCO must add the phase,

$$\theta_v = \frac{\pi}{4} - \theta_d. \qquad (15)$$

Combining equations (8), (9), and (15) gives,

$$E_4 = A_4 \sin\left(\omega t + \frac{\pi}{4} - \theta_d\right), \text{ and} \quad (16)$$

$$E_7 = A_7 \sin\left(\omega t + \frac{\pi}{4} + \theta_d\right). \quad (17)$$

Equations (16) and (17) show that signals $E_4$ and $E_7$ are conjugate in $\theta_d$. Combining equations (13) and (15) gives,

$$E_{12} = A_{12} \sin\left(\omega t + \frac{\pi}{4}\right). \quad (18)$$

Thus, the phase of the output signal is offset from the phase of the reference signal by $\pi/4$, but is completely independent of the delay in the two way portion of the signal path.

## Dynamic Analysis

The dynamic analysis shows the closed loop response and the error response of the system[5], taking into account the various gains, bandwidths and damping factors in the system (Figure 3).

The transfer functions for mixers M2 and M3 are,

MIXER M2 $$\theta_5(s) = k_m\left[\theta_4(s) - \theta_i(s)\right] \quad (19)$$

MIXER M3 $$\theta_8(s) = k_m\left[\theta_i(s) - \theta_7(s)\right] \quad (20)$$

where $k_m$ is the gain of the mixers. The transfer function of the phase detector is,

$$V_9(s) = k_d\left[\theta_5(s) - \theta_8(s)\right] \quad (21)$$

where $k_d$ is the gain of the phase detector, and
the ILF transfer function is,

$$V_{10}(s) = G_1 F(s) V_9(s) \quad (22)$$

where,

$G_1$ = additional gain of ILF, and

$$F(s) = \frac{s\tau_2 + 1}{s\tau_1}. \quad (23)$$

In equation (23), $\tau_2$ and $\tau_1$ determine the bandwidth of the ILF. Also, the transfer function of the VCO is,

$$\theta_v(s) = \frac{k_o V_{10}(s)}{s} \quad (24)$$

where $k_o$ is the gain of the VCO, and the transfer function of the phase lock loop is[6],

$$\theta_7(s) = H(s)\theta_{13}(s) \quad (25)$$

where,

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (26)$$

In equation (26), $\omega_n$ and $\zeta$ determine the bandwidth of the PLL. The term $e^{-ks}$ represents the time delay due to the change in length of the cable. Therefore the phase delay from the fiber optic transmitter to the mirror is,

$$\theta_{out}(s) = e^{-ks}\theta_v(s), \quad (27)$$

where k is the delay in seconds. The phase delay from the mirror to the PLL is,

$$\theta_{13}(s) = e^{-ks}\theta_{out}(s). \quad (28)$$

A manual phase shifter is used to keep the delay from mixer M2 to the fiber optic coupler and the delay from mixer M3 to the coupler the same. Thus delays added by the fiber optic transmitter, receiver, PLL and any other delays in the two paths are matched.

When equations (19) through (28) are combined, the closed loop transfer function becomes,

$$\frac{\theta_{out}(s)}{\theta_i(s)} = \frac{2k_d k_m k_o F(s)}{se^{ks} + k_d k_m k_o G_1 F(s)\left[e^{ks} + H(s)e^{-ks}\right]} \quad (29)$$

The error transfer function is calculated from the closed loop response as follows:

$$\frac{\theta_i(s) - \theta_{out}(s)}{\theta_i(s)} = \frac{\theta_{error}(s)}{\theta_i(s)} = 1 - \frac{\theta_{out}(s)}{\theta_i(s)}. \quad (30)$$

Equations (29) and (30) were evaluated using LOTUS 1-2-3. The system stability was examined by determining the frequency response of the closed loop transfer function and the error transfer function as various parameters were varied. The gains of the mixers, phase detector and VCO were determined by testing the components, but the bandwidths, damping factors and additional gains were varied for optimum system design. Damping factors were varied from 0.7 to 1.4. The bandwidths of the inner loop and the PLLs were determined from the spectral noise characteristics of the reference frequency (hydrogen maser), the VCO and the fiber optic link. Different delays could also be examined.

The PLL in the reference unit cleans up the signal and maintains a high signal to noise ratio and a constant amplitude into the phase detector. Its bandwidth is determined by the intersection of the spectral noise of the VCO and the fiber optic link (Figure 4). This allows the high fiber optic link stability to be attained for low offset frequencies and the low noise of the VCO at high offset frequencies.

The 7 dB bandwidth of the PLL is related to the cross over frequency by,

$$BW = 2\beta_l = 3.2 f_c \quad (31)$$

where,

$f_c$ = cross over frequency

$\beta_l$ = single sided noise bandwidth

$BW$ = double sided bandwidth.

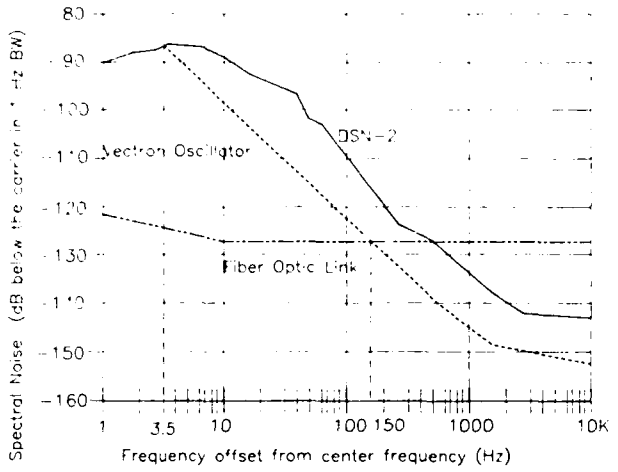



Figure 4. Plots of Spectral Noise of a Maser, VCO and Fiber Optic Link Verses Frequency

The optimum inner loop bandwidth is dependent on the noise spectrum of the reference frequency to be transmitted. Two conditions determine the bandwidth of the inner loop. First, the inner loop bandwidth must be much smaller (50 times) than the bandwidth of the PLL in the reference unit for system stability. Secondly, the inner loop bandwidth must be wider than typical variations in the fiber (10 min.). There is also a PLL at the remote unit to clean up the signal, but the signal out of this PLL is not returned to the reference unit and thus does not affect the system stability. The bandwidth chose for the PLL in the remote unit is dependent on the spectral noise of the reference source, but will be approximately the same as the inner loop bandwidth.

The LOTUS 1-2-3 program was used to evaluate the system with an inner loop bandwidth of 11 Hz and a PLL bandwidth of 475 Hz at the reference unit. The closed loop and error responses obtained with these bandwidths are shown in Figure 5.

From this analysis, it is theoretically possible to reduce phase variations at the remote unit by 124 dB at 1000 seconds averaging time. Such a factor corresponds to a phase reduction of approximately $10^6$ times.

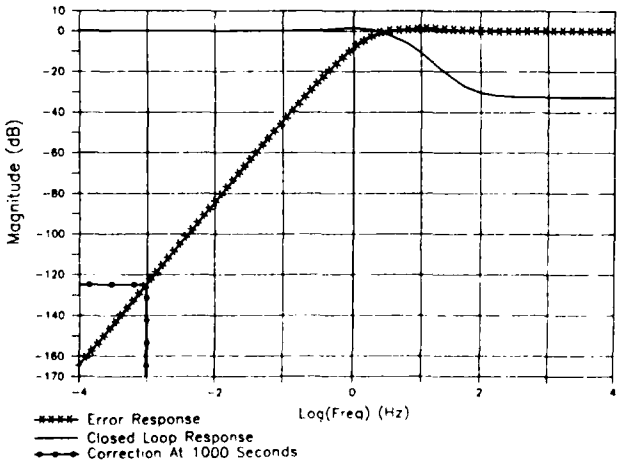



Figure 5. Plots of Closed Loop and Error Responses

## Measurements

Preliminary tests were performed on a 1 km length of fiber optic cable containing 4 fibers connected in series for a total length of 4 km to determine its temperature coefficient of delay. The cable was wound on a test rack that allowed circulation of air (Figure 6). The test rack was placed in an environmentally controlled test chamber, where temperature, pressure and humidity can be varied.

A 100 MHz signal was transmitted from a hydrogen maser through the fiber while the phase between the transmitted end and the received end were monitored. The temperature was then stepped from 15 to 35 °C (Figure 7).

In this arrangement, the phase between the transmitted end and the received end of the fiber optic cable changed 89 degrees over 9 hours. The measured temperature coefficient of delay for the fiber was 6.49 ppm/°C.

A breadboard stabilizer (figure 8) was assembled and tested using the parameters obtained from the dynamic analysis.

Tests were performed on the stabilizer by varying the temperature of the fiber and monitoring the signal phase across the link. The system was initialized by using

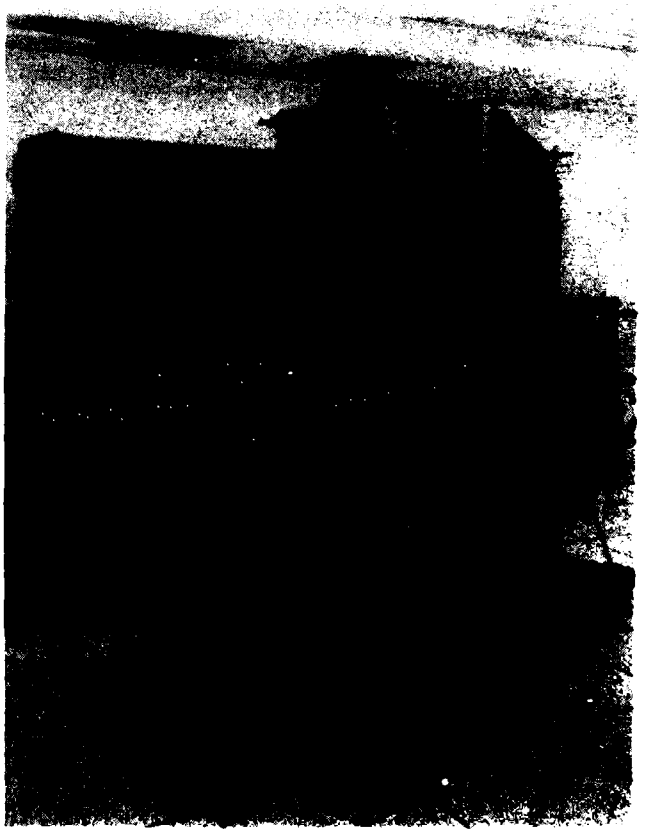


Figure 6. Fiber Optic Test Rack

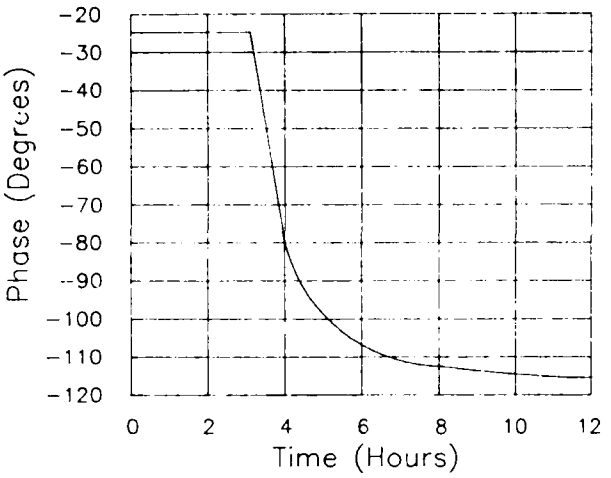



Figure 7. Phase Shift Across 4 Km Fiber for 20 C° Temperature Change

manual phase shifters to compensate for phase delays added by the the fiber optic transmitter, receivers, PLLs and other delays in the system. Figures 9 shows the phase difference in the reference unit between the reference and the transmitted signals and the phase difference between the reference and received signals after the ruond trip signal path. As shown in Figure 9 the transmitted and received signals at the reference unit are conjugate around 22 degrees. The phase difference between the signal at the receiver of the remote unit and the reference signal was also measured, (Figure 10). Also shown in Figure 10 is the conjugation error. Phase variations at the fiber optic transmitter and the receiver in the reference unit were approximately 90 degrees, while the phase at the remote unit varied only 2 degrees, for an overall correction of 45 times. By comparing the two graphs in Figure 11, we see that the phase variation at the remote unit is probably due to the conjugation error at the start of the test.

RF leakage and poor signal to noise ratio in the breadboard system appears to limit the correction factor to a value smaller than the theoretical limit. Additional power supply filtering has been added to the modules in the system and optical losses have been reduced. Further testing will determine the level of improvement achieved.

The amount of correction needed can be determined from Figure 11. This figure shows stability curves for a typical hydrogen maser, a 14 km fiber optic link, and an estimated plot of a 29 km link needed at the Goldstone Deep Space Communications Complex. Also shown is the stability limit imposed by the signal to noise ratio of the present fiber optic link.

The figure shows that a correction of 20 times is sufficient to reduce the link stability to the level imposed by the signal to noise ratio or to a level 10 times better than the hydrogen maser.

## Conclusion

A fiber optic cable stabilizer system has been designed, breadboarded and tested. The stabilizer is based on a conjugation system, and operates by comparing the input phase to the phase of the round trip returned signal. The performance of the stabilizer was analyzed, and based on the analysis a breadboard model was fabricated and tested. The test results indicate that the system has achieved adequate group delay reduction to be suitable for use in the DSN. Several needed improvements have also been identified. The system analysis also indicates that this approach will be adequate to meet the distribution requirements of future frequency standards.

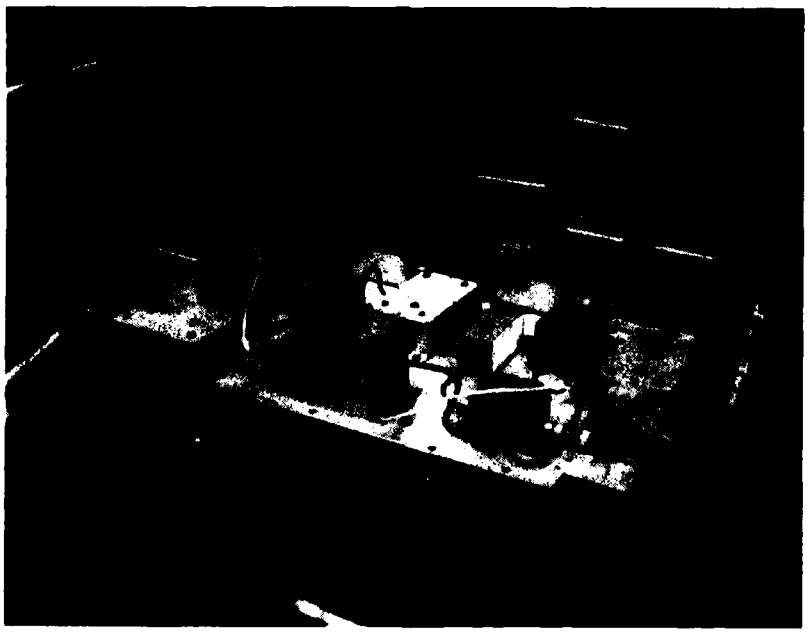


Figure 8. Breadboard Stabilizer

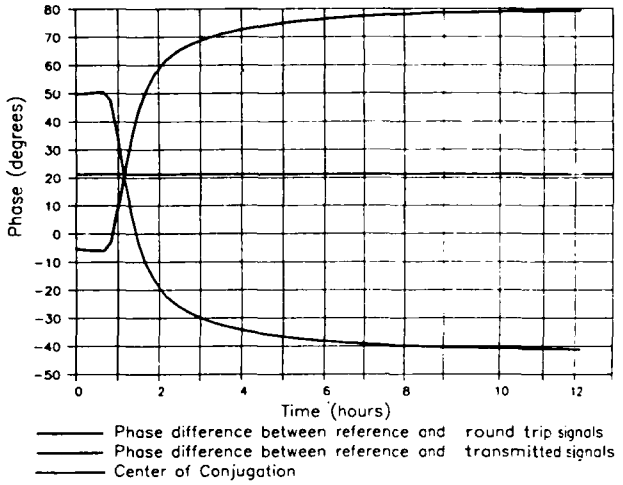



Figure 9. Phases at Mixers M2 and M3 in Reference Unit

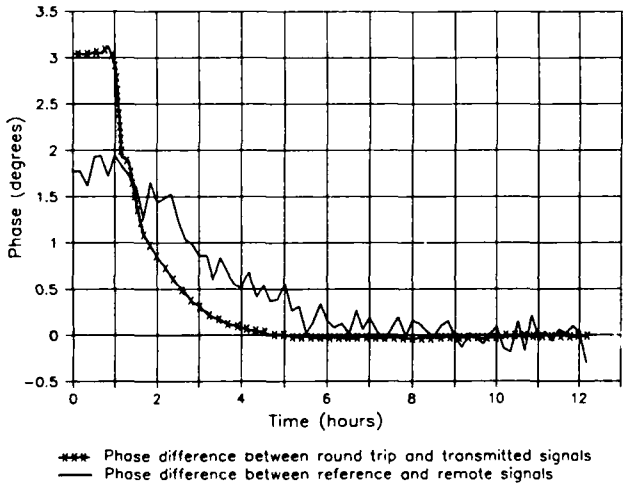



Figure 10. Phase Conjugation and Remote Unit Phase


## Acknowledgment

**This work represents the results of one phase of research carred out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.**

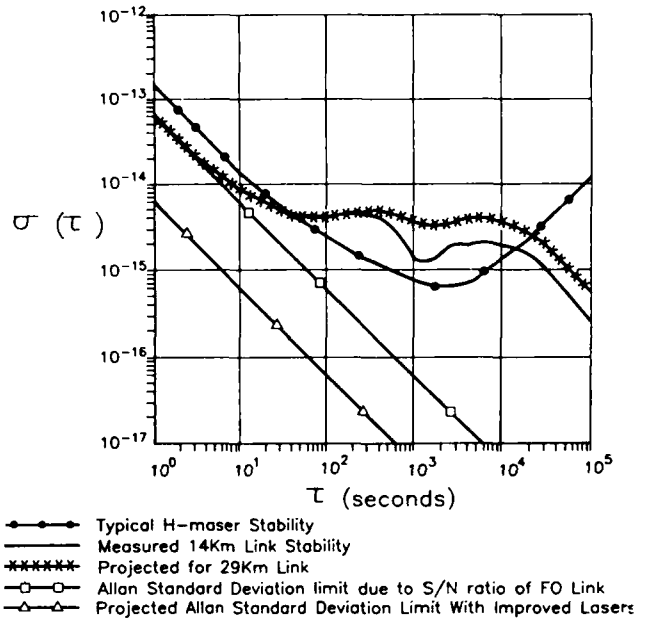




Figure 11. Frequency Stability Correction Required

## References


[1] JPL publication 400-333 1/88, "The Deep Space Network."

[2] G. Lutes, "Reference Frequency Distribution Over Optical Fibers: A Progress Report," Proc. 41st Ann. Frequency Control Symposium, pp. 161-166, May 27-29, 1987.

[3] G.F. Lutes, "Experimental Optical Fiber Communications Link," The Telecommunications and Data Acquisition Progress Report 42-59, pp. 77-85, Jet Propulsion Laboratory, July-August 1980.

[4] G. Lutes, "Optical Fibers For The Distribution Of Frequency And Timing References," Proc. 12th Ann. Precise Time and Time Interval (PTTI) Applications and Planning Meeting, pp 663-680, NASA Conference Publication 2175, Goddard Space Flight Center, Dec. 1980.

[5] A. P. Sage, Linear Systems Control, Matrix Publishers, Inc., 1978.

[6] F. M. Gardner, Phaselock Techniques, John Wiley & Sons, Inc., 1966.

42nd Annual Frequency Control Symposium - 1988

# ISRAEL'S NEW SYNCHRONIZED TIME SCALE, UTC(INPL)

A. Shenhar and W. Litman
National Physical Laboratory of Israel
Jerusalem 91904, Israel

A. Lepek† and A. Citrinovitch
Time and Frequency Limited
Holon 58117, Israel

D.W. Allan and T.K. Peppler
National Bureau of Standards
Boulder, CO 80303 USA

## Summary

The National Physical Laboratory of Israel (INPL) together with Time and Frequency Limited (TFL) and NBS is building a software clock to be used as the Israeli national time base, UTC(INPL). The software clock is based on several commercial Cs clocks (HP and TFL) whose outputs are routed sequentially through a TFL programmed switch into a time interval counter. The phase differences are processed in a personal computer using a procedure adopted from NBS to generate the software clock. The system also has an input from a common view GPS receiver for time comparisons. Steering of the software clock is possible from the PC keyboard.

The system has the capability to compare any additional clock (such as units under calibration) against the software clock, thus upgrading the accuracy of calibration.

This paper presents the principles of generating the software clock, its performance and the method to compare it to UTC and UTC(NBS). The paper also presents the optimization procedures for synchronizing and syntonizing UTC(INPL) with UTC as a coordinated time scale.

## Introduction

The UTC(INPL) is generated presently from one master Cs clock, HP-5061A* option 004 which is backed-up by two similar clocks and one TFL-5440A. The clocks are temperature stabilized within ± 0.5 °C. The system is equipped as well with two common view GPS receivers and the feedback for comparisons is made through the Bureau International des Poids et Mesures(BIPM) Circular-T. The stability over 10 days is better than 1e-13. The software clock system, which was installed in April 1988, runs now in parallel and upon completion of the run-in phase will be declared as the new UTC(INPL) generating system.

The software clock system generates its time base from the above four clocks by using an algorithm which filters the weighted average outputs of each clock. The outputs of the system, which are obtained in real time, are the time and frequency differences of each of the ensemble clocks from the software clock.

The advantages of using the software clock system over the previous one are listed below (some of these are not yet fully implemented): upgrades stability by ensemble averaging; filters individual clock's anomalies; takes care automatically of a failed clock, thus increasing reliability. Using a software clock enables additional functions: steering to UTC, comparison to other clocks, lock of a hardware clock onto the system and automated comparisons using common view GPS.

## Hardware

The guidelines in designing the system were high flexibility for future upgrades in a way that uses as much as possible existing instrumentation in a time and frequency lab.

Figure 1 is a block diagram of the software clock and Figure 2 is a photo of the complete UTC(INPL) system. The only additional piece of equipment that was custom designed is the switch matrix in Figure 1. The mode of operation is as follows: The phase differences of each pair of clocks in the ensemble are measured sequentially using the time interval counter. The divider is used to divide by two and generate a lower frequency at the "start" input of the counter to eliminate phase modulo ambiguity. The construction of the switches enables one to feed any clock at both inputs to the counter, thus making possible measurements such as system jitter and delays. Figure 3 is a short summary of the hardware features.

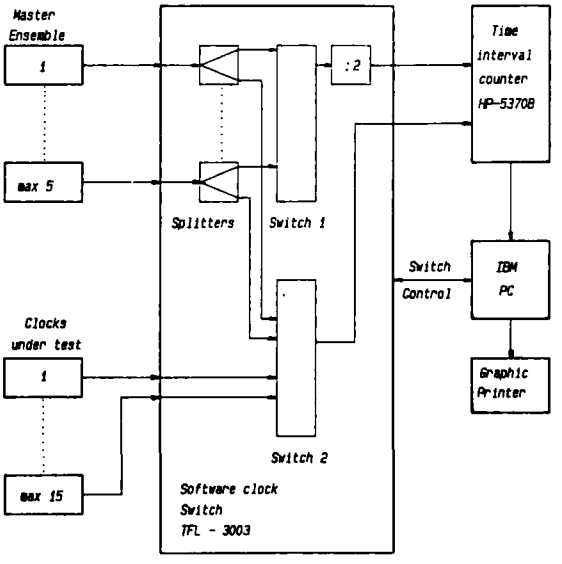



Fig. 1 A Block Diagram of the Software Block System.

†A. Lepek is a consultant to INPL.

Contribution of the U.S. Government, not subject to copyright.

*Identification of a commercial company does not imply recommendation or endorsement by the National Bureau of Standards, nor does it imply that any identified entity is the only or the best available for the purpose.

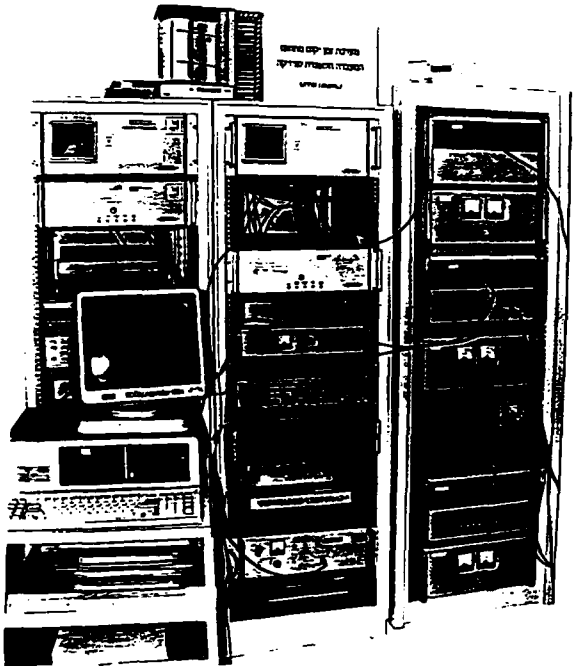


Fig. 2. The Complete Set-up of the UTC(INPL) System.

| | |
|---|---|
| Maximum ensemble clocks | 5 |
| Maximum clocks under test | 15 |
| Isolation between inputs | 120 dB |
| Ensemble clocks' frequency | 1 pps |
| Frequencies of tested units | 10 MHz and sub-harmonics down to 1 Hz |
| Switching time | 1 ms |
| Jitter | less then 30 ps |
| Computer synchronization output | by the counter's |
| Information storage | 20 Mb hard disk with diskettes backup |

Figure 3. A summary of the hardware features.

## The Main Algorithm

The algorithm is based on a simple procedure used successfully at NBS; it is described in [1] and is outlined here shortly. The inputs to the algorithm are the measured phases Xij of all clock pairs, with time intervals between measurement cycles of TO. TO was chosen to be 3 hours, an interval long enough to eliminate the influence of the system's jitter on the computed frequency offsets.

The procedure starts by assigning a high weight to the present master clock (77%) and equal small weights to the other clocks. The time constant of the exponential filter which allows the change of the weights (Nr in ref [1] notation) was chosen to be 15 days. The master's phase difference from the software clock at start-up is defined to be 0. Good convergence was obtained when deriving the weights from the square of the second differences of the phases; that is, the weights used in the next iteration to compute the average frequencies are obtained for each clock as follows:

$$E(t+T1) = [SSD(t) + Nr*E(t)]/(Nr+1) \quad (1)$$

$$W(t+T1) = k/E(t+T1) \quad (2)$$

where T1 is the time between iterations chosen in the present procedure to be 24 hours.

SSD is the squared second difference of each clock from the software clock,
E is the filtered squared second difference,
W is the weight to be used in the next iteration,
k is a dynamic normalization factor which keeps the sum of all W's equal to 1.

Once the offsets Yio from the software clock are computed, they are exponentially filtered with a dynamic time constant, Mi, which is related to the time interval (multiple of TO) for which Allan variance is a minimum (see ref. [1]), as follows:

$$\hat{Y}io(t+T1) = [Yio(t+T1)+Mi*\hat{Y}io(t)]/(Mi+1) \quad (3)$$

$$\hat{X}io(t+T1) = Xio(t) + \hat{Y}io(t)*T1 \quad (4)$$

where $\hat{Y}io$ and $\hat{X}io$ are the next predicted values of the frequency offset and phase respectively of the $i^{th}$ clock. Xio (the final phase difference from the software clock) is obtained as follows

$$Xio(t) = sum(j)\ \{Wj(t-T1)*[\hat{X}jo(t) - Xji(t)]\} \quad (5)$$

where Xji is the measured phase of clock j minus the phase of clock i.

The outputs of the system are the phases Xio and frequency offsets, $\hat{Y}io$ for each ensemble clock i at intervals T1. For monitoring purposes the two-sample Allan variance is computed for time intervals TO. The reasons to choose T1 = 24 hours were:

a) It differs from TO (=3 hours) by a power of 2, which makes it convenient in Allan variance calculations.

b) The diurnal cycle influences are minimized by taking one measurement every day at the same time.

c) It coincides with the BIPM circular-T with the measurements taken at 00 hours UTC.

Figure 4 shows the filtered SSD (see above) which is the basis to compute the weights. They evolve restricted by the time constant Nr from their initially assigned values.

Figure 5 shows the change of the weights with the time constant. The weight of clock 0 (the previous master) was initially assigned a high value to

enable a smooth start-up (actually the SSD were assigned initial values).

Figure 6 shows the Yio for the first month after the start-up. The Yio were given some initial arbitrary values which evolve with the time constant Nr towards the right values.

Figure 7 shows the evolution of the phases Xio. In this figure initial values are subtracted to allow insertion of all graphs into one scale.

Figure 8 shows the square root of the Allan variance of one clock. The value at time interval of 12 hours is sometimes higher than expected from the adjacent values and supports the probability of diurnal influences. As noted above,the choice of T1 considered this problem.

T1 being so long may impose some inconvenience from the point of view of accessibility, such as when frequent comparisons are needed or a lock of a less stable hardware clock to the software clock is desired. To overcome this problem it is possible to run the algorithm in parallel with two values of T1 i.e. T1a = 24 hours T1b = 3 hours, all the time constants generated with the run with T1a are used for the run with T1b. Since. at any time, the difference between Xioa and Xiob obtained in this way is known, the software clock can be hardware accessed at the shorter time intervals, T1b.

## Steering

The present system enables steering of its output to compensate for deviations from UTC. The method follows the guidelines described elsewhere [2] for steering to GPS.

The steering control loop is described in Fig. 10. The task is to keep the systems steered phase and frequency as close as possible to UTC. In the steered system, Xiutc, the phase of clock i with respect to UTC is

$$Xsutc = Xiutc - Xio - Xos \quad (6)$$

$$Xis = Xio + Xos \quad (7)$$

where

Xsutc is the steered output with respect to UTC,
Xos is the steering phase added to the software phase.
The phase of clock i from the steered system is now Xis.

At the start-up, for the previous master clock we define Xlo and Xos = 0, and therefore Xiutc = Xsutc.

To obtain convergence of Xsutc to zero we change Xos at the rate Ys defined for the next iteration as follows

$$Ys(t+T) = \{m*Ys(t) - [Xsutc(t) - Xsutc(t-T)]/T\}(m+1)-\ell*Xsutc/T \quad (8)$$

where

m is the exponential time constant for filtering Ys, T is 10 days, the BIPM Circular-T time interval, $\ell$ is a parameter which defines the phase steering.

The next Xos is defined simply by

$$Xos(t+T) = Ys(t+T)*T+Xos(t). \quad (9)$$

The values for the parameters m and $\ell$ were obtained using the simulation of ref [2] with BIPM Circular-T data files. The simulation assumed very conservative flicker FM levels for the INPL ensemble and for UTC (1e-13 and 2e-14 respectively). The former is about a factor of 2 larger than it will be and the latter is degraded because we have to use a predicted value to estimate UTC for the current time for servo control.

The parameter choices involve trade-offs between the best short-term or long-term stability. Figure 10 shows the simulated UTC performance used in the calculations and Figure 11 that of the clock ensemble.

Figure 12 is a typical output from which the parameters m and $\ell$ were chosen. The values $\ell$ = 6.0 and m = 3.2 were a reasonable trade-off choice. These values give rms time residuals from the simulation of 250 ns, which should also be conservative.

Unfortunately, because of the long time constants involved in the steering, we do not have yet meaningful results of steering the software clock to be shown in the present paper.

## Calibration

As shown in Figure 1, up to 15 clocks may be connected to the system. These may include GPS receivers for direct comparisons with the system. The better performance of the new time scale together with the automated measurements are expected to upgrade the precision of the services given by the INPL.

## Future prospects

The INPL is considering some additions to the present system. One such addition is a hardware clock with good short term stability which will be locked to the system and provide improved hardware time much the same as a microstepper but with much lower jitter.

## References

[1] F. B. Varnum, D. R. Brown, D. W. Allan and T. K. Peppler, "Comparison of time scales generated with the NBS ensemble algorithm," 19th Precise time and time interval symposium, Dec. 1987.

[2] W. J. Klepczynski, H. F. Fligel and D. W. Allan, "GPS time steering," 18th Precise time and time interval symposium, Dec. 1986.

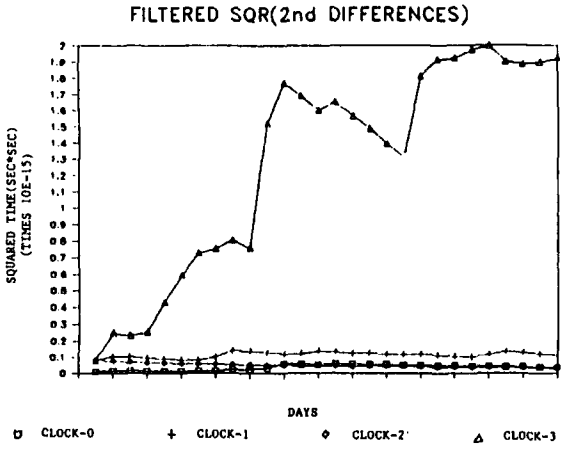



Fig. 4 Filtered Squared Second Differences From Which Clock Weights are Derived.

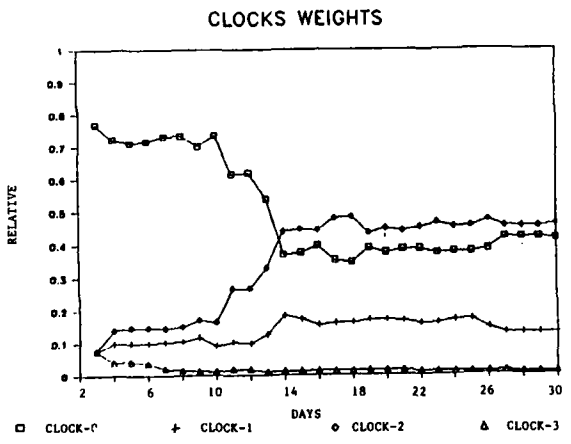



Fig. 5 Individual Clock Weights Versus Time.

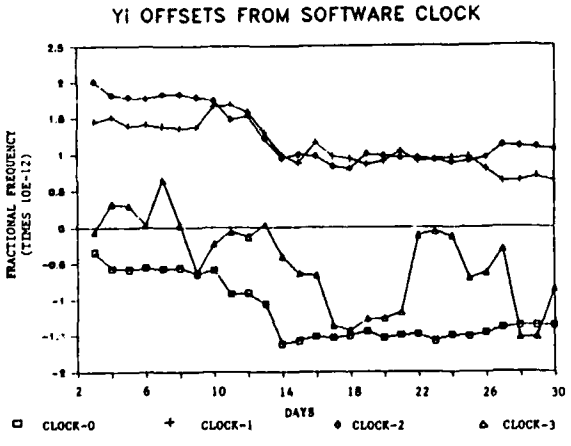



Fig. 6 Fractional Frequency Offsets of Individual Clocks Versus Time.

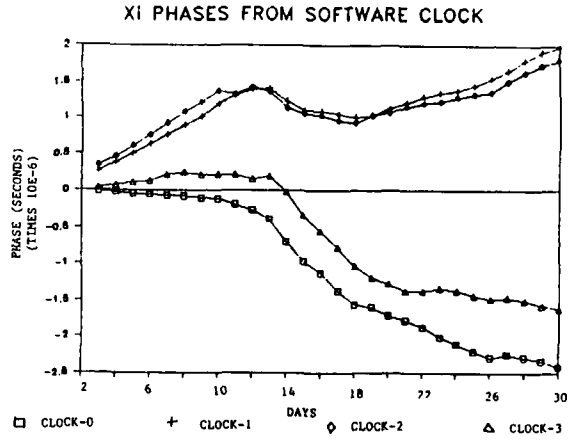



Fig. 7 Phase Evolution of Individual Clocks Versus Time.

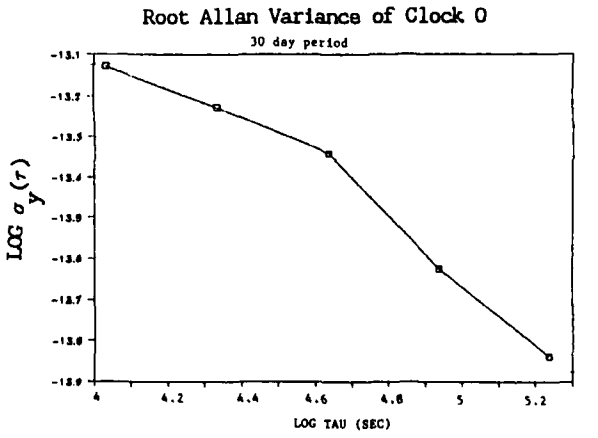



Fig. 8 Square Root of Allan Variance of Clock 0.

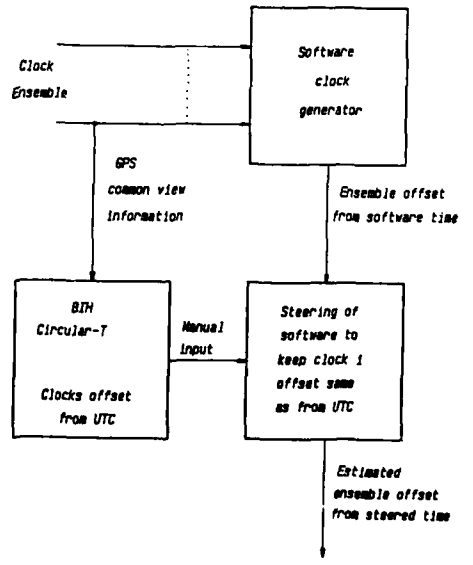



Fig. 9 Block Diagram of the Steering Control Loop.

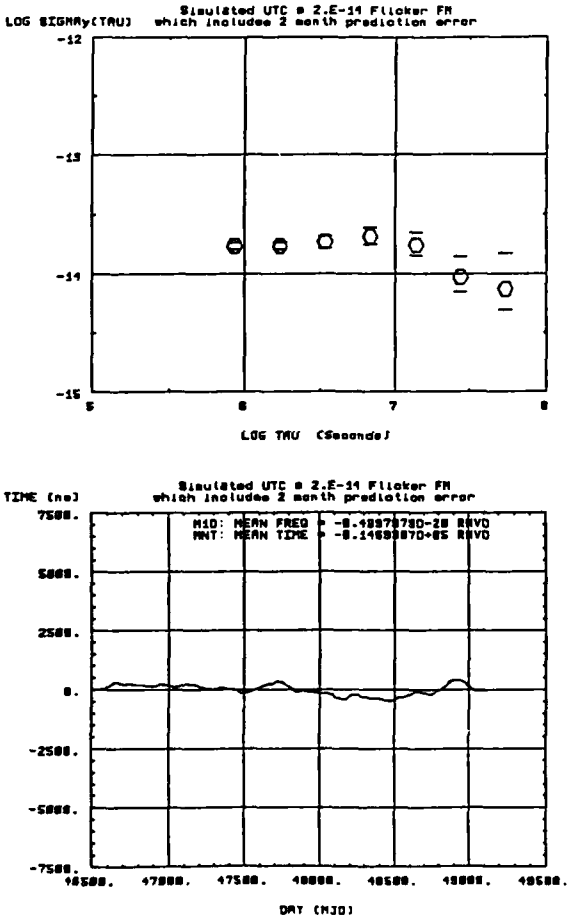



Fig. 10 Simulated UTC Performance.

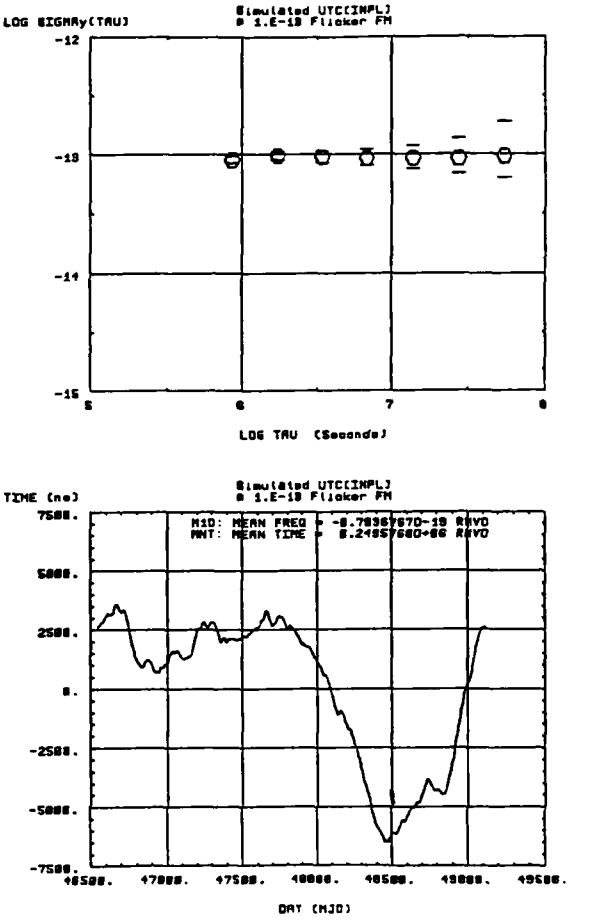



Fig. 11 Simulated UTC(INPL) Clock Ensemble Performance

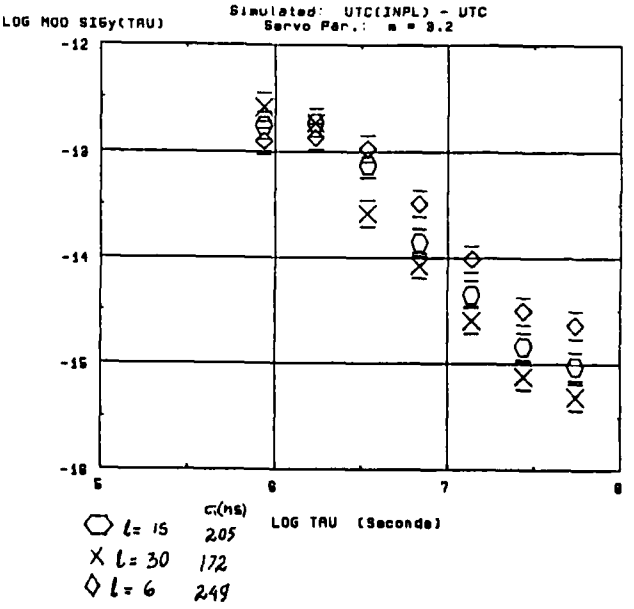



Fig. 12 Simulated Performance of UTC(INPL) - UTC, from which values for m and $\ell$ were chosen.

42nd Annual Frequency Control Symposium - 1988

# PERFORMANCE OF THE PTB'S PRIMARY CLOCKS CS2 AND CS1

A. Bauch, H. de Boer, B. Fischer, T. Heindorff, R. Schröder

Physikalisch-Technische Bundesanstalt
Braunschweig, Federal Republic of Germany

## Summary

Since summer 1985, two primary cesium atomic clocks have been in continuous operation in the PTB. In this paper the long-term performance of the two devices, CS1 and CS2, is described, based upon a 700 days data set. The output frequencies of the two clocks agree within their combined figures of uncertainty. Some features of the construction, and the evaluation of the uncertainty of the CS2 are discussed.

## 1. Introduction

The realization of the SI unit, the "second", and the production of the international time scale TAI is based upon the results of a few primary cesium clocks located all over the world. From the BIH reports, it can be understood that the PTB clocks CS1 and CS2 make a considerable contribution. It was in 1978 that the decision was taken to operate the CS1 continuously. During the decade of the clock's life, the reliability of the device was so excellent that the operation had to be discontinued due to failures for only some ten days in all. The CS1 has recently found a reputable competitor and successor in the CS2 clock. Its construction had been decisively influenced by the experience which had been gained with the CS1, so the essential features have been retained. For almost three years the CS2 has been in operation. Beside the intercomparisons with the CS1, numerous investigations have been undertaken to state its uncertainty figure correctly. It was, of course, determined as an estimate at the end of this limited time interval. In the past, confusion and doubt were common when the "quality" of primary clocks was judged. This was especially true as long as the systematic error inherent in international time comparisons considerably exceeded the uncertainties claimed. With the advent of the GPS the situation has changed. It is the authors' conviction that the uncertainty figure of the CS2, $15 \cdot 10^{-15}$ ($1\sigma$) will be confirmed in the future by continuing international intercomparisons. The CS2 was introduced for the first time at the CPEM'86 /1/. At that time only preliminary data were available. In this paper the performance of the two devices, CS1 and CS2, over an extended period is illustrated. In particular, the properties of the CS2 are pointed out. In the subsequent section some details of the construction are discussed which have a direct impact on the uncertainty budget. The latter is discussed in section 3. Finally, the intercomparison data are presented.

## 2. Description of the CS2

Most of the unique features of the PTB clocks are directly related to their common principle of construction: The axially symmetrical arrangement consists of a point beam source

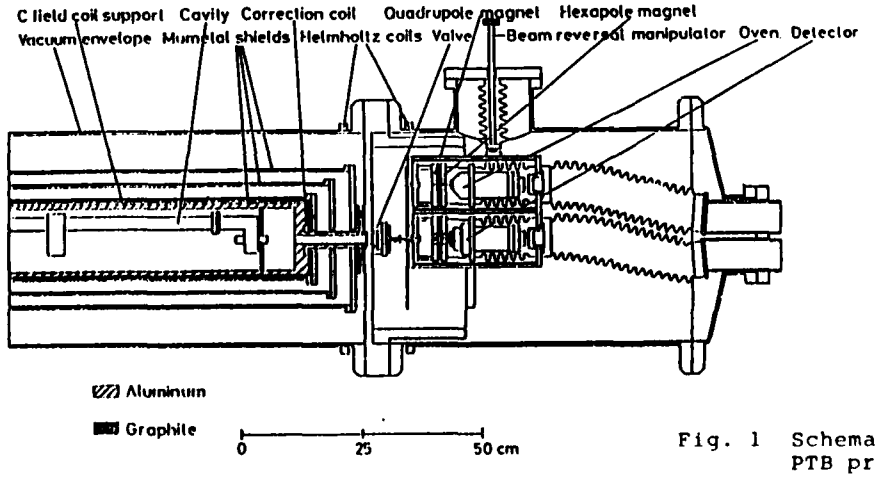



Fig. 1 Schematic drawing of one-half of the PTB primary clock CS2 (horizontal section, cavity rotated by 90°).

CH2588-2/88/0000-490 $1.00 C 1988 IEEE

and a point detector (1x1 $mm^2$), multipole fields for the state selection and velocity selection and an axial C-field generated by a solenoid. Some details of the construction of the CS2 are explained here and in /1/.

In Fig. 1 a horizontal section diagram of the CS2 is given. The CS2 is a two-beam apparatus. A beam reversal is accomplished by manually moving the oven/detector unit in each end chamber, the units gliding horizontally on ball bearings. Any interference from adjacent systems is essentially eliminated by shielding with graphite getter material. Beam reversal is thus a matter of 5 minutes. In the cesium ovens the reservoir containing 5g of cesium is heated to 420 K. The cylindrical nozzle, 0.5 mm long and 0.1 mm in diameter is held at an elevated temperature to avoid clogging. Each oven contains a valve to block the beam and seal the reservoir. From the above data the total flux is calculated to 1.6 · $10^{13}$/s, corresponding to a cesium consumption of ≈ 0.15 g/year.

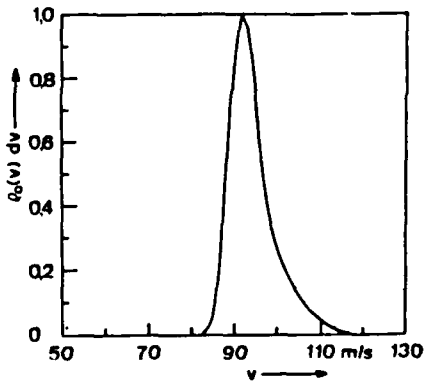



Fig. 2 Velocity distribution $\rho_o(v)$ of the atoms in the state (4,0) in the atomic beam of the CS2.

The four magnetic state selection systems are hexapole and quadrupole magnet combinations ("tandem") /1/. From the calculated trajectories, it can be roughly estimated that atoms emitted into a cone of 9° full angle are focussed to a paraxial beam. The imaging property, as related to the effective magnetic moment $\mu_{eff}$, thus depends on the quantum number m and determines the focussed velocity interval. This interval is centered around 95 m/s and covers about 0.5 % of the entire Maxwell-Boltzmann distribution in the case of (F = 4, m = 0)-atoms. The result of this simple calculation is an expected clock transition signal of 5.7 pA, not differing very much from the 3 pA found in practice. Fig. 2 depicts the velocity distribution $\rho_o(v)$, calculated from the (4,0)-(3,0) Ramsey pattern as described in 3C. The atomic beam may be said to be rather cool, since the half-width $\Delta v$ is less than 10% of the mean velocity v = 94,4 m/s. In the case of the other states, m ≠ 0, different values of v and $\Delta v$ are found and, in consequence, different populations of these states result. The recorded spectrum of the σ - transitions shown in Fig. 3 can thus be explained.

The construction of the Ramsey cavity has been described in /1,2/. Its geometrical parameters are: length of drift space, L = 76.7 cm, interaction length, $\ell$ = 23.26 mm. The field mode is such that the amplitude of the rf magnetic field seen by the atoms is cosinusoidal in shape. This is most advantageous in reducing pulling effects by neighbouring transitions /3/. The end-to-end cavity phase difference has been measured as $\varphi = 1.3 \cdot 10^{-4}$ rad (see 3D) without any noticeable aging during the period of operation.

From the parameters given so far, the frequency instability of the CS2 can be predicted. The linewidth of the clock transition line is 58.8 Hz, the line Q $1.56 \cdot 10^{8}$ and the signal-to-noise ratio is 1040/ $\sqrt{Hz}$. Thus $\sigma(2,\tau=1s) = 3 \cdot 10^{-12}$ is expected.

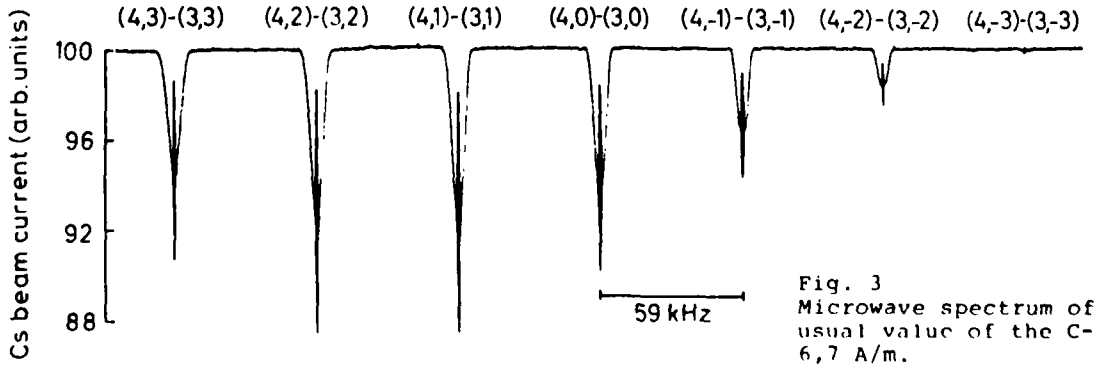



Fig. 3
Microwave spectrum of the CS2 at the usual value of the C-field strength 6,7 A/m.

Table I: Systematic CS2 frequency corrections and uncertainties in parts in $10^{15}$

| Type of Frequency Bias | Relative Frequency Bias | Relative Uncertainty |
|---|---|---|
| 1. C-field | - 326254 | 5 |
| 2. Second-order Doppler-effect | 48 | 1 |
| 3. Cavity phase | 260 | 10 |
| 4. Microwave power | - | 3 |
| 5. Cavity pulling | - | <0.1 |
| 6. Spectral impurities | - | 4 |
| 7. Rabi pulling | - | <0.1 |
| 8. Servo system | - | <2 |

## 3. The uncertainty budget of the CS2

Table 1 shows the systematic frequency corrections of the CS2 together with their uncertainties (1σ). They are given in relative units in parts in $10^{15}$ here and throughout this section. When UTC(PTB) is generated, the gravitational red-shift which amounts to 8, is taken into account. As long as no international recommendation concerning the black-body radiation shift is given, no such correction is applied. The total uncertainty of the CS2 is thus 15 (1σ, square root of the sum of squares). In the case of the CS1, a larger uncertainty in determining the cavity phase difference must be assumed. Its total uncertainty is therefore larger by a factor of 2 /4/.

Subsequently, a few comments are given on the determination and verification of some of the figures stated above. The reader is also referred to /5/.

### A C-field: measurement and stability

The line centre of the (4,1)-(3,1) transition is determined routinely with an uncertainty of 0.1 Hz. The Zeeman frequency $f_z$ and the C-field correction is then calculated. The uncertainty contribution due to the measurement error is 1.1. In the past, the value of $f_z$ was occasionally readjusted to the arbitrary value of 58.84 kHz. Deviations $\delta f_z$ of $f_z$ from the value adjusted once were recorded in the period until the next readjustment (e.g. change of C-field direction). The $\delta f_z$ values are depicted in Fig. 4.

The air-conditioning in the PTB's clock hall apparently reduced the instability of $f_z$: The main trouble is the temperature affecting the length of the C-field coil support made of aluminum with a temperature coefficient of $23.8 \cdot 10^{-6}$/K.

The CS2 frequency is corrected for the changes of $\delta f_z$ twice a week. The temperature is usually stable within 0.3 K during such periods. This results in a total uncertainty of the C-field correction of 5, including the effects of the temperature coefficient of the current supply.

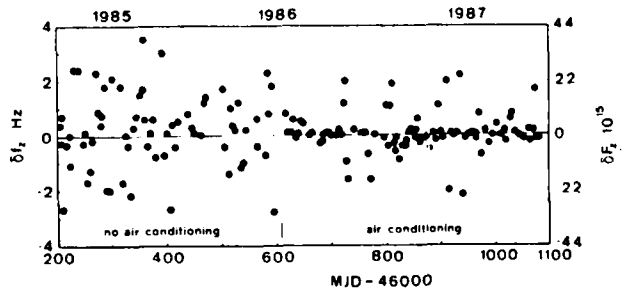



Fig. 4 Zeeman frequency changes of CS2, $\delta f_z$, in Hz and expressed as relative transition frequency corrections, $\delta F_z$. MJD: Modified Julian Date.

### B C-field: inhomogeneity and related phenomena

When a solenoid enclosed by Mumetal shields is used, a very small inhomogeneity of the C-field in the cavity region is attainable. It amounts to 0.005 A/m, compared with the operational C-field strength of 6.7 A/m. No asymmetry of the (4,3)-(3,3) Ramsey pattern is measureable and the contrast of its fringes is the same as that of the (4,0)-(3,0) pattern. The contribution of the uncertainty of the C-field correction (cf. 3A) caused by the use of $(\bar{H})^2$ instead of $\overline{H^2}$ in the calculation is negligible. However, great care has to be taken to ensure an adiabatic passage of the atoms along their path between both state-selecting magnets. In the openings of the magnetic shielding small guiding field coils and pairs of Helmholtz coils outside the vacuum vessel are placed to produce a magnetic field smoothly increasing from the low C-field value to the strong fields in the magnets without a rotation of the field direction. Following the ideas published in /6/, indications of Majorana transitions were sought.

Since the velocity distributions of the beam atoms in the different states (1,m) are quite different, they are considerably altered when the magnetic field conditions are worsened to such an extent that Majorana transitions are induced. This is shown in a change of the linewidth and the contrast of the Ramsey fringes. The (4,3)-(3,3) pattern is the most sensitive of all and therefore this has been examined.

During the Rabi pulling measurements (see 3E), values of the C-field strength between 0.7 A/m and 13 A/m were used. No measureable influence (except on the centre frequency!) was found as long as the Rabi pedestals were reasonably well separated. The same is true when the C-field direction is reversed by reversing the currents through all the above mentioned coils. Furthermore, within the statistical error of 10 ($1\sigma$), the CS2 frequency is the same, regardless of the C-field direction.

Majorana transitions do not therefore seem to be a problem in the CS2, although, there is still a lack of proof that the CS2 frequency (as well as the frequencies of all existing primary clocks) is not biased by related effects. Since no theoretical prediction is yet known, a contribution to the uncertainty budget was omitted.

## C Second-order Doppler correction

The velocity distribution $\rho_o(v)$, see Fig. 2, of the atoms in the (4,0) state is obtained by Fourier analysis of the resonance pattern. It is assumed that the transition probability $P_o$ is described by

$$P_o(\omega) = \cos^2((\omega_o-\omega) \cdot L_D/v). \qquad (1)$$

with: $\omega_o$: resonance frequency
v: velocity of the atoms

$L_D$ is the effective length to describe $P_o$ by a simple formula like (1). When the equation V.37 in /7/ is expanded using $|\omega_o-\omega| \ll b$ (b: interaction strength), $L_D$ turns out to be

$$L_D = L + 8 \cdot \ell/\pi^2 \qquad (2)$$

in the case of a cosinusoidally shaped interaction field /8/.

Due to the small width of $\rho_o(v)$, the optimum excitation condition is fulfilled for all atoms in a good approximation which allows (1) to be used. In the Fourier transformation the integration interval of $\omega$ has to overlap all Ramsey fringes. Then $|\omega_o-\omega|$ becomes of the order of b, which may cause a slight error in the shape of $\rho_o(v)$.

The second-order Doppler correction is calculated taking into account the specific modulation scheme during frequency measurement, in a way similar to that described by De Marchi /9/. The result is 48. The uncertainty of 1 is estimated from the results of several calculations in which the input data set as well as integration parameters were varied.

In addition, the method of Audoin et al. /10/ has been applied. The results agree if the same value of $L_D$ is used.

## D Cavity phase difference

The uncertainty contribution related to the cavity phase difference is the largest of all and deserves a detailed analysis. First the evaluation of the systematic uncertainty will be illustrated, then the measurement results presented.

During the process of mechanical alignment prior to the completion of the CS2, two holes in the middle vacuum chamber were used to realize the symmetry axis by a tightened steel wire (see "valve" in Fig. 1). The axes of the C-field coil support and of the magnetic shields as well as the cavity openings were adjusted in relation to the wire with a precision of less than 0.05 mm.

The next step was to adjust the symmetry axis of the magnet pairs 1 and 2, in use in the beam directions 1 and 2, with similar precision. Fig. 1 shows (as an example) oven 1 with polarizer 1 and detector 2 with analyzer 2. When numerous beam reversals were simulated, the reproducibility of the position adjusted once was found to be better than 0.02 mm.

Finally, the CS2 was put into operation and the ovens and detectors were adjusted to obtain maximum beam signal in both beam directions. The two signals differ by 3%, the mean velocity of the (4,0) atoms by 1.7%. The oven and detector adjustment has never been changed by more than 0.02 mm. From these findings we conservatively estimate the error of beam retrace after a reversal to be 0.1 mm.

The distributed cavity phase difference in the CS2 type microwave interaction region was $9.4 \cdot 10^{-5}$ rad/mm in the vertical direction /1,2/. The uncertainty contribution of $\delta F_\varphi$ can thus be estimated to be 10.

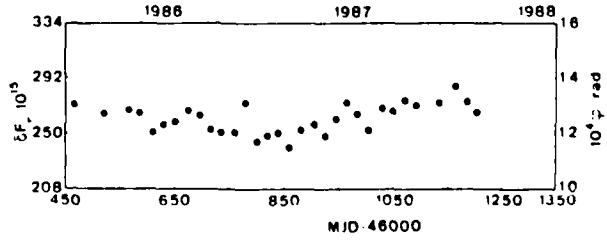



Fig. 5 The cavity phase difference $\varphi$ and the corresponding frequency correction $\delta F_\varphi$ of CS2 between 1986 and 1988. CS1 is used as reference clock.

The cavity phase difference $\varphi$ is determined by measuring the frequency difference of CS2 between both beam directions, $2 \cdot \delta F_{\varphi}$. Just before and after a beam reversal, a 10-day average of the frequency difference between the CS2 and a reference clock is calculated. Fig. 5 demonstrates that the cavity phase difference $\varphi$ of the CS2 is very stable. The result is $\delta F_{\varphi} = 260$ with a $1\sigma$ deviation of the individuals of 10. This instability is in accordance with the ordinary frequency instability observed, as can be inferred from Fig. 8. Thus the mechanical imperfections of the beam reversal do not contribute much to the instability of $\delta F_{\varphi}$. They may, however, have a systematic effect which has been estimated above.

E Rabi pulling

The Rabi pulling due to (4,m)-(3,m) transitions /3/ has been studied in detail /11/. Although the amplitudes and linewidths of the $\sigma$-transitions are quite different (cf Fig. 3), the Rabi pulling is no problem since the interaction field is of a cosinusoidal shape in the CS2 cavity. A reduced C-field and Zeeman frequency has been used to check the validity of the published theory /3/. Fig. 6 depicts the measurement results. The agreement between calculation and experiment is satisfactory. This allows a realistic statement of the uncertainty contribution to be made: It remains smaller than 0.1 as long as $f_z$ is above 50 kHz.

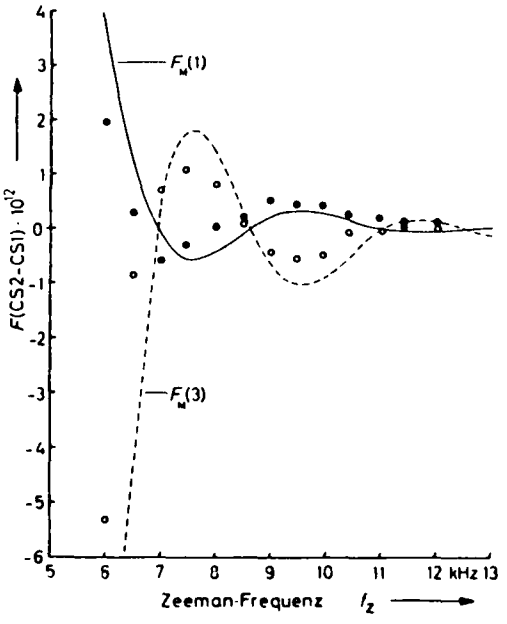



Fig. 6 Dependence of the relative frequency F(CS2-CS1) upon the Zeeman frequency $f_z$ of CS2.
•• Measured values $F_M(1)$ using 60 Hz modulation width,
—— corresponding theoretical data.
oo Measured values $F_M(3)$ using 180 Hz modulation width.
- - - corresponding theotetical data.

## 4. Long-term time and frequency comparison between the two primary clocks

The time scales derived from the CS1 and the CS2 are compared in Fig. 7. The data points are five-day averages. Except for the first few days, a frequency offset appears between the two clocks, so that F(CS2)>F(CS1). The overall average is $25 \cdot 10^{-15}$, a value within the limit of the root of the sum of the squared total $1\sigma$ uncertainty values, $34 \cdot 10^{-15}$ or of the sum of these figures for CS1 and CS2, $45 \cdot 10^{-15}$. The instability of the frequency data is worse than expected from the estimated short-term stability of the clocks.

The expected frequency instability of the CS2 at $\tau = 1s$, $\sigma\,(2,\tau=1s) = 3 \cdot 10^{-12}$, cannot be measured unambiguously with the equipment available. The frequency instability of the CS1 is unfortunately larger than the calculated value of $6 \cdot 10^{-12}$, due to noise at the beam signal which is caused by vacuum problems. The comparison data therefore exhibit an instability of $\sigma_A\,(2,\tau=1s)$ of $10.6 \cdot 10^{-12}$.

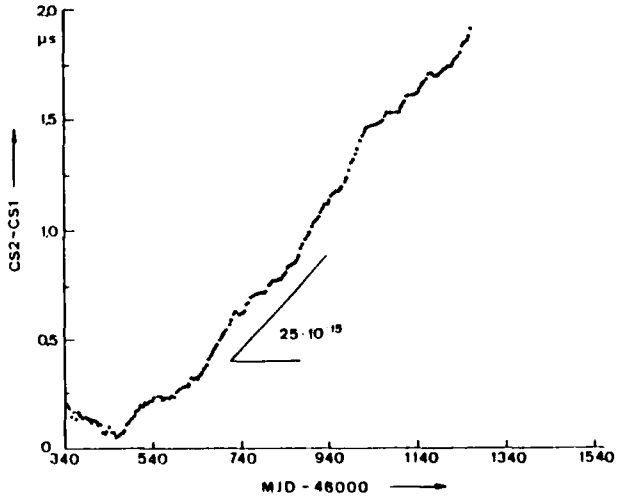



Fig. 7 Comparison of the time scale readings of the CS1 and the CS2 between 1986 and 1988.

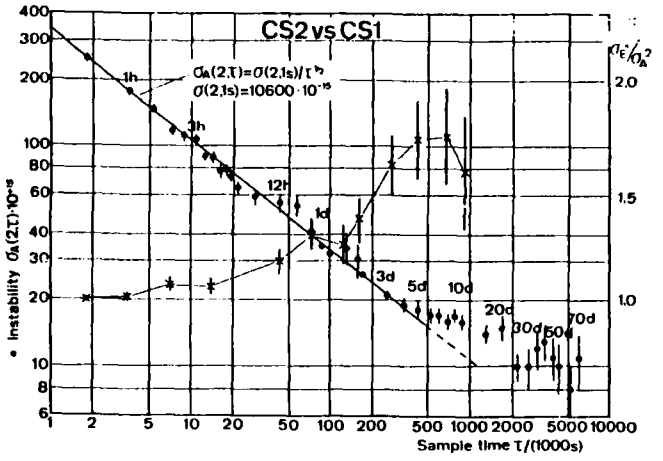



Fig. 8 Two-sample-standard deviation $\sigma_A(\tau)$ (left scale) and ratio $(\sigma_E(\tau)/\sigma_A(\tau))^2$ (right scale) of CS2 versus CS1 as a function of the sample time $\tau$.

This is produced to a large extent by the CS1, as can be deduced from a three-corner analysis with an HP supertube as a third clock.

Fig. 8 depicts the $\sigma$ - $\tau$ diagram of the frequency comparisons. Between 2000 s $<\tau<$3d, the frequency instability shows the expected behaviour of a white frequency noise process due to the shot noise of the atomic beam. Having calculated, in addition, the classical variance $\sigma_E^2(\tau)$ of the data set, we can make the "ratio-1 test" /12/. We note an increasing difference from the value 1 of the ratio $\sigma_E^2/\sigma_A^2$ as can be seen from Fig. 8. It is evident that the dominant noise process begins to change. With equation (7) in /12/, the exponent $\mu$ ($\sigma^2(2,\tau)\sim\tau^{\mu}$) has been calculated to be -0.70 at $\tau$ = 1 day and -0.45 at $\tau$ = 10 days. Prolonged measurements will be necessary to decide with sufficient statistical significance the noise process dominant at very long measurement times. However, a flicker level just below the $10^{-14}$ level must be suspected.

## 5. Conclusion

The two PTB primary clocks, CS1 and CS2, have proved their reliability in the past and will be valuable in the future for realizing the PTB time scale as well as the TAI. Up to now, their output frequencies agree within the uncertainty figures. However, the apparent systematic offset as well as the long-term instability merit future investigations. The completion of two other clocks, CS3 and CS4, has been somewhat delayed, but may be expected in 1989. They will allow three-corner comparisons between primary clocks in the PTB to be made, and facilitate such research work. With the new devices, it is the prime objective to reduce the uncertainty contribution due to cavity phase difference: They will operate with a slower atomic beam (v ≈ 75 m/s) running vertically.

## References


/1/ A. Bauch et al.: IEEE Trans. IM-36, 613-616 (1987).

/2/ A. Bauch, T. Heindorff, R. Schröder: IEEE Trans. IM-34, 136-138 (1985).

/3/ A. De Marchi, G.D. Rovera, A. Premoli: Metrologia 20, 37-47 (1984).

/4/ K. Dorenwendt: Proc. IEEE 74, 137-140 (1986).

/5/ A. Bauch et al.: Proc. of the 2nd EFTF, Neuchâtel, 1988, to be printed.

/6/ A. Bauch: Ph.D. thesis (in German), PTB-Bericht Opt-20 (1986).

/7/ N.F. Ramsey: Molecular Beams. London and New York: Oxford Univ. Press (1956).

/8/ P. Thomann, Oscilloquartz S.A., private communication.

/9/ A. De Marchi, G.D. Rovera, A. Premoli: IEEE Trans. UFFC-34, 582-591.

/10/ C. Audoin, P. Lesage, A.G. Mungall; IEEE Trans. IM-23, 501-508 (1974).

/11/ A. Bauch, K. Dorenwendt, T. Heindorff: Metrologia 24, 199-203 (1987).

/12/ D.W. Allan: IEEE Trans. IM-36, 646-654 (1987).

42nd Annual Frequency Control Symposium - 1988

# RESEARCH ON THE OPTICALLY PUMPED CESIUM BEAM FREQUENCY STANDARDS

G. Théobald, V. Giordano, M. de Labachelerie, A. Hamel,
N. Dimarcq, P. Cérez and C. Audoin

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

## Abstract

After having recalled the interest of applying optical methods to the state selection and the state detection of atoms in a cesium beam frequency standard, we summarize the work done at the Laboratory on the physics of optical pumping applied to a cesium beam. We shall give results, depending on the optical transition and the light polarization chosen, on i) the population difference which can be achieved and ii) the mean number of fluorescence photons which can be obtained for each incoming atom when a pumping transition is used for the atom detection. We shall indicate the limit of validity of the model and give consequences of the non linear Hanle effect on the efficiency of the state preparation and on the intensity of the fluorescence light.

We shall present practical results achieved with an experimental set-up in which a single semiconductor laser is used for the state selection and the atom detection. This device uses a 21 cm long separation between the two microwave interaction regions. It gives a 500 Hz wide linewidth which is currently observed with a signal to noise ratio equal to 10,000 in a 1 Hz noise bandwidth.

An experimental electronic system has been realized which includes innovative features. It has been used to measure the short term frequency stability of the optically pumped cesium beam frequency standard. The result is $\sigma_y(\tau) = 2 \times 10^{-12}/\sqrt{\tau}$ for $1 \text{ s} < \tau < 500 \text{ s}$. For longer sampling times the set-up, in its present state, does not enable to achieve significant frequency stability results. Nevertheless, the capability of optical methods for improving by one order of magnitude the cesium frequency standard frequency stability is proven. Discussion of possible improvements shows that another order of magnitude can be gained in a standard size cesium beam frequency standard.

Information will be given on the work done with extended cavity semi-conductor lasers which enhance greatly the spectral purity of the laser light. These devices are of interest for using cycling transitions for the atom detection. Their capability of improving the detection signal to noise ratio will be reported.

## 1. Introduction

The possibility of applying optical methods to the state preparation and to the state detection of atoms in a beam is known since a long time [1]. It has been confirmed about 20 years ago using a rubidium 87 beam and spectral lamps [2]. More recently, the advent of reliable semi-conductor lasers in the near infrared has made very attractive the application of these methods to the cesium beam frequency standard [3] and a number of laboratories and companies have engaged research works and development activities in this field [4-15].

In this paper, we shall summarize the work done until now at Laboratoire de l'Horloge Atomique with the objectives i) of understanding the details of the interaction between the cesium atoms of a beam and a laser light beam and ii) of investigating experimentally the properties of the optically pumped cesium beam frequency standard. Additional results will be presented elsewhere [16].

## 2. Interest of optical methods applied to the cesium beam frequency standard

At first, we shall briefly recall the expected impact of optical methods applied to the state preparation and to the state detection of atoms in the cesium beam frequency standard.

### 2.1. Short and medium terms frequency stability

The impact of optical methods on the short term frequency stability is essentially related to the increase of the flux of detected atoms. This proceeds from the removal of the magnet state selectors.
i) These state selectors deflect the atom trajectories with an angle which depends greatly on the atom velocity and, consequently, they attenuate the flux of atoms striking the hot wire detector by a factor of several hundreds.
ii) The need of using a very narrow detector for avoiding the detection of atoms in an undesired state disappears. It follows that the cross section of the detection zone can be increased noticeably when optical detection is implemented.
iii) Considering atoms in the hyperfine sub-levels $F = 3$, $m_F = 0$ and $F = 4$, $m_F = 0$, magnetic state selection eliminates from the beam atoms in one of these states. On the contrary, as saturated optical pumping with a single laser transfers all atoms of one of the hyperfine levels $F = 3$ or $F = 4$ to the other one, optical state preparation provides a gain of a factor of about two in the number of detected atoms. In a more involved configuration, two-laser optical pumping is able to provide a gain almost equal to 16 [17].

However, optical methods may bring drawbacks. The line quality factor is decreased since the magnet velocity selection on the low side of the maxwellian distribution does not occur any longer. But this effect is more than compensated by the achievable increase of the flux of atoms detected. Other detrimental effects are due to the additional shot noise of the stray light and, mainly, to the intrinsic noise of the optical detector. This point will be discussed in Section 3. In Section 4, the optical transition leading to the largest microwave signal will be specified. In Section 5, the achievement of a frequency stability given by $\sigma_y(\tau) = 2 \times 10^{-12}/\sqrt{\tau}$ will be described and prospects for improvement including use of an extended-cavity semiconductor laser will be given in Section 6.

### 2.2. Frequency shifts

There are good reasons to believe that optical methods are intrinsically able to reduce a number of annoying frequency shifts [18] and, consequently, to improve the long term frequency stability as well as the

CH2588-2/88/0000-496 $1.00 C 1988 IEEE

accuracy of cesium beam frequency standards.
i) Optical pumping with linearly polarized light populates equally the hyperfine levels having opposite values of the magnetic quantum number $m_F$. It follows that the amplitude of the seven $\Delta F = \pm 1, \Delta m_F = 0$ lines of the microwave spectrum of the cesium atom varies symmetrically around the central $F = 4, m_F = 0 \longleftrightarrow F = 3, m_F = 0$ line as shown in Fig. 1. In that case, the Rabi pulling frequency shift [19,20] should disappear along with its associated inaccuracy and with the frequency change related to a variation of the value of the static magnetic induction.
ii) Another benefit of the absence of Rabi pulling relies on the possibility of decreasing the value of the magnetic induction applied to the atoms without introducing a harmful effect. This would reduce still more the sensitivity to ambient magnetic field changes.
iii) Removal of state selector magnets creating strong magnetic fields will improve the magnetic field homogeneity in the whole region including the two microwave interaction regions and the drift region between them. This will decrease greatly the related frequency shifts.
iv) Indirect frequency shifts due to Majorana transitions should disappear, as a consequence of the better magnetic field homogeneity.
v) The optical detection enables the beam reversal method to the easily implemented, thereby providing a routine means of measuring the cavity phase shift.
vi) The detection angle of the atomic trajectories is approximately two orders of magnitude smaller in an optical pumping region than in a magnet state selector ($10^{-4}$ radian instead of $10^{-2}$ rad.). This is favorable to a better estimate of the distributed cavity phase shift or to its measurement by beam reversal.

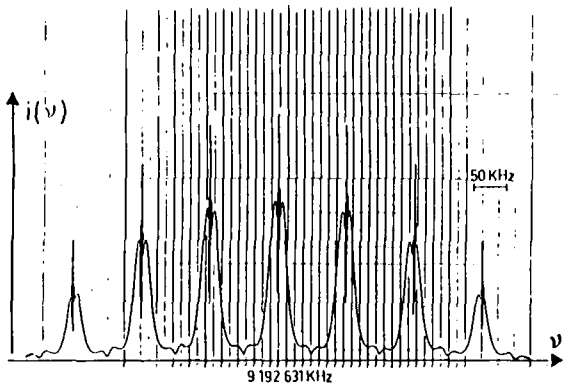



Fig.1

Microwave spectrum of the optically pumped cesium atoms. Note the symmetry of the spectrum around the central line.

The light shift, which is specific to double resonance methods will be very small [21,22] since the light flux in the microwave interaction regions is solely due to the cesium atoms fluorescence which is coming from a distant optical interaction region. It should not prevent an improvement of the long term frequency stability and of the accuracy.

## 3. Noise in the optical detection of atoms

An estimate of the short and medium term frequency stability is given by :

$$\sigma_y(\tau) \simeq \frac{\tau^{-1/2}}{2Q_\ell \left[\mathcal{S}/\mathcal{N}\ (B = 1\ \mathrm{Hz})\right]^{1/2}}, \qquad (1)$$

where $Q_\ell$ is the atomic quality factor and the quantity $\left[\mathcal{S}/\mathcal{N}\ (B = 1\ \mathrm{Hz})\right]^{1/2}$ represents the amplitude signal to noise ratio in a noise bandwidth B equal to 1 Hz.

In the ideal case where the signal to noise ratio is determined by the atomic beam shot noise, we have :

$$\frac{\mathcal{S}}{\mathcal{N}} = \frac{\overline{I_a}}{2B}, \qquad (2)$$

where $\overline{I_a}$ is the useful flux of atoms in either the state $F = 3, m_F = 0$ or the state $F = 4, m_F = 0$. The shot noise limit, shown in Fig. 2, is reached with a surface ionization detector, assuming a complete detection of the $Cs^+$ ions and neglecting the electron multiplication related noise. In presently manufactured cesium beam frequency standards, we have $\overline{I_a} \simeq 2 \times 10^6 s^{-1}$, giving $\left[\mathcal{S}/\mathcal{N}\ (B = 1\ \mathrm{Hz})\right]^{1/2} \simeq 10^3$.

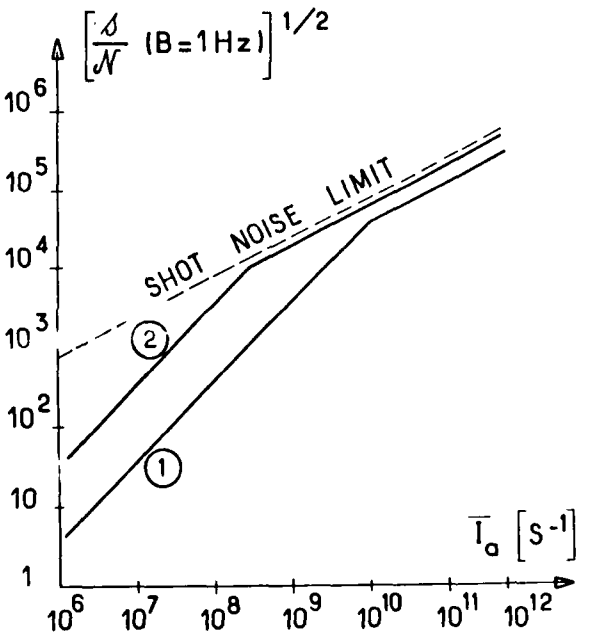



Fig. 2

Examples of variation of the amplitude signal to noise ratio in a 1 Hz noise bandwidth versus the flux of incident atoms in the $F = 3, m_F = 0$ or the $F = 4, m_F = 0$ state.
Curve 1 with $\beta = 4$ (pumping transition),
Curve 2 with $\beta = 30$ (cycling transition).

In the case of optical detection and assuming that the statistics of the fluorescence photon emission follows a Poisson distribution, it can be shown [23] that we have

$$\frac{\mathcal{S}}{\mathcal{N}} = \frac{(\overline{I_a})^2}{2F(\overline{I_a} + I'_b)B}, \qquad (3)$$

where F is the noise factor of the optical detection of atoms. It is given by :

$$F = 1 + \frac{1}{\beta\gamma\eta}. \qquad (4)$$

$\beta$ is the mean number of fluorescence photons emitted per cesium atom in the state F = 4, $m_F$ = 0 or F = 3, $m_F$ = 0. $\gamma$ and $\eta$ are the photon collection efficiency and the photodetection quantum efficiency, respectively. In the experiment reported below, we have $\beta$ = 4, $\gamma$ = 0.15 and $\eta$ = 0.85. It follows that we have F = 3.0. In Eq. (3), $I'_b$ represents the effect of additional noise, not related to the beam shot noise. In the case of a photocell, it is given by :

$$I'_b = \frac{2\eta \overline{I_s} + F_A(P_n R/e)^2}{2(\beta^2\gamma^2\eta^2 + \beta\gamma\eta)} \qquad (5)$$

Two causes contribute to the value of $I'_b$. The first one is the shot noise associated with the flux, whose mean value is $\overline{I_s}$, of stray photons reaching the photodetector. The second cause is the photodetector thermal noise and the current to voltage converter noise. In that respect, $P_n$ and R are the noise equivalent power and the responsivity of the photocell, respectively. e is the elementary charge. $F_A$ is the noise factor of the current to voltage converter. Physically, $I'_b$ is a fictitious flux of detected atoms, whose shot noise is equivalent to the stray light shot noise, plus the photocell thermal noise worsened by the amplifier noise.

In our set-up, the light rejection factor is $2.5 \times 10^7$. With an incident light beam power of 1 mW at 0.85 µm, we have $\overline{I_s} = 1.7 \times 10^8\ s^{-1}$ and the stray photon flux contributes negligibly to $I'_b$. We also have $P_n = 2.5 \times 10^{-14}\ W.Hz^{-1/2}$ and $R = 0.6\ A\ W^{-1}$. Assuming $F_A$ = 1.8, the theoretical value of the noise factor of the current to voltage converter, we thus obtain $I'_b = 1.1 \times 10^{10}\ s^{-1}$.

For $\overline{I_a} \ll I'_b$, the signal to noise ratio is determined by the additional noise and we have :

$$\frac{\mathcal{S}}{\mathcal{N}}(\overline{I_a} \ll I'_b) = \frac{(\overline{I_a})^2}{2FI'_b B}\ , \qquad (6)$$

whereas for $\overline{I_a} \gg I'_b$, it is shot noise limited and we obtain :

$$\frac{\mathcal{S}}{\mathcal{N}}(\overline{I_a} \gg I'_b) = \frac{\overline{I_a}}{2FB}\ . \qquad (7)$$

The two different behaviours are shown in Fig. 2, where the two curves are derived from numerical data specified above, but for $\beta$ = 4 in the case of a pumping transition and for $\beta$ = 30 in the case of a cycling transition.

Some conclusions of interest can be derived from the figure

i) In the present state of the art of the optical detector design and of the photocell noise level, the optical detection method requires a noticeably larger atom flux than the surface ionization method, in order to achieve a 1 Hz amplitude signal to noise ratio larger than $10^3$. Fortunately, the required increase follows from the magnet state selector removal, as stated in Section 2.

ii) For a given value of the photon yield per atom, the shot noise limit is more closely approached, and for a smaller value of the atom flux, for collection efficiencies as close as unity as possible. However, one must take care that the stray photon flux does not increase up to detrimental values.

ii) For a given detector geometry, the shot noise limit is also more closely approached, and for a smaller value of the atom flux, if the photon yield per atom $\beta$ is increased. In this respect, when a pumping transition is used for the detection, it must be chosen carefully. In ordinary conditions, the transition is sufficiently saturated so that the laser frequency fluctuations do not matter usually. Large values of $\beta$, of the order of several tens, can be obtained with a cycling transition. But in that case, saturation effects do not occur for the usual light beam intensities ; consequently, the laser power spectral purity must be high enough to limit the additional noise caused by the laser frequency fluctuations.

## 4. Choice of the optical transition

Here, we shall only consider the most simple experimental arrangement in which a single laser is used to perform both the state preparation and the optical detection of the microwave resonance. Results in the case of two-laser pumping have been published elsewhere [17].

Fig. 3 shows, schematically, the configuration of the single laser optically pumped cesium beam tube. The light is linearly polarized, its electric field may be either parallel to the static magnetic induction $B_o$ ($\pi$ polarization) or perpendicular to $B_o$ ($\sigma$ polarization).

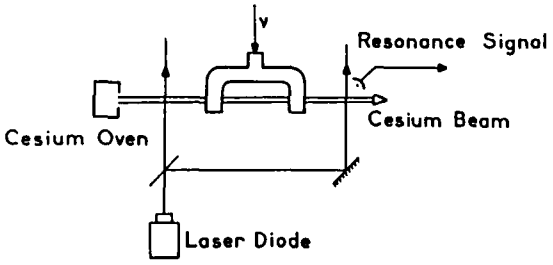



Fig. 3

Schematic representation of a single laser optically pumped cesium beam tube (not to scale).

### 4.1. Physical background

A population difference between the F = 3, $m_F$ = 0 and F = 4, $m_F$ = 0 ground-state hyperfine sub-levels is created in the first optical interaction region. The change of this population difference under the effect of the microwave $\Delta F = 1$, $\Delta m_F = 0$ transition is probed in the second optical interaction region, where the flux of fluorescence photons is proportional to both the probe absorption rate and the flux of atoms in the absorbing state.

Due to the arrangement of the cesium hyperfine levels, shown in Fig. 4, an incident linearly polarized laser light is absorbed by the atoms causing $F' - F = 0, \pm 1$ transitions, with $m_F - m_{F'} = 0$ or $\pm 1$ depending on whether the laser light is $\pi$ or $\sigma$ polarized. Once atoms are in an excited level, they decay by spontaneous emission to one of the ground state sub-levels. After many cycles of absorption followed by spontaneous emission, atoms are transferred from one hyperfine ground state level to the other one, with a given distribution between the (2F + 1) Zeeman sublevels. In particular, a population difference may be created between the two F = 3, $m_F$ = 0 and F = 4, $m_F$ = 0 clock sublevels. As a consequence of the linear light polarization and of symmetry properties of the transition probabilities

among the various Zeeman sublevels, this distribution is symmetric with respect to the sign of the $m_F$ values. This is the reason why the cesium atom microwave spectrum of Fig. 1 is symmetrical around the F = 3, $m_F$ = 0 $\longleftrightarrow$ F = 4, $m_F$ = 0 transition.

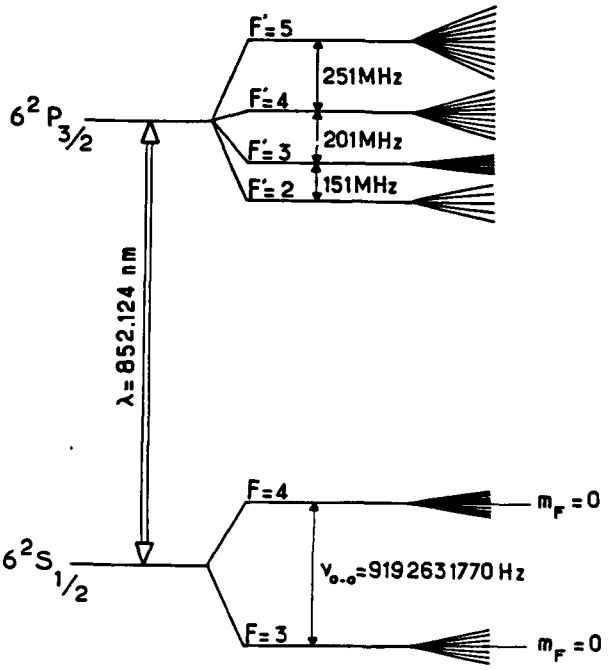



Fig.4

Energy-level diagram corresponding to the $D_2$ line of cesium atom. Each F or F' hyperfine level has (2F + 1) or (2F' + 1) Zeeman sublevels which are characterized by the magnetic quantum numbers $m_F$ and $m_{F'}$, respectively.

In ordinary experimental conditions where the laser spectral width is noticeably larger than the natural width of the cesium $D_2$ line and where the magnetic induction is small, it is justified to calculate the evolution of the level populations using the rate equations approximation. Then, between 21 to 27 coupled differential equations are obtained, depending on the F to F' optical transition considered.

4.2. State preparation

In the state preparation region, the population of the ground state sublevels is computed, assuming that all the ground state sublevels are equally populated while all the excited ones are actually depleted at the beginning of the interaction. We define the fractional population difference Δn of the two $m_F$ = 0 ground state hyperfine sublevels by :

$$\Delta n = \frac{n(F = 4,\ m_F = 0) - n(F = 3,\ m_F = 0)}{n} , \qquad (8)$$

where n is the total population of the cesium ground and excited state levels. Fig. 5 shows an example of variation of Δn versus the interaction time for the different possible pumping transitions, assuming the light intensity is I = 10 mW. $cm^{-2}$. One sees that Δn reaches a steady state value for most of the transitions after a light-atom interaction time of about 3 μs. Since, in practice the time spent in the light beam is much larger than this value, the level populations are saturated at the exit of the preparation region.

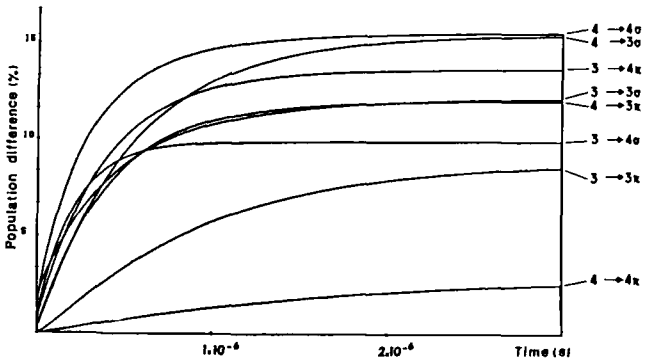



Fig.5

Time evolution of the absolute value of the fractional population difference Δn between the $m_F$ = 0 ground state sublevels, for different optical transitions. It is assumed I = 10 mW.$cm^{-2}$.

The highest values of Δn, equal to - 15.5 % and to - 15.4 % are obtained for the (4-4,σ ) and the (4-3, σ ) transitions, respectively.

4.3. Optical detection of the microwave resonance

In the detection region, the number $\mathcal{F}$ of photons emitted is equal to the number of atoms which have decayed from the excited state sublevels to the ground state during the optical interaction time T. We have :

$$\mathcal{F} = \int_0^T \Gamma \sum_{m_{F'}} n_{m_{F'}}(t)\, dt , \qquad (9)$$

where $m_{F'}$ is the magnetic quantum number of the particular excited state F' which is reached under the laser excitation. Γ is the decay rate of the cesium atom in the excited state. The value of $n_{m_{F'}}(t)$ is obtained by solving again the rate equations but with initial conditions given by the atomic populations at the output of the Ramsey cavity.

The quantity $\mathcal{F}$ has been calculated in two different cases. Firstly, when the microwave excitation is off, or largely detuned, it is found that the value of $\mathcal{F}$ is zero for all the optical transitions which are able to give a significant signal. Secondly, when the microwave transition is on resonance the value of $\mathcal{F}$ is calculated assuming optimum power and monokinetic atoms. To take the velocity distribution into account, the result is multiplied by the Ramsey probability averaged over the modified Maxwellian distribution of the atom velocities, which is close to 0.75. The photon yield of the beam is thus equal to S = 0.75 $\mathcal{F}$ . The number of fluorescence photons per atom in a given $m_F$ = 0 ground state, β , is given by β = $\mathcal{F}$ / Δn.

4.4. Results

The computed values of the population difference $\Delta n$, the signal S and the mean number of photons per atom in a given $m_F = 0$ ground state, $\beta$, are given in Table 1 for all the possible pumping transitions. It is assumed that the laser intensity is equal to 3 mW.cm$^{-2}$ in both optical interaction regions and that the time spent by the atoms in the laser light beam is 20 μs. It has been verified experimentally that the ratios between the different S values are in agreement with the theoretical ones [17]. Table 1 shows clearly that the (3-3, σ) transition is the best suited for use in the single laser configuration.

| Transition | Δn(%) | S | β |
|---|---|---|---|
| 3-3,π | - 8.5 | 0 | 0 |
| 3-3,σ | - 2 | 0.37 | 4.0 |
| 3-4,π | 13.7 | 0.17 | 1.7 |
| 3-4,σ | 9.9 | 0.12 | 1.6 |
| 4-3,π | - 12.1 | 0.12 | 1.3 |
| 4-3,σ | - 15.4 | 0.15 | 1.3 |
| 4-4,π | 2.5 | 0 | 0 |
| 4-4,σ | - 15.5 | 0.28 | 2.4 |

TABLE 1. Single laser configuration. Value of Δn the relative population difference created in the preparation region, of S the number of fluorescence photons per cesium atom emitted in the detection region and of β the number of fluorescence photons per cesium atom in a given $m_F = 0$ ground state sublevel emitted in the detection region. It is assumed I = 3 mW.cm$^{-2}$ and an interaction time of 20 μs.

It is worth noting that in the detection region, as well as in the state preparation region, the populations reach their steady state value after an interaction time of the order of a few μs, which is much smaller than the time of the passage across the light beam. It follows that the resonance signal does not depend, practically, on the laser amplitude and frequency noises.

4.5. Population leakage

In the physical model considered above, the possibility of exciting neighbouring optical transitions is ignored. However, the laser power spectrum being Lorentzian, with a linewidth of about 25 MHz, optical excitation is expected even far from the laser line center and a leakage of atoms on excited levels 151, 201 or 251 MHz away (see Fig. 4) becomes possible. Bearing this in mind, rate equations have been modified to include the case where two optical transitions share in common the same ground state. Calculations show that, in the case of the pumping transitions, this effect leads to a slight increase of the pumping time, but with no significant change of the population steady state value.

4.6. Influence of the magnetic field on the optical state preparation and detection.

The physical model leading to the results of Table 1 concerns only atomic populations and neglects laser-induced Zeeman atomic coherences. However, this assumption is not fulfilled when σ -polarized light is used and when the magnetic field $B_o$ is decreased to a value smaller than 100 to 200 milligauss about. When the magnetic field is weak, the laser generates Zeeman coherences in the ground state. They are then connected by the laser radiation to the excited state populations, preventing the excited state population to grow up. Such a version of the non-linear Hanle effect leads to the observation of a minimum of the fluorescence intensity when the magnetic induction is set to zero.

Fig. 6 shows a record of the fluorescence signal as a function of the magnetic induction when the laser is tuned to the (3-3,σ) and (4-4,σ) transitions. The half width of the dip is about 60 mGauss. In this experiment, the laser intensity was 2 mW.cm$^{-2}$. Two consequences clearly appear :

i) hyperfine pumping cannot be completely achieved. It results a diminution of the relative population difference Δn. For instance, in the case of the (4-4, σ) transition a density matrix calculation performed with a laser intensity of 3 mW.cm$^{-2}$ gives Δn = - 11.3 % when $B_o = 0$ instead of the value predicted by the population rate equation model which is equal to - 15.5 % (see Table 1). The reduction amounts to 27 %. It is 20 % in the case of the (3-3,σ) transition.

ii) the fluorescence intensity decreases noticeably which is detrimental to the clock signal amplitude. In the case of the (3-3, σ) transition, when the magnetic induction is varied from 300 to 60 mG, the peak-valley amplitude decreases by a factor of about 1.7, while the background fluorescence light increases very significantly. More detailed results and their quantitative theoretical interpretation will be given in a forthcoming paper [16].

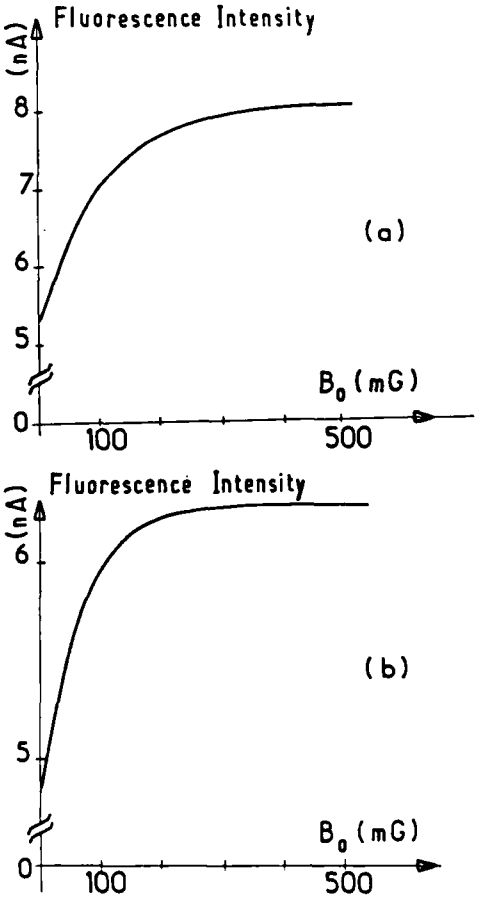



FIGURE 6.
Experimentally recorded variation of the fluorescence intensity versus the magnetic induction applied to the light-atom interaction region. a) for the (3-3,σ) transition and b) for the (4-4, σ) transition. In these experiments we had I = 2 mW.cm$^{-2}$.

## 5. Frequency stability of an experimental optically pumped cesium beam frequency standard

An experimental optically pumped cesium beam frequency standard has been operated with the goal to measure, at first, the short term frequency stability capability of the device.

### 5.1. Cesium beam tube

This is a laboratory device, in which the separation between the oven and the optical detection zone is 57 cm. The oven's collimator consists of a stack of corrugated foils. Its cross-section is 4 x 2 $mm^2$. The oven temperature is set around 100° C. A hot wire detector enables to measure the atomic beam axial intensity. In the conditions of the reported experiment, it is equal to 6.1 x $10^{14}$ $atoms.sr^{-1}.s^{-1}$. The distance between the two arms of the U-shaped microwave cavity is 21 cm. Two cylindrical magnetic shields made of μ-metal surround the Ramsey cavity and the optical interaction regions. Four current carrying rods produce the static magnetic induction.

The light source is a single-frequency Ga-Al-As Hitachi HLP 1400 laser diode operating around 852 nm with an optical power of a few milliwatts. The laser is frequency locked on the 3-3 transition using the fluorescence light emitted in the first optical interaction region. The light is σ polarized and the light power is about the same (1 mW) in the two interaction regions.

Light-atom interaction zones are identical and designed in order to minimize the scattered laser light. In the optical detection zone, the beam cross section is 10 x 5 $mm^2$. The area of the silicon photodiode is large, equal to 1 $cm^2$. Its noise level is low, the NEP being 2.5 x $10^{-14}$ $W.Hz^{-1/2}$. The photon collection efficiency is 0.15. The flux of stray photons collected is 2.5 x $10^7$ times smaller than the photon flux in the laser beam crossing the detection zone.

Fig. 7 shows the Ramsey pattern obtained. 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 defined as the peak to valley current difference divided by the standard deviation of the peak current is equal to 10,000 in a 1 Hz noise bandwidth.

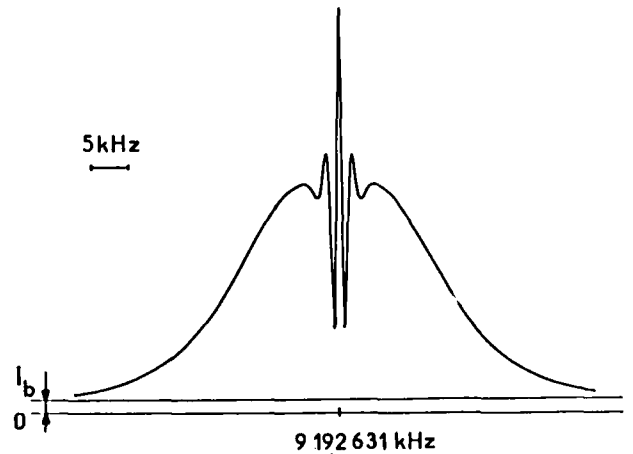



Fig.7

Experimentally recorded response of a single laser optically pumped cesium beam tube.

### 5.2. Frequency control electronic system

Possible designs for the generation of the interrogation signal and for the processing of the beam tube response are being studied. The goal is to investigate means to take full advantage of the short to long term frequency capability of optically pumped cesium beam tubes.

Fig. 8 shows a simplified block-diagram of the frequency control electronic system used in the present work. It is described in more details in reference 14.

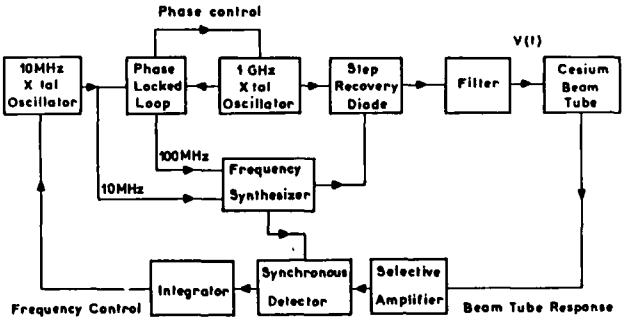



Fig.8

Simplified block-diagram of the frequency control electronic system associated with the optically pumped cesium beam tube.

The quartz crystal oscillator to be frequency locked to the cesium hyperfine transition is a high quality 10 MHz QAS quartz crystal oscillator. A 1 GHz quartz crystal oscillator is phase locked to the 10 MHz oscillator. The loop includes frequency dividers. There is no frequency multiplier used in the 10 MHz to 1 GHz part of the chain. This ensures :
i) the absence of any sideband at 10 MHz and at its harmonics and
ii) the easiness of the implementation which is mainly due to the absence of tuned circuits. The frequency of the 1 GHz signal is multiplied up to 9 GHz by means of a step recovery diode.

The frequency synthesizer generates a square-wave frequency modulated signal around 192.6 MHz, which is mixed to the 9 GHz signal. We have deliberately chosen a large value of the additional frequency in order to render negligible the frequency shift -and thus its possible variation related to a power change- due to unbalanced spurious frequency components in the microwave interrogation spectrum [24]. In our set-up, this frequency shift is smaller than 1 x $10^{-14}$. This contrasts with the value of this shift in some manufactured cesium beam frequency standards where the level of the 12.6 MHz or of the 7.4 MHz is such that the frequency shift may amount up to $10^{-12}$.

Square wave frequency modulation has been chosen since it is favorable to the achievement of a good frequency stability of the controlled source [25, 26] and since it avoids the use of an analog phase modulator which might introduce frequency shifts in the case of even harmonic distorsion of the modulation waveform [27]. The amplitude of modulation is equal to 190 Hz and its frequency is equal to 95 Hz. Two phase locked oscillators generate continously fixed frequencies separated by 380 Hz around 7.4 MHz. Their output signals are multiplexed in a logic gate, then mixed to a 200 MHz

signal. The lower sideband obtained is properly filtered.

It is known that amplitude modulation of the interrogation signal at odd harmonics of the modulation frequency may be the source of a false error signal in the frequency control loop [27] . In order to avoid such an effect, it is necessary to ensure that the amplitude of the interrogation signal is the same during the two half periods of modulation. For that purpose, the amplitude changes of the interrogation signal are synchronously detected, using a signal at the modulation frequency as a reference. Then, the error signal controls the amplitude of the output of one of the two sources at 7.4 MHz. The residual fractional amplitude modulation is approximately equal to $2 \times 10^{-6}$. Moreover, a number of frequency offsets such as, for instance, those related to the second order Doppler effect, the cavity phase shift, the cavity mistuning, the spurious spectral components and the imperfections in the electronics depend on the microwave interrogation level. A 1 % variation of this level gives a fractional frequency offset change of the order of $10^{-14}$. A control of the microwave level is thus desirable. It is implemented by detecting the level of the 9.2 GHz signal and by acting upon the polarization current of the step recovery diode. A level stability of 0.2 % per degree Celsius is achieved.

The power available at 9.2 GHz is 0.3 mW. Fig. 9 shows the spectrum of the interrogation signal. Spurious components due to electronic imperfections are at least 60 dB below the signal power. The related fractional frequency shift is smaller than $1 \times 10^{-14}$.

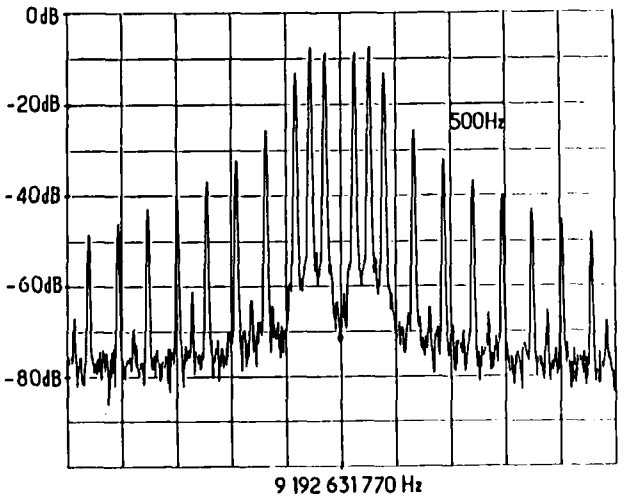



Fig. 9

Power spectrum of the square wave frequency modulated microwave interrogation signal. The level referred as 0 dB is that of the signal without modulation.

At present, the processing of the cesium beam response is made classically. The first harmonic of the response is selectively amplified and synchronously detected at 95 Hz. The error signal feeds an integrator whose output voltage is applied to the frequency control varactor of the 10 MHz quartz crystal oscillator. The value of the loop time constant is set to 0.2 s.

5.3. Measured short term frequency stability

The frequency stabiity of the frequency controlled 10 MHz quartz crystal oscillator has been measured with respect to a hydrogen maser. Fig. 10 shows the Allan standard deviation obtained. The result is :

$$\sigma_y(\tau) = 2 \times 10^{-12}/\sqrt{\tau} \quad , \qquad (10)$$

for $1 \text{ s} < \tau \lesssim 500$ s. For $\tau > 500$ s, systematic effects due to imperfections of the microwave cavity and of the magnetic shield of the set up in its present version limit the achievable frequency stability.

To our knowledge this measured short term frequency stability is the best obtained to date with an optically pumped cesium beam frequency standard. It confirms the interest of this kind of frequency standard.

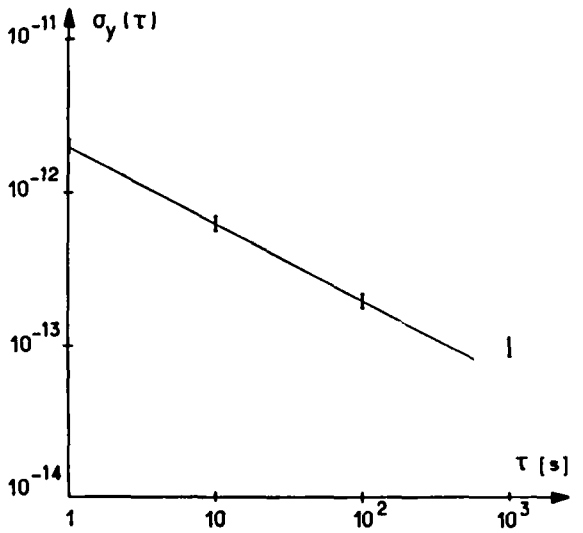



Fig. 10

Measured short term frequency stability of an experimental optically pumped cesium beam frequency standard.

## 6. Prospects for improvement of the frequency stability

An estimate of the frequency stability, which is more accurate than that given by Eq. (1) is the following [18] :

$$\sigma_y(\tau) = F^{1/2} \frac{(2I'_b + I_{pv})^{1/2}}{2Q_\ell \, I_{pv} \, \tau^{1/2}} \quad , \qquad (11)$$

where $I_{pv}$ is the peak to valley flux of detected atoms. In our experimental set up, we have $I_{pv} = 6.4 \times 10^9 \text{ s}^{-1}$ and $Q_\ell = 1.8 \times 10^7$. The amplificator noise factor is approximately equal to 6 or 7 instead of the theoretical value of 1.8. This gives $I'_b = 3.9 \times 10^{10} \text{ s}^{-1}$ and $\sigma_y(\tau) \simeq 2.2 \times 10^{-12}/\sqrt{\tau}$, in agreement with the measured value.

6.1. Improvement of the frequency stability achieved

Equations (11), (4) and (5) enable to discuss the possible improvement of the frequency stability achieved. In the present discussion, it is assumed that the total flux of atoms crossing the detection region is kept unchanged.

1°) The noise factor of the current to voltage converter connected at the photocell output is at least 3 times its theoretical value. It follows that a gain of a factor 1.6 on $\sigma_y(\tau)$ can be achieved.
2°) If the frequency stability were limited by the beam shot noise, i.e. with F = 1 and $I'_b \ll I_{pv}$ in Eq. (11), the frequency stability would be improved by a factor of 3.6. As shown by Eqs. (4) and (5), this can be obtained by increasing sufficiently the mean photon yield $\beta$, up to values of several tens. Such values can be achieved by using a cycling transition for the optical detection, provided that the laser spectral purity is sufficiently enhanced.
3°) Assuming two-laser optical pumping, the flux of atoms in state F = 3, $m_F = 0$ or F = 4, $m_F = 0$ is increased by a factor of almost 8. It follows an improvement of the frequency stability by a factor almost equal to 2.8 in the case $I'_b \ll I_{pv}$.

We thus see that improvement of the frequency stability by a factor of the order of 10 is theoretically possible.

### 6.2. Scale effects

Equations (11), (4) and (5) can also be used to look at scale effects.
1°) If the flux of detected atoms is changed, $\sigma_y(\tau)$ varies as $I_{pv}^{-1}$ for $I'_b \gg I_{pv}$ and as $I_{pv}^{-1/2}$ for $I'_b \ll I_{pv}$.
2°) If the separation L between the two microwave interaction regions and if the separation between the oven and the detection zone are scaled by the same factor, whereas the axial beam intensity is kept constant, then $\sigma_y(\tau)$ does not depend on L in the case $I'_b \ll I_{pv}$, but varies as L in the case $I'_b \gg I_{pv}$. The latter result means that, in the present state of the art, where $I'_b$ is larger than $I_{pv}$, the frequency stability of an optically pumped cesium beam tube is improved when its overall length is decreased, the beam intensity beeing kept constant.

We conclude from Sections 6.1. and 6.2. that a frequency stability of a few $10^{-13}$ for $\tau$ = 1 s can be achieved in the usual size of a manufactured cesium beam frequency standard.

### 6.3. Work on extended-cavity semi conductor lasers

Semi conductor lasers are compact, reliable and easy to tune devices. However, they have drawbacks : i) it exists dead wavelength ranges, located at unpredictable places which may vary under aging effects, which cannot be reached by either temperature or current tuning, ii) their linewidth is larger than 10 MHz, which causes a large noise of the fluorescence signal when they are used to excite cycling transitions and iii) they are very sensitive to optical feedback, i.e. to external parasitic reflections of the laser beam into the diode cavity, which perturb the lasing frequency and may enhance frequency noise. Although it is known that controlled optical feedback can greatly reduce the laser linewidth [28,29], this method cannot remove the drawback i) above. We have thus chosen to study extended-cavity semi conductor lasers [30] which can oscillate at any frequency of the diode gain spectral band and which provide a sufficient improvement of the linewidth. These lasers can be designed compact and easy to operate. They are very well suited for the optical detection of cesium atoms [31].

The linewidth of semi conductor lasers, mainly limited by spontaneous emission phase noise, is inversely proportional to the square of the laser optical cavity length L. This remark leads to the concept of the extended cavity laser. Fig. 11 shows the principle of this device.

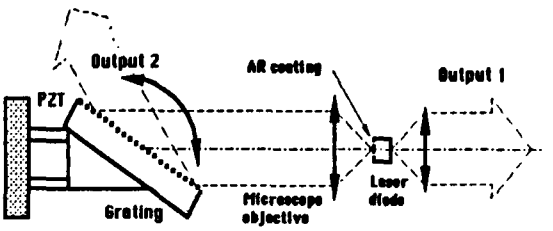



FIGURE 11

Principle of an extended-cavity semi conductor laser.

One of the diode cleaved mirrors is suppressed. The light beam emerging from the anti-reflection coated facet is collimated by a large aperture microscope objective. The cavity is ended by a Littrow-mounted diffraction grating. The laser wavelength is imposed by the grating. Its rotation enables to get any wavelength in the diode gain spectral band. A tuning range larger than 30 nm around 850 nm has been obtained. Fine tuning is provided by mounting the grating on a piezoelectric translator. In this arrangements the laser is no longer subject to mode hops after aging of the semi conductor material. Spectral analysis of the beat note between two such lasers gives an instantaneous laser linewidth between 10 to 50 kHz, which is convenient for the optical detection of cesium atoms. It has been verified experimentally that the extended cavity semi conductor laser is much less sensitive to optical feedback than ordinary laser diodes.

The decrease of the fluorescence noise when an extended cavity semi conductor laser is used to optically detect cesium atoms has been observed experimentally, as shown in Fig. 12.

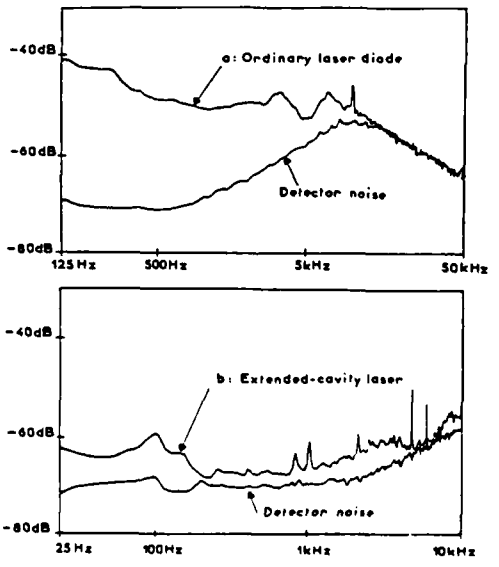



FIGURE 12

Spectral analysis of the fluorescence noise for a laser locked to the top of a cesium beam line a) ordinary laser diode, b) extended-cavity laser.

## 7. Conclusion

We have demonstrated experimentally, with the simplest possible set-up using a single laser, that optical methods applied properly to the cesium beam frequency standard can improve the short term frequency stability by an order of magnitude. Discussion of possible improvements show that another order of magnitude may be gained in a standard size cesium beam frequency standard. This can be obtained by i) a complete state preparation using two lasers and ii) an optical detection by means of a cycling transition excited by an extended-cavity semi conductor laser, for instance.


## Acknowledgements

The authors are grateful to CEPE who supplied the quartz crystal oscillators and filters used in this work. They are indebted to Vincent Candelier whose contribution was determinant in the design of the frequency control system and in the measurement of the frequency stability.


## References


[1]. A. Kastler, Journal de Physique et le Radium, 11 (1950), 255

[2]. P. Cérez, M. Arditi and A. Kastler, Comptes-Rendus Académie des Sciences, Paris, 267B (1968), 282 ; M. Arditi and P. Cérez, IEEE Transactions on Instrumentation and Measurement IM-21 (1972), 391

[3]. M. Arditi and J.L. Picqué, Journal de Physique - Lettres 41 (1980), L-379

[4]. G. Avila, E. de Clercq, M. de Labachelerie and P. Cérez, IEEE Trans. on Instr. and Meas. IM-34 (1985), 139

[5]. Shang Song-quan, Wu Xin-xin, Yao Shu-tong, Xie Linzhen and Wan Yi-qiu, Chinese Physics Letters 2 (1985), 557

[6]. A. Derbyshire, R.E. Drullinger, M. Feldman, D.J. Glaze, D. Hilliard, D.A. Howe, L.L. Lewis, J.H. Shirley, I. Pascaru and D. Stanciulescu, Proc. of the 39th Annual Symposium on Frequency Control, Philadelphia, USA (1985), p. 18

[7]. S. Oshima, Y. Nakadan and Y. Koga, private communication (1987)

[8]. R.E. Drullinger, J. Shirley, D.J. Glaze and L.W. Hollberg, Proc. of the 40th Annual Symposium on Frequency Control, Philadelphia, USA (1986), p. 428

[9]. T. Mc Clelland, I. Pascaru, J. Zacharski, N.H. Tran and M. Meirs, Proc. of the 41st Annual Symposium on Frequency Control, Philadelphia, USA (1987), p. 59

[10]. V. Giordano, A. Hamel, G. Théobald, P. Cérez, C. Audoin and V. Candelier, Metrologia 25 (1988), 17

[11]. J. Umezu, B. Komiyama, H. Saitho and Y. Ohta, Proc. of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988). To be published

[12]. P. Thomann, H. Schweda and G. Busca, Proc. of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988). To be published

[13]. E. de Clercq, A. Clairon and B. Dahmani, Proc. of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988). To be published

[14]. V. Candelier, V. Giordano, A. Hamel, G. Théobald, P. Cérez and C. Audoin, Proc. of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988). To be published

[15]. C. Jacques and P. Tremblay, Proc. of the 42nd Annual Frequency Control Symposium, Baltimore, USA (1988), this issue

[16]. G. Théobald, P. Cérez and V. Giordano, Proc. of the Fourth Symposium on Frequency Standards and Metrology, Ancona, Italie (1988). To be published

[17]. G. Avila, V. Giordano, V. Candelier, E. de Clercq, G. Théobald and P. Cérez, Physical Review A 36 (1987), 3719

[18]. J. Vanier and C. Audoin in "The Quantum Physics of Atomic Frequency Standards". Adam Hilger, Bristol, UK, 1989, in press

[19]. A. de Marchi, G.D. Rovera and A. Premoli, Metrologia, 20 (1984), 27

[20]. A. de Marchi, IEEE Trans. on Ultrasonics, Ferroelectronics and Frequency Control, UFFC-34 (1987), 598

[21]. A. Brillet, Metrologia, 17 (1981), 147

[22]. J.H. Shirley, Proc. of the 39th Annual Symposium on Frequency Control, Philadelphia, USA (1985), 22

[23]. V. Giordano, V. Candelier, A. Hamel, C. Audoin, G. Théobald and P. Cérez, Optics Communications (1988), in press

[24]. C. Audoin, M. Jardino, L.S. Cutler and R.F. Lacey, IEEE Trans. on Instr. and Meas. IM-27 (1978), 325

[25]. C. Audoin, V. Candelier and J. Vanier, IEEE Trans. on Ultrasonics, Ferroelectronics and Frequency Control, UFFC-34 (1987), 573

[26]. A. de Marchi, G.D. Rovera and A. Premoli, IEEE Trans. on Ultrasonics, Ferroelectronics and Frequency Control, UFFC-34 (1987), 582

[27]. F.L. Walls, IEEE Trans. on Ultrasonics, Ferroelectronics and Frequency Control, UFFC-34 (1987), 592

[28]. B. Dahmani, L. Hollberg and R. Drullinger, Optics Letters 12 (1987), 876

[29]. A. Clairon, B. Dahmani, Ph. Laurent and Ch. Bréant, Proc. of the 2nd European Forum on Time and Frequency, Neuchâtel, Switzerland (1988). To be published

[30]. M.W. Fleming and A. Mooradian, IEEE Journal of Quantum Electronics, QE-17 (1981), 44

[31]. M. de Labachelerie, K. Diomandé and N. Dimarcq, Proc. of the 2nd European Frequency and Time Forum, Neuchâtel, Switzerland (1988). To be published

42nd Annual Frequency Control Symposium - 1988

# CALCULATIONS ON THE EFFICIENCY OF OPTICAL PUMPING OF A CESIUM ATOMIC BEAM BY LASERS OF FINITE LINEWIDTH

C. Jacques
Laboratory for Basic Standards
Division of Physics
National Research Council
Ottawa, Canada K1A 0R6

and

P. Tremblay
Laboratoire de recherches sur les oscillateurs et systèmes
Département de Génie électrique
Université Laval
Cité universitaire
Québec, Canada G1K 7P4

## Abstract

The calculations of the efficiency of the optical pumping in the cesium ground state hyperfine levels by means of lasers have been done until now with rate equations which do not include coherence effects. On the other hand, a laser with a finite linewidth tuned to a particular transition excites as well neighboring transitions. This paper presents a development of rate equations with coherence in which the effects of laser spectrum wings and linewidth are included. We evaluate the optical pumping efficiencies in the cesium ground state hyperfine levels when two lasers pumping schemes are used under different conditions of polarization, tuning, laser linewidth and power densities.

## Introduction

Preparation of a cesium atomic beam for a frequency standard by optical pumping has been discussed by various authors[1,2] and many experiments have been done. The advantages of optical pumping in a cesium atomic frequency standard are well known. It should increase the signal to noise ratio, neglecting possible spurious effects in the pumping and the detection regions, by preparing all the cesium atoms in one level, either 6s $^2S_{1/2}$ F=(3,0) or F=(4,0). It reduces Rabi pulling. By replacing the selecting magnets, it enormously reduces the asymmetry of the beam, and the possibility of Majorana transitions due to leaks of the strong magnetic field into the C field region (Fig.1).

Nevertheless, many points have still to be investigated thoroughly, especially the question of time and efficiency of optical pumping in relation to the power, linewidth and polarization of the lasers used. Previously, two theoretical models have been used: one based on a phenomenological approach giving rise to rate equations without including the effects of coherence between states, approach valid for low intensity and broadband optical pumping, and the approach including coherence. The latter has been applied only in the case of monochromatic radiation from the lasers. There is, however, an intermediate zone, both in laser intensity and linewidth which has not been studied. The interest of the study lies in the fact that there is an important discrepancy between the two models for monochromatic lasers where the rate equations without coherence give a value of nearly 100% for the efficiency of pumping for some choices of tuning and polarizations, while the rate equations with coherence give 14%. On the other hand, with a laser linewidth of the order of a few tens of MHz, the much simpler model without coherence works very well. The use of laser diodes locked to external cavities and having linewidths less than 1 MHz presents a theoretical challenge in the sense that it is difficult to evaluate their pumping efficiency. Furthermore, a laser with a finite linewidth tuned to a particular transition will excite neighboring transitions by the wings of these transitions and that of the laser spectrum. The wings have low intensities, but with time the excitation to neighboring levels has a significant effect on the efficiency of pumping.

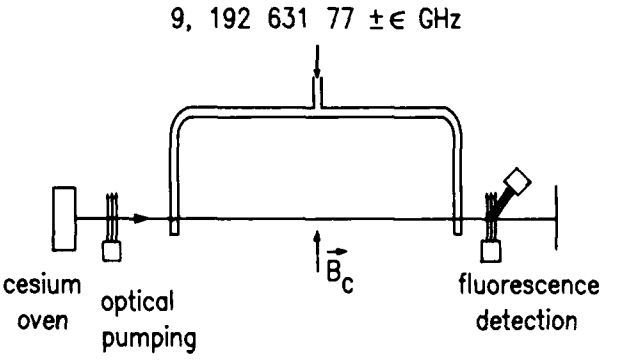



Fig. 1 Schematic representation of an optically pumped cesium beam frequency standard.

In this paper we will describe a development of rate equations with coherence in which we include the effects of laser spectrum wings and linewidth by means of fluctuating laser fields. This formalism gives the same results with the same initial conditions as the previous two models, with coherence for monochromatic lasers and without coherence for low intensity lasers having large linewidths, where they respectively apply.

## Theory

The rate equations without coherence take into account the usual three processes related to the matter-light interaction: spontaneous emission $A_{eg}$, absorption $W_{ge}$ and stimulated emission $W_{eg}$ between an upper level e and a lower level g. In the case of the cesium atom, g represents the first hyperfine level of the ground state 6s $^2S_{1/2}$ F=3, f represents the second hyperfine level 6s $^2S_{1/2}$ F=4 and e represents the excited state 6p $^2P^0_{3/2}$ (Fig.2).

CH2588-2/88/0000-505 $1.00 C 1988 IEEE

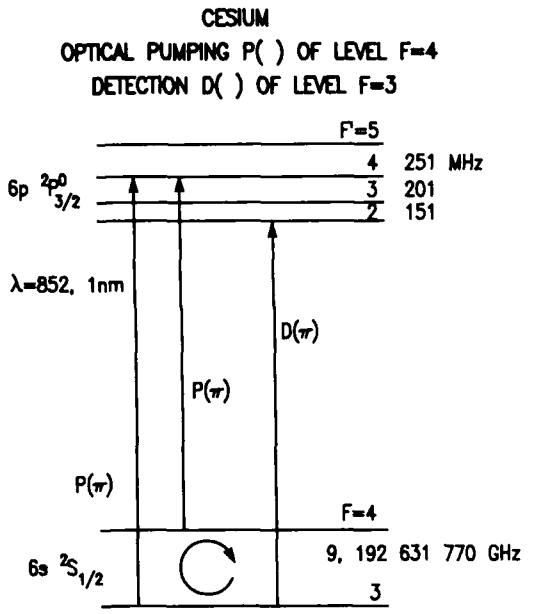



Fig. 2 Energy level diagram of the cesium atom.

The set of rate equations without coherence characterizing the time evolution of the populations of each Zeeman sublevel can thus be written as:

$$\frac{dn_{e_i}}{dt} = \sum_j W_{f_j e_i} n_{f_j} + \sum_k W_{g_k e_i} n_{g_k} - \left(\sum_j W_{e_i f_j} + \sum_k W_{e_i g_k}\right) n_{e_i} - \left(\sum_j A_{e_i f_j} + \sum_k A_{e_i g_k}\right) n_{e_i} \quad (1a)$$

$$\frac{dn_{f_j}}{dt} = -\sum_i W_{f_j e_i} n_{f_j} + \sum_i W_{e_i f_j} n_{e_i} + \sum_i A_{e_i f_j} n_{e_i} \quad (1b)$$

$$\frac{dn_{g_k}}{dt} = -\sum_i W_{g_k e_i} n_{g_k} + \sum_i W_{e_i g_k} n_{e_i} + \sum_i A_{e_i g_k} n_{e_i} , \quad (1c)$$

with

$$W_{er} = \frac{3\lambda_{er}^3}{8\pi ch} A_{er} \frac{\Gamma + \Gamma_\ell}{\left(\frac{\Gamma + \Gamma_\ell}{2}\right)^2 + (\omega_{er} - \omega_\ell)^2} \quad (1d)$$

where

$\lambda_{er}$ = wavelength between levels e and r (f or g),

$\Gamma$ = natural linewidth,

$\Gamma_\ell$ = linewidth of the laser,

$\omega_{er}$ = angular frequency between levels e and r (f or g)

$\omega_\ell$ = angular frequency of the laser.

The equation (1d) contains a Lorentzian which permits the evaluation of the wing effect for the neighboring transitions as well as the detuning of the laser around the main transition.

The rate equations with coherence is briefly outlined below. The starting point is the equation of the evolution of the density matrix with time[3]

$$\frac{d\rho(t)}{dt} = \frac{1}{i\hbar}\left[ H(t), \rho(t) \right] + R(t) \quad (2)$$

where $\rho_{ii}(t)$ = population of level i

$\rho_{ij}(t)$ = coherence between levels i and j

$H(t)$ = Hamiltonian

$R(t)$ = relaxation related and spontaneous emission terms

The Hamiltonian H(t) is the sum of the unperturbed Hamiltonian $H_o$ and the interaction term V(t), which is

$$V(t) = -\vec{D} \cdot \vec{E}(t) \quad (3)$$

where $\vec{D}$ is the electric dipole moment and $\vec{E}(t)$ is the total electric field.

The electric field for one laser is expressed as

$$\vec{E}(t) = \vec{e}_\ell E_\ell \{1 + \varepsilon_\ell(t)\} \cos\left\{ \omega_\ell + \phi_\ell(t) + \theta_\ell \right\} , \quad (4)$$

where $\vec{e}_\ell$ = polarization vector,

$E_\ell$ = intensity of electric field,

$\varepsilon_\ell(t)$ = relative amplitude fluctuations,

$\phi_\ell(t)$ = phase fluctuations,

$\theta_\ell$ = phase.

$\varepsilon(t)$ and $\phi(t)$ are zero mean stochastic processes. We assume that they are wide-sense stationary and mean square continuous.

This electric field could be expressed in the following manner,

$$\vec{E}(t) = \frac{\vec{e}_\ell E_\ell}{2}\left[ \zeta_\ell(t)\, e^{i\omega_\ell t} + \zeta_\ell^*(t)\, e^{-i\omega_\ell t} \right] , \quad (5)$$

where $\zeta_\ell(t) = [1+\varepsilon_\ell(t)]\, e^{i(\phi_\ell(t) + \theta_\ell)}$

The function $\zeta_\ell(t)$ represents all the fluctuations of the electric field.

We assume that the power spectral density, which is the Fourier transform of the autocorrelation function, $G_{\zeta_\ell}(\tau)$, of each laser is of Lorentian shape. This means that

$$G_{\zeta_\ell}(\tau) = e^{-\frac{1}{2}\Gamma_\ell|\tau|} . \quad (6)$$

Since we use two lasers, $\vec{E}(t) = \vec{E}_1(t) + \vec{E}_2(t)$.

We substitute equations (3) and (5) in equation (2) and make the following change of variables:

$$\rho_{eg}(t) = \delta_{eg}(t)\, e^{-i\omega_1 t} \qquad (7a)$$

$$\rho_{ef}(t) = \delta_{ef}(t)\, e^{-i\omega_2 t} \qquad (7b)$$

$$\rho_{fg}(t) = \delta_{fg}(t)\, e^{-i(\omega_1 - \omega_2)t} \qquad (7c)$$

Since $\zeta_1(t)$ and $\zeta_2(t)$ are stochastic processes, they cannot be evaluated in a deterministic manner. As a consequence, $\rho(t)$ is also a stochastic process, and in that sense one must look at its statistical properties. We are interested in the mean of the populations of each Zeeman sublevel. Thus we must consider the following values: $\langle\rho_{ii}(t)\rangle$, where i means e, f or g, $\langle\zeta_1(t)\delta_{eg}(t)\rangle$, $\langle\zeta_2(t)\delta_{ef}(t)\rangle$ and $\langle\zeta_1(t)\zeta_2^*(t)\delta_{fg}(t)\rangle$.

We obtain for these variables equations characterizing their time evolution. These equations are similar to those obtained for the monochromatic approach of the density matrix formalism, where we make the following substitutions,

$\langle\rho_{ii}(t)\rangle$ for $\rho_{ii}(t)$,

$\langle\zeta_1(t)\,\delta_{eg}(t)\rangle\, e^{-i\omega_1 t}$ for $\rho_{eg}(t)$,

$\langle\zeta_2(t)\,\delta_{ef}(t)\rangle\, e^{-i\omega_2 t}$ for $\rho_{ef}(t)$,

and

$\langle\zeta_1(t)\,\zeta_2^*(t)\,\delta_{fg}(t)\rangle\, e^{-i[\omega_1-\omega_2]t}$ for $\rho_{fg}(t)$.

The only changes appear in R(t),

$$R_{ee'} = -\Gamma\,\langle\rho_{ee'}\rangle\,, \qquad (8a)$$

$$R_{eg} = -\tfrac{1}{2}\,[\,\Gamma + \Gamma_1\,]\,\langle\zeta_1\delta_{eg}\rangle\, e^{-i\omega_1 t} \qquad (8b)$$

$$R_{ef} = -\tfrac{1}{2}\,[\,\Gamma + \Gamma_2\,]\,\langle\zeta_2\delta_{ef}\rangle\, e^{-i\omega_2 t} \qquad (8c)$$

$$R_{fg} = -\tfrac{1}{2}[\,\Gamma_1 + \Gamma_2\,]\,\langle\zeta_1\zeta_2^*\delta_{fg}\rangle\, e^{-i(\omega_1-\omega_2)t} \qquad (8d)$$

$$R_{gg'} = \Gamma \sum_{\substack{ee' \\ m_e - m_{e'} = m_g - m_{g'}}} a_{eg}\,\langle\rho_{ee'}\rangle\, a_{e'g'} \qquad (8e)$$

and

$$R_{ff'} = \Gamma \sum_{\substack{ee' \\ m_e - m_{e'} = m_f - m_{f'}}} a_{ef}\,\langle\rho_{ee'}\rangle\, a_{e'f'} \qquad (8f)$$

where

$$a_{er} = (-1)^{1+I+J_e+F_e+F_r-m_e}\,\sqrt{2F_e+1}\,\sqrt{2F_r+1}\,\sqrt{2J_e+1} \times \begin{pmatrix} F_e & 1 & F_r \\ -m_e & q & m_r \end{pmatrix} \times \begin{Bmatrix} F_e & 1 & F_r \\ J_r & I & J_e \end{Bmatrix} \quad (r=f \text{ or } g)$$

## Results

The calculations entail solving numerically sets of linear differential equations. For the time evolution calculation we used the program RKF45, published by G.E. Forsythe et al[4]. The asymptotic state was evaluated by inverting the matrix of constant coefficients of the set of differential equations using the programs DGECO and DGESL by Dongarra et al[5].

We have studied different optical pumping schemes that use two lasers, since it appears that this is the most efficient way to maximise the optical pumping efficiency - the population difference between the two hyperfine levels F=(3,0) and F=(4,0) of the ground state 6s $^2S_{1/2}$ of cesium - neglecting possible spurious effects. The results of the time evolution of the population difference are shown in fig. 3 and fig. 4. The curves in fig. 3 correspond to rate equations without coherence and with wing effect. They are almost similar to what Avila et al[2] have calculated. The most drastic changes appear in fig. 4 when we add the coherence with the density matrix model. The efficiency of optical pumping decreases a little for the (3-4π, 4-4π)* scheme, but decreases drastically for all the other schemes which include σ polarizations. This means that, even with a laser linewidth of 20 MHz, the coherence between the levels induced by the laser radiation is present and that it has an important effect in σ polarization. Because of these results, our calculations have been concentrated on the (3-4π, 4-4π) scheme since it appears to be the most promising.

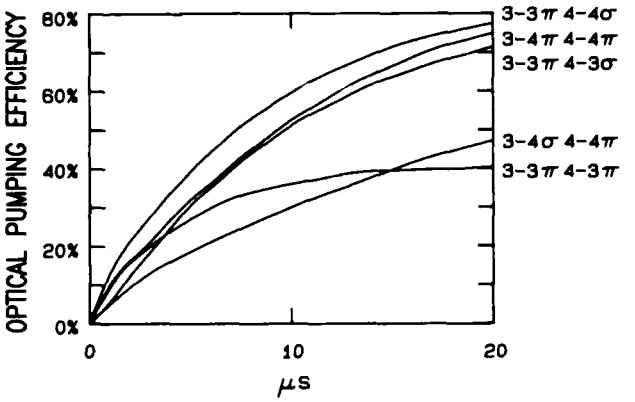



Fig. 3 Time evolution of the optical pumping efficiency as a function of the pumping scheme. The calculation is made with the rate equations without coherence and with wing effect; the laser linewidth is 20 MHz at a power density of 3 mW/cm$^2$.

* The notation (F-F'$p_1$, F-F' $p_2$) means the choice of the hyperfine levels, F for 6s $^2S_{1/2}$ and F' for 6p $^2P^0 3/2$, and polarization $p_1$ and $p_2$ to which the two lasers are tuned and polarized.

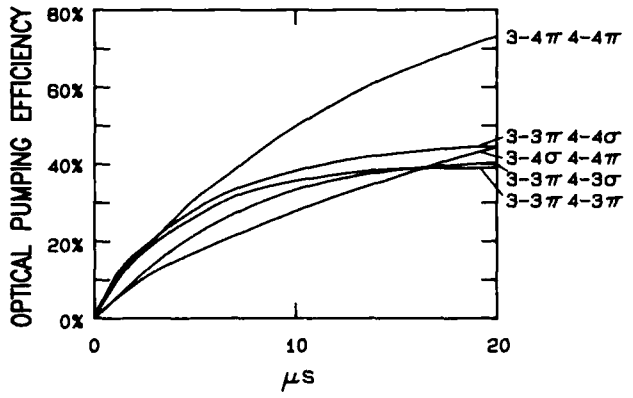



Fig. 4 Time evolution of the optical pumping efficiency as a function of the pumping scheme. The calculation is made with the rate equations with coherence and wing effect; the laser linewidth is 20 MHz at a power density of 3 mW/cm$^2$.

There are two ways to induce more coherence between the levels with laser radiation: higher power density or narrower linewidth. Fig. 5 illustrates the effect of coherence on the time evolution of the efficiency of optical pumping for a power density of 10 mW/cm$^2$ and laser linewidth of 300 kHz. The population inversion grows rapidly due to the power and the concentration of that power near the center line of the transition. But, as soon as the coherence builds up, the population inversion suddenly slows down.

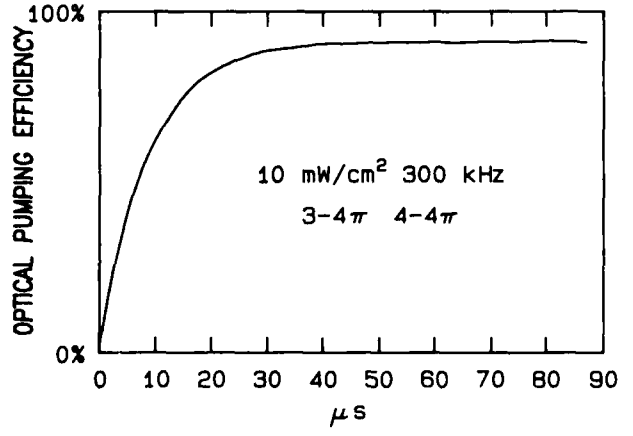



Fig. 5 Time evolution of the optical pumping efficiency. The calculation is made with the rate equations with coherence and wing effect. Both lasers are tuned and polarized (3-4π 4-4π); the laser linewidth is 300 kHz at a power density of 10 mW/cm$^2$.

This effect is most remarkable in Fig. 6, at 10 mW/cm$^2$ and with monochromatic lasers, where the population inversion grows very rapidly at the beginning. Fig. 7 shows the left part of the growing curve with the slowing down at an efficiency of 14%. This last result is interesting because the level of population difference at 14% is the maximum attainable with the density matrix formalism without the wing effect. The excitation to the neighboring levels help in building the population difference up to 46%. At a narrow laser linewidth, the wing effect increases the optical pumping efficiency as time goes on.

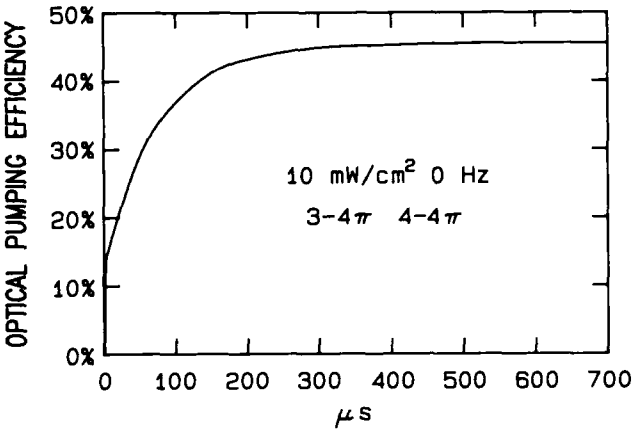



Fig. 6 Time evolution of the optical pumping efficiency. The calculation is made with the rate equations with coherence and wing effect. Both lasers are tuned and polarized (3-4π 4-4π); the lasers are monochromatic at a power density of 10 mW/cm$^2$.

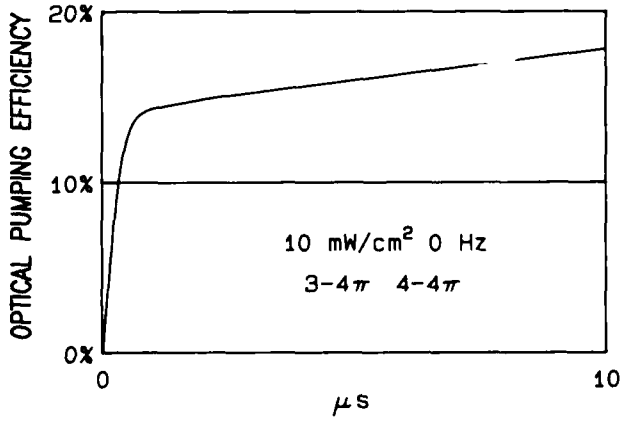



Fig. 7 First 10 μs of Fig. 6

We have studied the details of the dependence of the asymptotic optical pumping efficiency on the laser linewidth. In fig. 8, the two models, without and with coherence are compared.

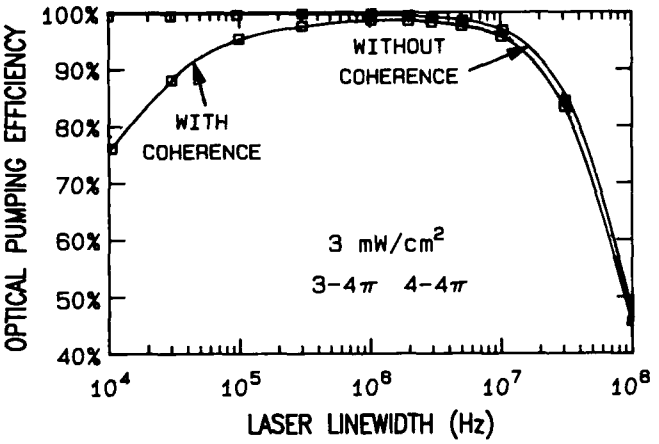



Fig. 8 Asymptotic optical pumping efficiency as a function of laser linewidth for the two models; with and without coherence. Both lasers are tuned and polarized (3-4π 4-4π) at a power density of 3mW/cm$^2$.

Because of the wing effect, it is impossible to reach 100% asymptotic efficiency. With narrow laser linewidths, the coherence decreases the efficiency, while the wing effect does the same for wide linewidths of more than 10 MHz. The fact that the two models converge for linewidths larger than 10 MHz indicates the decreasing importance of the coherence effect, which is expected.

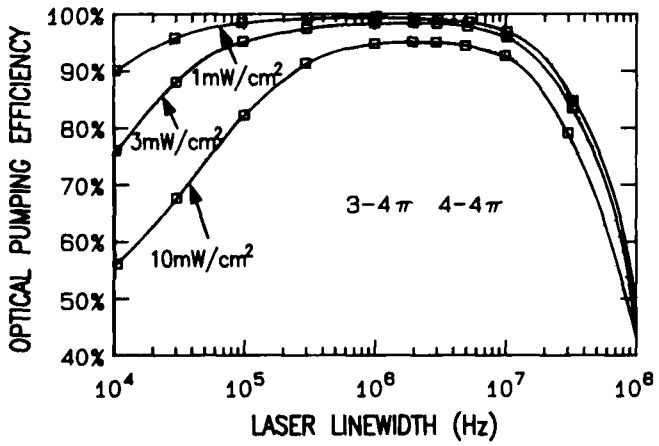



Fig. 9 Asymptotic optical pumping efficiency as a function of laser linewidth for the rate equations with coherence for various power densities of 1, 3, and 10 mW/cm$^2$. Both lasers are tuned and polarized (3-4π 4-4π).

Fig. 9 compares the optical pumping efficiency versus laser linewidth at different power densities of 1, 3 and 10 mW/cm$^2$. It is interesting to note that the increase of power density corresponds not only to an increase in coherence at narrow linewidths, but also to an increase in wing effects at larger linewidths, thus to less optical pumping efficiency. These curves correspond to the asymptotic state. If the time of optical pumping is limited, it is expected that the shapes of these curves would be different. For a finite pumping time, the optimum conditions both in power and linewidth for maximum efficiency could be calculated using our model.

## Conclusions

A development of rate equations with coherence in which the effects of laser spectrum wings and linewidth are included has been presented. This formalism is equivalent to rate equations with coherence for monochromatic lasers and to rate equations without coherence for lasers having large linewidths relative to the natural linewidth of the involved atomic transitions. This corresponds to laser linewidths of more than 10 MHz, relatively to the 5 MHz linewidth of the cesium transition $6s\ ^2S_{1/2} - 6p\ ^2P^o_{3/2}$. We were able to evaluate optical pumping efficiencies for lasers of intermediate linewidth where coherence and wing effects are both present. These two effects decrease the optical pumping efficiency, the coherence for narrow laser linewidth and wing effect for large laser linewidth. In fact, the calculations confirmed that the most efficient laser linewidths are slightly smaller than the natural linewidth of the atomic transition. We also found that the effect of coherence is important when one laser is σ polarized in the two laser optical pumping schemes; thus, it is preferable to avoid σ polarization in order to have a better optical pumping efficiency.

## Acknowledgement

The authors wish to thank Dr. Jacques Vanier for his numerous and helpful comments.

## References

1) E. de Clercq, M. de Labachellerie, G. Avila, P. Cerez and M. Têtu, Laser diode optically pumped cesium beam, J. Physique 45 (2), 239-47 (1984).

2) G. Avila, V. Giordano, V. Candelier, E. de Clercq, G. Theobald, and P. Cerez, State selection in a cesium beam by laser-diode optical pumping, Phys. Rev. A 36(8), 3719-28 (1987).

3) C. Cohen-Tannoudji, Atoms in strong resonant fields, Frontiers in laser spectroscopy, Vol. 1, Les Houches, Session XXVII (1975), edts. R. Balian, S. Haroche, and S. Liberman, North-Holland, Amsterdam (1977).

4) G.E. Forsythe, M.A. Malcolm, and C.B. Moler, Computer Methods for Mathematical Computations, Prentice-Hall Inc., Englewood Cliffs, (1977).

5) J.J. Dongarra, C.B. Moler, J.R. Bunch, and G.W. Stewart, LINPACK User's Guide, SIAM, Philadelphia, (1979).

42nd Annual Frequency Control Symposium - 1988

# CESIUM GETTERING BY GRAPHITE AND COLLOIDAL GRAPHITE (AQUADAG)

N. D. Bhaskar and C. M. Kahla
Chemistry and Physics Laboratory, The Aerospace Corporation
P. O. Box 92957, Los Angeles, California 90009

## SUMMARY

Solid blocks of graphite are placed inside cesium (Cs) beam tubes (CBT) to getter the background Cs vapor. Additional gettering is accomplished by coating the inner surfaces of the beam tubes with colloidal graphite (Aquadag). Keeping a low Cs background environment is crucially important in order to maintain and realize the full potential of the standard. Graphite getters play a particularly important role in space borne clocks with required useful lifetimes of at least 7.5 years. From our earlier investigations we identified Poco CZR-2 as the graphite possessing superior Cs gettering capability. Currently we are investigating the Cs gettering properties of colloidal graphites containing graphite particles suspended in various types of solvents (termed "primary vehicle") and binders. The gettering performance was studied by measuring the Cs-sticking coefficient on stainless steel surfaces coated with colloidal graphite. We studied Electrodag 121, Electrodag 179 and Aquadag E, all manufactured by Acheson Colloids Company. The three types contain water as the primary vehicle with different binders (organic and inorganic). Out of the three, only Aquadag E shows any significant Cs gettering capability. The other two behave essentially like uncoated metal surfaces. We present these results together with possible reasons for this behavior.

## INTRODUCTION/BACKGROUND

In the cesium (Cs) beam tube clocks (CBT) only a small fraction of the Cs atoms effusing from the oven travel through the microwave cavity where they undergo the 9192.6 MHz microwave transition. The remaining fraction of the effusing Cs atoms are collimated and gettered by slits and baffles made of graphite placed at various locations inside the clock. This is schematically shown in Fig. 1. The inner surfaces of the beam tube are also coated with colloidal graphite (Aquadag) to improve the gettering. Should the graphite getters cease absorbing Cs, the increasing background of Cs atoms would rapidly obscure the true signal needed to operate the standard reliably. The clocks will also be rendered unusable if the insulators in the electric feedthroughs get shorted by the conductive coatings of the background Cs atoms. Getter failure is then a life-limiting process in the Cs clock. For the clocks to be reliable, the gettering lifetime should be at least as long as the guaranteed lifetimes of the tubes. Standard failure is of particular concern in satellite applications, such as GPS, where reliable operation over the specified tube lifetime, 7.5 years, is required.

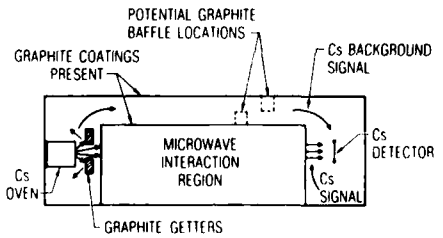



Fig. 1. Schematic representation of a Cs Beam Tube, showing the locations of graphite getters. The exposed metal surfaces inside the tube are also coated with Aquadag to further reduce the migration of Cs atoms to the detector region.

The practice of putting graphite blocks as getters inside the beam tubes dates back to the early days of the development of Cs beam frequency standards. Although the usefulness of graphite as a Cs getter has all along been qualitatively recognized, no systematic study of the Cs gettering properties of graphite has so far been reported. We have been experimentally investigating the gettering properties of graphite and colloidal graphite in order to be able to reliably characterize the type and amount of graphite to be used in space-borne clocks.

Many materials are effective surface and bulk getters (Ref. 1). In surface getters the gas molecules collide and react with the surface. There is little diffusion into the material. In bulk getters the efficiency is determined by the diffusion of the impinging molecules into the bulk. In both cases, the characteristics of the surface and the interaction between the bombarding atoms and the surface should be such that the incident atoms stick well to the surface. Stated quantitatively, the prerequisite for efficient gettering is that the sticking coefficient $\gamma$ be close to 1. Therefore, study of the sticking coefficient enables one to address the problems of gettering from a microscopic view. Another approach is to study the bulk gettering properties of graphite, e.g., how much Cs a particular weight of graphite may absorb. In our laboratory we are investigating from both the microscopic (sticking coefficient experiments) and the macroscopic (bulk gettering) approaches.

In the next section we present the details of our experimental procedure for measuring the sticking coefficient $\gamma$.

## EXPERIMENTAL DETAILS AND DATA

### Measurement of Sticking Coefficient ($\gamma$)

To measure the sticking coefficient $\gamma$ of Cs on graphite we use an atomic beam apparatus. The experimental technique is illustrated in Fig. 2. A well-collimated beam of Cs atoms effusing from an oven is directed towards a graphite target. The reflected beam at a particular angle $\theta$ is measured using a surface ionization detector. We measured the sticking coefficient of Cs on graphite relative to a paraffin-coated surface, which is known to be perfectly nonsticking. The sticking coefficient $\gamma$ is defined as

$$\gamma(\theta) = 1 - \frac{\text{reflected signal from graphite at } \theta^{\circ}}{\text{reflected signal from paraffin at } \theta^{\circ}}$$

CH2588-2/88/0000-510 $1.00 C 1988 IEEE

We are studying the evolution of $\gamma$ as a function of time for various incident Cs atomic beam intensities. Typical background pressure in the atomic beam apparatus is $< 1 \times 10^{-8}$ torr.

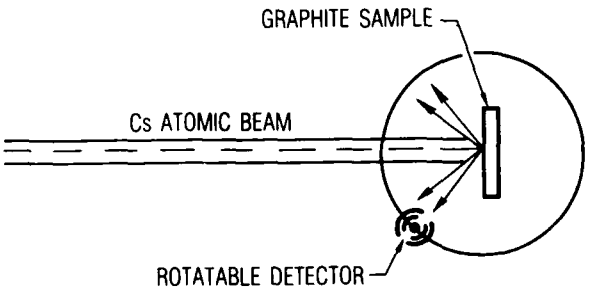



Fig. 2. Experimental arrangement for measuring the sticking coefficient $\gamma$ of Cs on graphite.

Graphite samples used in the present experiments are first baked externally in a quartz tube furnace at 1000°C under liquid-nitrogen-trapped rough vacuum for a minimum of two hours, by which time the residual gas pressure is found to drop to a few microns. Glass seals are used everywhere in order to prevent the contamination of graphite samples. The inner walls of the quartz oven are thoroughly cleaned before the start of baking. The graphite sample is then loaded into a copper holder inside the atomic beam apparatus and pumped down to ~ $1 \times 10^{-7}$ torr. The sample is baked to 350°C in high vacuum for several hours until the background pressure is below $1 \times 10^{-7}$ torr. The vacuum chamber containing the graphite target is kept isolated from the source chamber containing the Cs oven, except while measuring the reflected beam signal. This is accomplished by placing a gate valve between the two chambers. Physical isolation is necessary because even at room temperature we detected a weak Cs beam effusing from the oven. Inadequate and improper baking of the graphite samples gets reflected in the rapid decrease of the Cs sticking coefficient. When properly degassed, graphite can be made more gas-free than most common metals used in high vacuum systems. It is generally believed that graphite degassed at 1000 - 1100°C is "well outgassed."[2]

The residual gas pressure in our apparatus is usually less than $1 \times 10^{-8}$ torr. This is accomplished primarily by baking the vacuum chambers at about 100°C for about 24 hours. This improvement in the high vacuum resulted in a marked decrease in the background signal recorded by the surface ionization detector. Under best operating conditions, the typical background signal is $< 10^{-15}$ A whereas the reflected Cs signal at 70° is ~ $10^{-11}$-$10^{-12}$A. The collision rate of the background gas at the residual gas pressure of $1 \times 10^{-8}$ torr is ~ 27 monolayers/hour, which is comparable to the collision rate of Cs on the graphite target. It is conceivable that what we are observing is actually a gas-phase reaction catalyzed by the surface--in other words, reaction of the impinging Cs atoms with the residual gas atoms on the surface. The evidence counter to this is the fact that our experimental data essentially remains unchanged when the Cs beam flux is increased by a factor of two and residual gas pressure is reduced by about a factor of two.

The variation of $\gamma$ with time for three different types of graphite and an an aquadag coated surface is shown in Fig. 3. The percentage porosity of the respective samples is also indicated in Fig. 3. The sudden steps in the data are due to fluctuations in the incident atomic beam which remains very stable only for about 24 hours. A differential measurement, with continuous monitoring of both the incident and the reflected beams would be more desirable for these types of long term experiments. Suitable corrections for the beam fluctuation were made in computing $\gamma$ (Ref. 3).

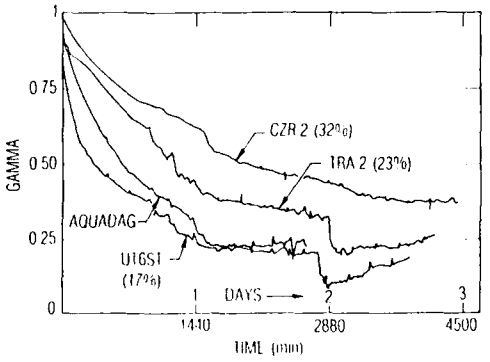



Fig. 3. Long term (60-75h) variation of the sticking coefficient as a function of time. At the end of 60-75h of exposure to the Cs beam, the sticking coefficient drops to ~ 0.2-0.4, depending on the type of graphite. The Cs beam intensity used in these measurements is ~ $4 \times 10^{2}$ atoms/cm$^2$ sec. POCO CZR-2 exhibits a distinctly slower rate of decrease of $\gamma$ than all the other types investigated.

Colloidal graphite is primarily used in the electrical/electronic industry as conductive or resistive coatings and impregnant. For example low resistance colloidal graphite coatings on television tubes bleed off static charges and confine harmful radiation. These coatings protect computers and related electronic equipment from the adverse effects of electromagnetic interference and electrostatic discharge. A wide variety of coatings have been developed to meet the needs of the hi-tech electronic industry. The use of colloidal graphite to getter Cs in beam tubes is fortuitous. In light of this it will not be surprising if many of the commerically available colloidal graphite coatings show very poor Cs gettering capabilities. Our experimental studies clearly confirm this behavior.

Colloidal dispersions are homogeneous suspensions of extremely small particles in fluids. Aquadag and Electrodag are suspensions of high purity graphite in a wide range of liquid carriers. Aquadag and Electrodag are registered trade names used by Acheson Colloids Company, who first developed these coatings. The term 'colloidal' refers to the size of the individual graphite particles which is about a micron. The particles tend to be flake-like. As the particles become finer the ratio of their surface area to weight becomes very large. This is important since many of

the properties of colloidal graphite are dependent on the surface available. Graphite is dispersed in two main groups of fluids, aqueous and nonaqueous. Aqueous fluid, of course, is water whereas nonaqueous fluids can be organic or inorganic solvents, resins and oils. Dispersions are made in these various fluids to satisfy certain definite uses. The various solvent systems are frequently used to put graphite down in a conductive film or in a film which is to be lubricating in its dry state after the solvent has been evaporated. In order to obtain stable and durable coatings for a wide variety of environmental conditions binders are also added. Binders also prevent the flocculation of the finely dispersed graphite particles.

Colloidal graphite surfaces were prepared by coating precleaned stainless steel substrates with aquadag using a brush. Several coatings were applied. The coated substrate was cured at about 100°C for an hour inside the vacuum chamber. The substrate was exposed to the Cs atomic beam after it had cooled to the room temperature. The gettering performances of three different colloidal graphites were assessed by measuring the Cs sticking coefficients of stainless steel substrates coated with these suspensions.

Aquadag E: This consists of processed micrographite with deionized water as the primary vehicle. It has a gel like consistency. The variation of the sticking coefficient $\gamma$ with exposure time is shown in Fig. 4. For comparison, the variation in $\gamma$ for POCO CZR-2 is also included. Clearly, after a day of exposure to the Cs flux, the surface coated with Aquadag E exhibits very poor Cs gettering capability. This is in sharp contrast to POCO CZR-2, which, under similar conditions, continues to maintain $\gamma$ at 0.2-0.3 even after several days of Cs exposure.

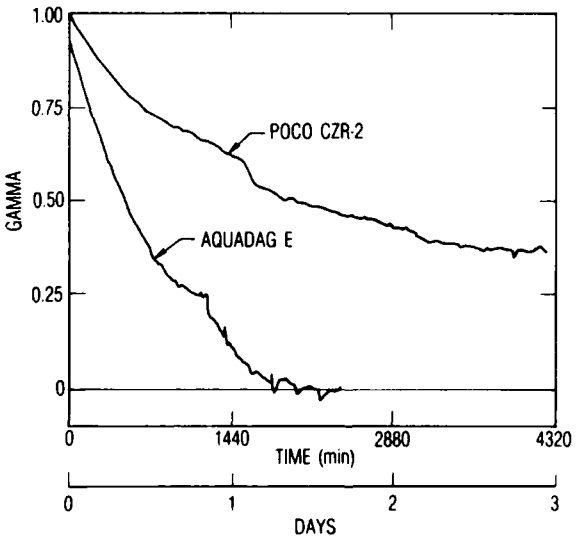



Fig. 4. Variation of $\gamma$ with time for Aquadag E. For comparison data for POCO C2R2 is also shown.

Electrodag 121: This is composed of graphite particles in water along with inorganic binders. This forms a hard and durable coating which can withstand the high temperature (350°C) baking required in CBT fabrication. Variation in $\gamma$ with time for substrates coated with Electrodag 121 is shown in Fig. 5. Within a few minutes of exposure to the incident Cs atomic beam, $\gamma$ drops essentially to zero. This typically is the behavior of uncoated metal surfaces. In other words, metal surfaces coated with Electrodag 121 exhibit no enhancement of the Cs gettering efficiency over uncoated surfaces. The very poor gettering efficiency of Electrodag 121 is in sharp contrast to Aquadag E.

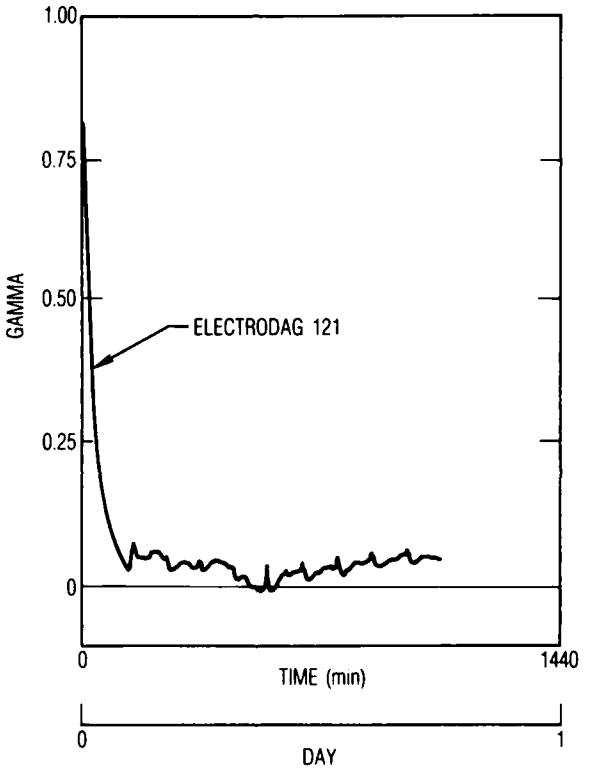



Fig. 5. Variation of $\gamma$ with time for Electrodag 121

Electrodag 179: This consists of graphite particles in water with potassium silicate as binder. It is also low in organics to reduce outgassing under vacuum. Electrodag 179 forms a hard, durable coating which can withstand up to 550°C. The experimental data for the Cs sticking coefficient on surfaces coated with Electrodag 179 is shown in Fig. 6. It exhibits very poor Cs gettering efficiency.

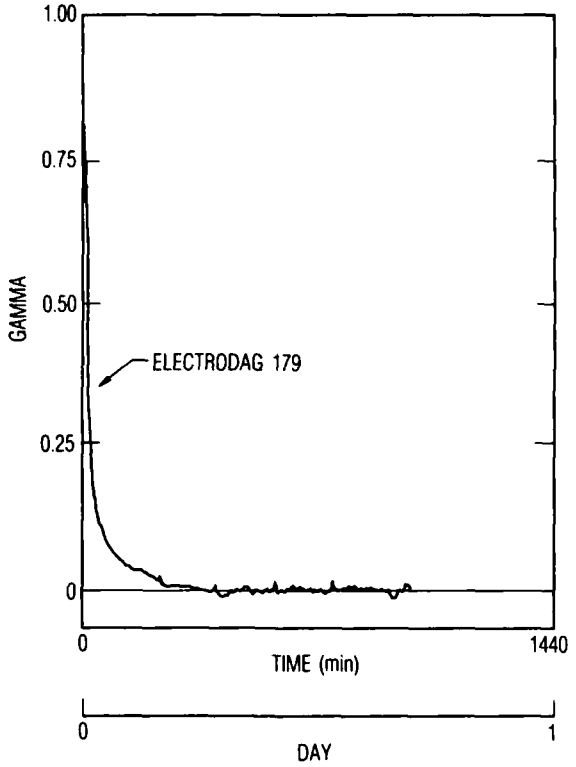



Fig. 6. Variation of γ with time Electrodag 179

## DISCUSSION AND CONCLUSIONS

Our experimental results represent the first quantitative investigation of the Cs gettering capabilities of colloidal graphite. Our investigations clearly show that collioidal grapite coated surfaces do not getter Cs as efficiently as solid blocks of graphite. In a CBT, substantial amount of solid graphite blocks are placed at various locations and the bulk of the gettering takes place at the graphite blocks.

This apparent dependence in the Cs gettering efficiency on the type of colloidal graphite prompted us to take a careful and detailed look at the properties and composition of different types of colloidal graphite. The Cs gettering capability of Aquadag E is far superior to Electrodag 121 and 179. Although the three colloidal suspensions contain water as the primary vehicle, they have different binders. The fractional binder content in Aquadag E is the smallest the exact type and amount of binder used is proprietary to the manufacturer. Due to the relatively small amount of binder Aquadag E has gel like consistency. The maximum service temperature of Aquadag E coatings is only 200°C which is substantially lower than the baking temperatures used in the fabrication of beam tubes. Electrodag 121 and 179 contain larger fractions of binder content relative to Aquadag E. These have the liquid consistency. For example Electrodag 179 contains potassium silicate as the binder which helps in forming durable and hard coatings which can withstand temperatures up to 550°C. It may be that the binder and the graphite pigment form a complex matrix which although improves adhesion and durability perhaps greatly reduces the graphite surface adsorption sites where impinging Cs atoms can stick quasi permanently. Better Cs gettering capability of Aquadag E may perhaps be due to the relatively small binder content. This opens up many new questions: Can colloidal graphite suspensions be formulated which is as good as solid graphite in gettering Cs? Our experience indicates that without suitable binders the colloidal graphite coatings flake off. Are there binders which improve the adhesion and durability of these coatings without substantially reducing the Cs gettering capabilities? Experimental investigations are currently in progress to address these issues.

A difficult problem to address experimentally is getter performance after a number of years of Cs exposure. Theoretical approaches may be the most productive when analyzing long term gettering and, in particular, sticking coefficient evolution. While a rigorous model of Cs gettering by graphite does not yet exist, the present experiments are of value in arriving at simple models. Graphite is composed of coke grains which are held together with pitch residue. Cs atoms that stick to the graphite surface must penetrate into the graphite or the sticking coefficient would rapidly go to zero. The penetration of Cs into graphite is typically treated as a "grain boundary" process (Ref. 4). Processes of this type are characterized by two diffusion mechanisms proceeding at dramatically different rates. Between the grains diffusion proceeds rapidly. Penetration into the grains proceeds at a much slower rate. However, due to the large grain surface area, a major fraction of Cs should be able to penetrate into the grains in reasonable periods of time. This is consistent with our observations. At present we are using bulk exposure and sticking coefficient data to develop a theoretical model of the gettering process. This should allow accurate prediction of sticking coefficient behavior even after many years of Cs exposure.

## ACKNOWLEDGEMENT

This is supported by the United States Air Force SpaceDivision under Contract No. F04701-85-C-0086.

## REFERENCES

1. J. F. O'Hanlon, A User's Guide to Vacuum Technology, pp. 206-214, J. Wiley and Sons, Inc. New York (1980).

2. G. A. Beitel, "The Use of Graphite in High and Ultrahigh Vacuum: A Review," J. Vac. Sci. Technol. 8, 647 (1971).

3. N. D. Bhaskar, C. M. Kahla, R. P. Frueholz, and R. A. Cook, "Cesium Gettering by Graphite", Proc. 41st Annual Freq. Cont. Symp. (1987), pp. 47-53.

4. D. Chandra and J. H. Norman, "Diffusion of Cesium through Graphite", J. Nucl. Mat. 62, 293 (1976).

42nd Annual Frequency Control Symposium - 1988

# THE EFFECT OF HUMIDITY ON COMMERCIAL CESIUM BEAM ATOMIC CLOCKS

by

James E. Gray, Howard E. Machlan, David W. Allan
Time and Frequency Division
Boulder, CO 80303

## Summary

Because a dependence of the frequency of commercial cesium beam clocks on humidity has been observed by others, NBS chose to control the humidity in one of the environmental chambers for the NBS clock ensemble. On 9 February 1988 the relative humidity was changed from a few percent to 48%. All of the clocks underwent a change in frequency. The resulting frequency changes were of different sign and of varying magnitudes among the clocks. In some, the changes were an order of magnitude larger than the flicker noise FM levels associated with the corresponding clocks.

The implications of such changes are quite significant in the generation of International Atomic Time (TAI) and of Universal Coordinated Time (UTC). The possibility of annual variations in the time scales of the principal timing centers as well as in TAI and UTC will be discussed. These results imply the potential for improved long-term frequency stability performance for individual clocks as well as for clock ensembles.

## Introduction

The authors know of only two primary timing centers that have actively controlled the humidity around their clocks, the University of Graz in Austria and PTB in the Federal Republic of Germany. The latter discontinued humidity control a few years ago due to equipment failure. A recent analysis by Bava et. al. [1] at the Instituto Elettrotecnico Nazionale at Corso, Italy documented the clear effect of humidity on the frequency of the clocks in the IEN clock ensemble. Work by Freon at the Paris Observatory has indicated both a cause and a possible cure for this problem [2], The presence of an apparent annual term in the International Atomic Time (TAI) scale has been a concern for several years. Figure 1 is a plot of TAI versus Guinot's optimally smoothed 1988 post analysis time scale, Terrestrial Time scale of the Bureau International des Poids et Mesures (TTBIPM88). Over the ten year period, ten cycles are apparent with amplitudes of a few parts in $10^{14}$.

Because of the above results and concerns, NBS made a decision to conduct an experiment involving servo control of the humidity in one of the environmental chambers (EC2) in the NBS clock ensemble.

## Effect of Humidity on NBS Clock Ensemble (Internal Estimate)

On 9 February 1988, a humidistat was turned on in one of the chambers of commercial cesium beam frequency standards which are a part of the National Bureau of Standards atomic time scale ensemble. The relative humidity increased from its dry season value of 12% to a controlled value of 42%, while the temperature remained stabilized at 23 degrees Celsius. The sensor used was a commercial capacitive-type with a manufacturer's stated accuracy of 4%. There were four cesium standards in the chamber at the time, and each of them shifted to a new frequency within about two hours after the humidity control was turned on. The frequency shift was estimated by comparing the average frequency measured during the first two weeks of humidification with that during the previous two weeks. Assuming that the frequency change is linearly proportional to the change in relative humidity, we can express the results as humidity coefficients.[1]

| SERIAL NUMBER | FREQUENCY SHIFT | COEFFICIENT PER % |
|---|---|---|
| HP323 | +1.1 E-13 | +3.6 E-15 |
| HP324 | -1.8 E-13 | -6.1 E-15 |
| HP2165 | -1.3 E-13 | -4.4 E-15 |
| HP2315 | -0.5 E-13 | -1.6 E-15 |

Using these coefficients, we can calculate the extent to which seasonal variations in humidity should have affected the frequency of these clocks, and then compare the actual frequency variations with those predicted. For a period of about 400 days before the humidistat was turned on, we read the humidity off a commercial strip-chart recorder (capacitive-type sensor) once a week. The humidity varies slowly in our laboratory, so that a weekly reading should adequately represent the conditions in the chamber. The manufacturer of the recorder claims the instrument to be accurate to 4% between 20% and 80% relative humidity. About half of our readings are below 20% so that the accuracy of this recorder is questionable. However, we have two recorders and they have always agreed with one another within the 1% resolution that we can read the strip-charts. A more accurate (1%) commercial instrument was utilized in the servo controlling the humidity. This instrument featured a Dunmore type hygro-scopic salt detection method. In steady state its reading was 48% -- differing from the 42% of the capacitive- type sensor. Since the above coefficients are based on a change in humidity, the capacitive-type instrument was used for those calculations -- the particular Dunmore-type sensor we had purchased was limited in range in order to achieve the 1% accuracy.

Contribution of the U.S. Government, not subject to copyright.

Figure 2 shows the recorded humidity in the chamber during the 400 days. We multiply this recorded humidity by the coefficient for HP323 to find the effect of humidity on the frequency of HP323. Figure 3 shows the humidity effect plotted on the same scale as the observed frequency of HP323. The frequency measurements are one week averages. The appearance of figure 3 suggests that humidity variations are the principal cause of seasonal variations in the frequency of HP323.

Although HP324 is of the same make, model and adjacent serial number, its behavior during the 400 days is rather different. Its frequency responds to two influences with about equal magnitude. One of them is the humidity, and the other is the signal level in its electronics, as indicated by the "second harmonic" reading on its panel meter. During the 400-day interval, we increased the amplifier gain to compensate for the decline in signal amplitude with ageing of the cesium-beam tube. At that time, we observed a frequency shift of +2.3 E-13 for a change in the second harmonic reading from 15 to 40. We can form an amplitude coefficient with this information just as we did with the humidity. The amplitude coefficient is +9.2 E-15 per unit on the second-harmonic meter. Multiplying this coefficient by the recorded amplitude gives the expected variation in frequency due to variation in signal amplitude. It has the form of a sawtooth, declining as the tube ages, and rising sharply when we increase the amplifier gain. To explain the behavior of HP324, we must add this amplitude effect to the humidity effect. Figure 4 shows the result. These two effects together account for most of the frequency variations of HP324 during the 400-day period.

We have no long-term data for HP2165 because of repeated repair problems. HP2315 was in operation during the later part of the 400-day period. Figure 5 shows that its frequency begins to shift at day 47100 for unknown reasons. Humidity is clearly not the only cause of frequency shifts in this clock.

The data presented so far indicates that variations in humidity generally affect the frequency of cesium beam standards but is not necessarily the only cause of frequency variations. We also consider the behavior of the other clocks in our time scale ensemble. These are located in another chamber, where the humidity has never been intentionally altered. Consequently, we do not know the humidity coefficients of the clocks in this chamber. Nevertheless, we can examine the long-term variations in their frequencies and compare these variations with the humidity record for the chamber. In many instances, the resemblance of the two curves is so great that we are sure humidity variations are the major cause of the observed frequency variations. We can even estimate the magnitude of the humidity coefficient -- whatever is needed to make the two curves match. This indirect procedure permits us to extend our observations to a larger population of clocks, though with less confidence in the certainty of the results.

Figures 6 & 7 show two instruments that seem to be affected by humidity.[2] The frequency curve for FTS 113 did not resemble the humidity curve.[3] Figure 8 shows the last few months of the life of the cesium tube in FTS 217.[4] The record is short, and the humidity profile so simple, that we cannot be sure whether we are seeing a humidity effect. Figure 9 shows OSQ61, whose frequency appears to track the humidity.[5]

The humidity coefficients for these four instruments have been chosen to match the calculated curve to the measured frequency curve. The value for FTS217 is particularly uncertain, but it is included because it is the only example we have of an option 004 instrument of this make. We can be reasonably confident of the other coefficients.

| SERIAL NUMBER | COEFFICIENT PER % |
|---|---|
| HP1316 | +3.8 E-15 |
| HP352 | -11. E-15 |
| FTS217 | -9. E-15 |
| OSQ61 | +40. E-15 |

Changes in relative humidity cause significant changes in the frequency of most of the commercial cesium beam frequency standards that we have examined. For some of the most stable of these clocks, changing relative humidity is the predominant cause of long term instability. The sensitivity of each clock to changes in humidity seems to be an individual matter. Clocks of the same make, model and even adjacent serial numbers do not necessarily behave the same. For this reason, we caution the reader not to suppose that the numbers given here are typical of the makes and models mentioned. We simply do not have enough information for such generalizations. We do suggest that users of such instruments should consider the possibility that their instruments may vary in frequency with changes in humidity, even though our data are too limited to predict it.

The frequency measurements reported here are made with respect to our AT1 time scale. We have neglected the humidity coefficient of this time scale. The coefficients given above contain an error to the extent that AT1 has a non-zero humidity coefficient. We can now estimate that coefficient. AT1 is a weighted average of the members of its ensemble. We can multiply each clock's humidity coefficient by its weight in AT1 and add them up. This simplified procedure overlooks the effect of the non-commercial devices that contribute to AT1, and it overlooks the fact that the composition of AT1 changes as the assortment of available clocks changes. Nevertheless, the coefficient obtained in this way is enough of an approximation to give us a rough idea of the magnitude of the error it contributes to the coefficients given above. We estimate the humidity coefficient of AT1 to be +0.9 E-15 per percent. The humidity coefficients of the several clocks have cancelled to some extent. This coefficient not only represents an error in the reported values of the clock humidity coefficients given above, but it also estimates the effect of the humidity in our laboratory upon UTC(NBS), which is obtained from AT1 by adding coordination corrections. Apparently, UTC(NBS) goes about 0.1 microsecond early during our summer and about 0.1 microsecond late during our winter. These deviations should be reduced in the future to the

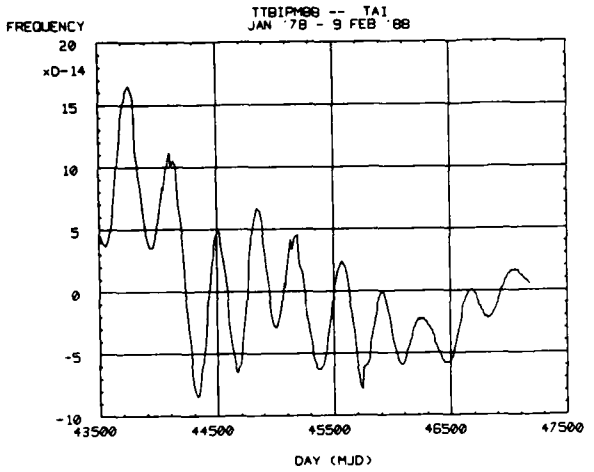



FIGURE 1. FREQUENCY OF (TTBIPM--TAI) JANUARY 1978 - FEBRUARY 1988

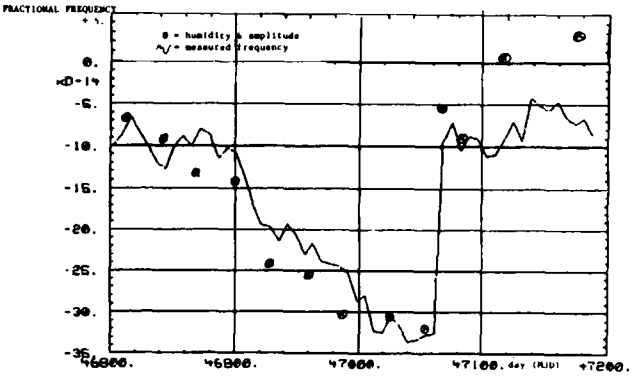



FIGURE 4. HP324 - EFFECTS OF HUMIDITY AND SECOND HARMONIC AMPLITUDE

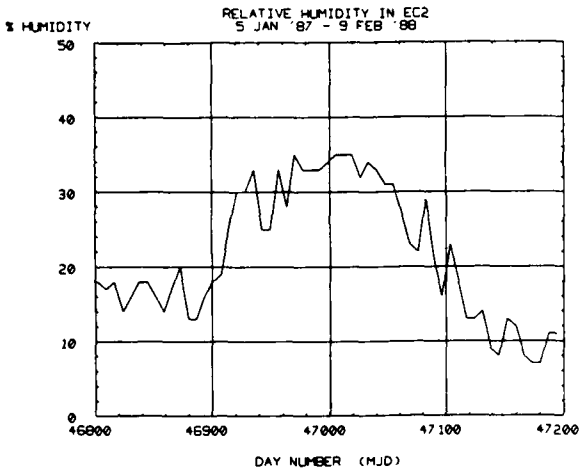



FIGURE 2. RELATIVE HUMIDITY IN CHAMBER EC2

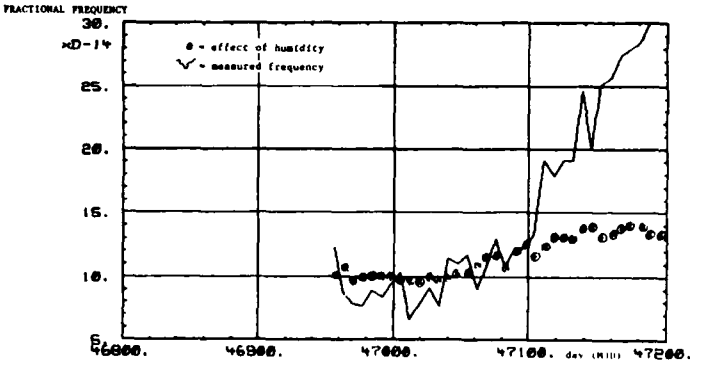



FIGURE 5. HP2315 - EFFECT OF HUMIDITY

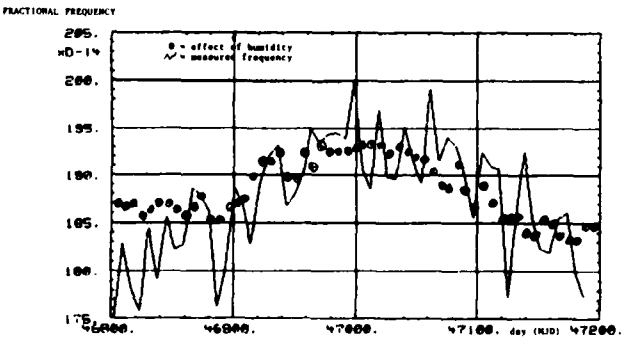



FIGURE 3. HP323 - EFFECT OF HUMIDITY

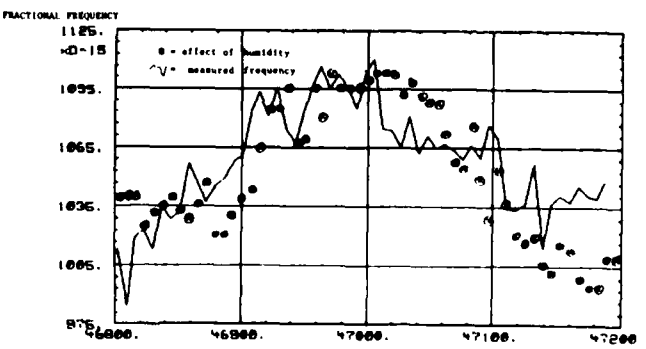



FIGURE 6. HP5110 - EFFECT OF HUMIDITY

extent that we apply humidity control to the clocks of our ensemble.

## External Estimate of the Humidity Dependence of Clocks

The internal estimate discussed above overlooks any humidity-induced effects which may be common to all of the NBS clocks. For that reason, it is especially interesting to compare NBS(AT1) with other time scales. Probably the best external scale for comparison is the TTBIPM88 scale. Figure 10 is a plot of the frequency of this terrestrial time scale versus the NBS(AT1) ensemble. Comparing figure 10 with figure 2 shows an apparent correlation. An approximate estimate of the coefficient is minus three parts in $10^{15}$ per percent change of the relative humidity. This is of opposite sign and about three times larger than what was estimated from internal measurements as stated above. The frequency stability between these two time scales is plotted in figure 11. From the humidity coefficient, one can infer a fractional frequency stability caused by humidity acting on the NBS ensemble. This humidity-caused instability is plotted in figure 12. Figures 11 and 12 agree within the uncertainties, showing that the model is self-consistent.

One may conclude that the NBS clock ensemble probably has been perturbed in the past by environmental humidity variations. Since 9 February 1988, the environmental chamber EC2 has been controlled within about 1% at 48% relative humidity. The effect of humidity on these clocks should therefore not degrade the stability of NBS ensemble to more than a few parts to $10^{15}$ due to this cause. Environmental chamber number three is yet to be dealt with.

Several of the clocks contributing to international atomic time are not protected from variations in humidity. Most of these clocks are in the northern hemisphere and they could have seasonal intercontinental correlations due to this effect. The annual variations plotted in figure 1 could well be driven by this effect. The fact that the internal and external estimate of the humidity coefficients for NBS(AT1) were of opposite sign and a factor of three different in magnitude could be explained if the annual variations were not in NBS(AT1), but in TTBIPM88. There is additional evidence for that possibility as discussed in reference [3]. In actuality there are probably some annual variations in both time scales. Another possible explanation is that all of the NBS clocks have a humidity coefficient much larger than measured because relative measurements detect only differences between clocks, rather than the full effect. Further study is needed to ascertain, with some degree of certainty, where these annual variations come from. The probability that relative humidity changes are an important concern for time scale environments seems now evident.

## Conclusion

The data presented in this paper indicate that variations in relative humidity are a major contributor to the long-term instabilities in at least some cesium atomic clocks. Several of the atomic clocks contributing to the NBS ensemble have been affected, and the same kind of clocks are used in the generation of International Atomic Time as well as at the primary timing centers throughout the world.

There are insufficient data at this point to conclude where the annual variations are. A study is underway to answer some of these questions. It seems probable that seasonal humidity variations could be an important cause of the annual variations observed between time scales.

Long-term instabilities in earth-bound time systems have become increasingly important with the discovery of millisecond pulsars. The effects of humidity may cause terrestrial time scales to change a few parts in $10^{14}$ over the course of half a year. This is the level of estimated instabilities in the millisecond pulsar PSR 1937+21 for the same integration time. At an integration time of one year this pulsar has an estimated instability of about one part in $10^{14}$. This pulsar may assist in determining the sources of these annual variations. Now that variations in relative humidity are recognized as a probable cause of annual variations in terrestrial time scales, controlling the humidity or compensating for the effect of humidity should improve the long-term performance of earth bound atomic clocks. This improvement may be very important for characterizing the long-term performance of millisecond pulsars.

Footnotes

[1]Each of these four clocks was a model 5061A option 004 made by Hewlett Packard. Identification of a commercial company does not imply recommendation or endorsement by the National Bureau of Standards, nor does it imply that any identified entity is the only or the best available for the purpose.

[2]Serial HP1316 is a Hewlet-Packard 5061A option 004, and serial number HP352 is a Hewlett-Packard 5061A containing a standard tube.

[3]Serial number FTS113 is a Frequency and Time Systems model 4050.

[4]Serial number FTS217 is a Frequency and Time Systems model 4050 option 004.

[5]Serial number OSQ61 is an Oscilloquartz model 3200.

## References


[1] E. Bava, F. Cordara, V. Pettiti, and P. Tavella, "Analysis of the Seasonal Effects on a Cesium Clock to Improve the Long-Term Stability of a Time Scale," 19th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 1-3, 1987, Redondo Beach, CA.

[2] Dr. W. Lewandowski, private communication.

[3] D. W. Allan, "A Study in Long-term Stability of Atomic Clocks," 19th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, December 1-3, 1987, Redondo Beach, CA.

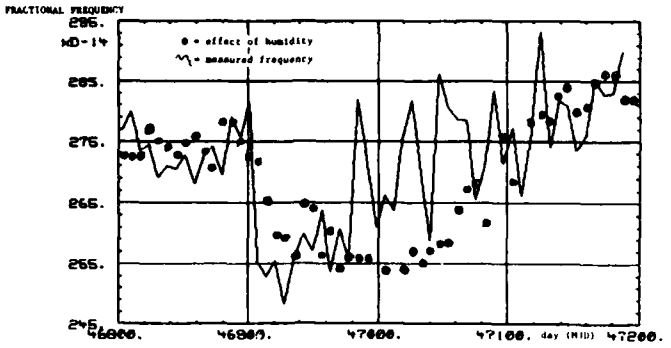



FIGURE 7. HP352 - EFFECT OF HUMIDITY

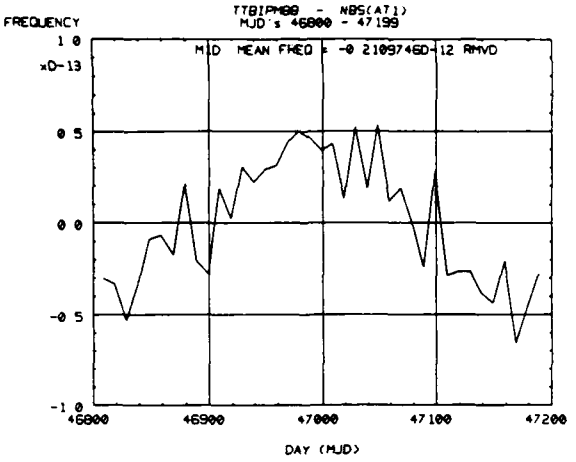



FIGURE 10. FREQUENCY OF (TTBIPM88 - NBS(AT1) MJD 46800 - 47199

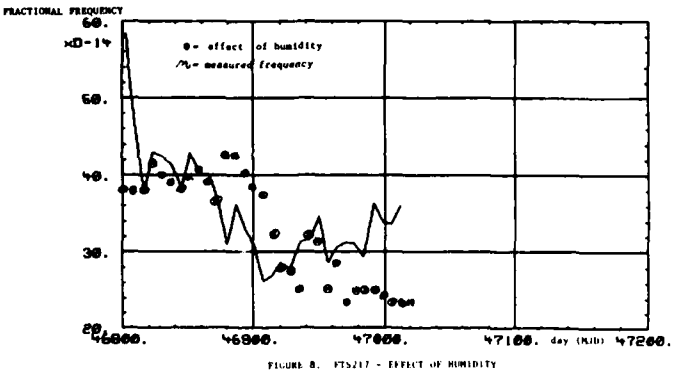



FIGURE 8. FTS217 - EFFECT OF HUMIDITY

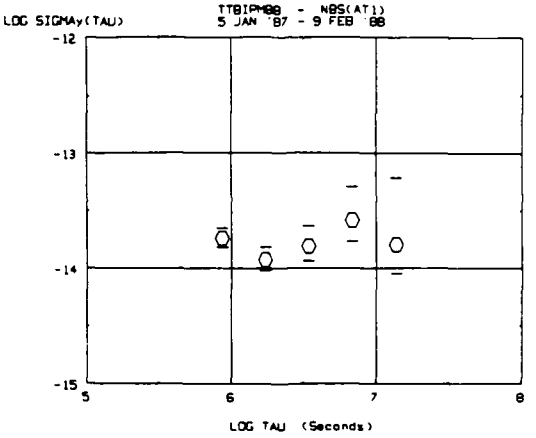



FIGURE 11. INSTABILITY OF NBS(AT1) vs. TTBIPM88

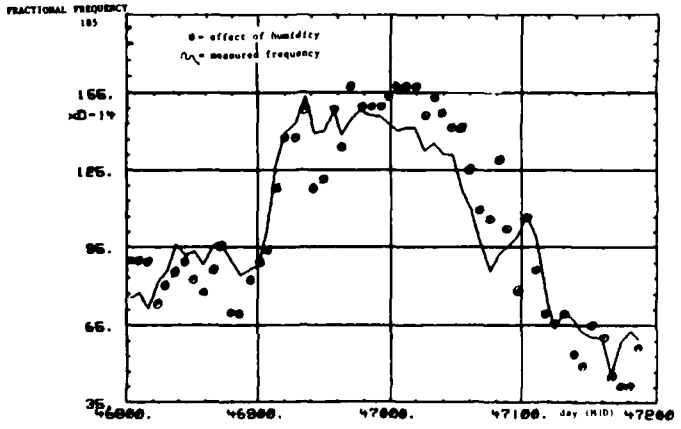



FIGURE 9. OSOM1 - EFFECT OF HUMIDITY

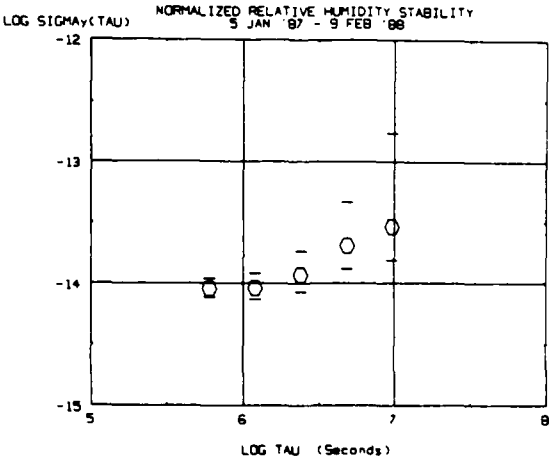



FIGURE 12. HUMIDITY-INDUCED INSTABILITY PREDICTED FOR NBS(AT1) FROM EXTERNAL ESTIMATE OF HUMIDITY VERSUS FREQUENCY COEFFICIENT

42nd Annual Frequency Control Symposium - 1988

# A NOVEL COMPACT RUBUDIUM FREQUENCY STANDARD WITH A LOW SENSITIVITY TO MAGNETIC AND VIBRATIONAL DISTURBANCES

A. Stern, A. Hertz, Y. Zarfaty and A. Lepek
TFL, Time & Frequency Ltd., P.O.B. 1792
Holon 58117, Israel.

## ABSTRACT

TFL has developed a novel Rubidium Frequency standard which has a compact and ruggedized design.

Two improvements are incorporated into the traditional design:

- An electronic method which improves the frequency stability under vibration and accelerations.
- An electronic method to eliminate the sensitivity to magnetic fields.

These schemes are incorporated into our new Rubidium Frequency Standard. This unit has a very high modularity using partialy hybridized circuits. It has a good short term stability better than 3E-11/SQR(tau) between 1 and 500 seconds.

The OCXO incorporates a low g sensitivity SC cut quartz crystal.

## 1. INTRODUCTION

In this paper we describe a Rubidium Frequency Standard (RFS) which was designed in TFL. This unit, which is modeled TF4000A, has a compact and ruggedized design aimed at military applications. The design is based in part on F.E.I.'s model FE5606A, whereas most of the packaging and some electronics were redesigned at TFL[1]. Most of the electronics is packaged in discrete form. However it also includes 6 hybrides:

4 hybrides in the crystal oscillator unit, 1 hybrid in the Physics Package and 1 hybrid in the RF multiplier board. Special emphasis in the design was placed a modularity and ease of assembly as elaborated below.

In addition to the basic design, novel schemes have been invented at TFL to improve the RFS performance under vibration , acceleration and magnetic field *. These schemes are being studied now and are described below. They are added as options to our basic unit.

## 2. DESIGN APPROACH

The unit, shown in figure 1, was designed to meet the needs in airborn and mobile ground application. It has compact dimensions of 3.25x3.25x4.00 inch and was carefully designed to withstand military environmental conditions.

Figure 2 brings two views of an open unit. The unit is comprised of a Physics Package, a Crystal Oscillator package and a printed boards basket that includes 5 boards. The boards slide and plug into the mother board and each is readily removable with no interference to the other modules.

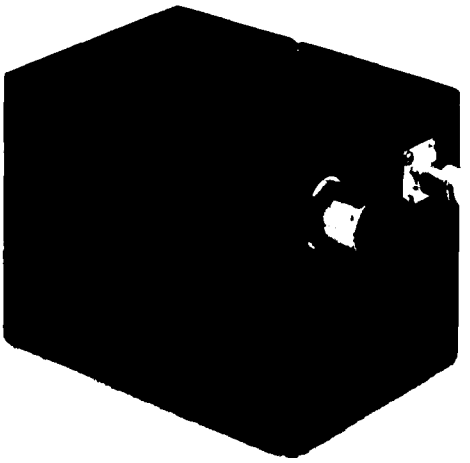


Figure 1: Rubidium Frequency Standard Model TF4000A.

* Patent applications pending

CH2588-2/88/0000-519 $1.00 C 1988 IEEE

Figure 2: TF4000A an open unit.

Most of the adjusting trimers are accessible directly, so that system tuning is most conventiently performed.

A very high degree of modularity continues to apply also to each module. Especially, the Physics Package is made of submodules that are easily removable with no need to dissassemble the package from the RFS assembly.
For example, the lamp Assembly is a completely independent submodule that includes its own connector and can be replaced if fails.

The cavity in the Physics Package operates in the pseudo TE101 mode[2] and includes both the resonance and the filter cell. A hybridized preamplifier connected to the photo diode is also included in the package. A magnetic inner shield surrounds the cavity only, while the outer shield composes the package box and surrounds both the inner shield, the lamp and the preamplifier.

The Crystal Oscillator (OCVCXO) is mounted on the top of the Physics Package. This is an ovenized Colpitts which uses a very low "g" sensitivity SC cut crystal (3E-10/g). The oscillator exhibits excellent features such as, short-term stability of 1E-11/sqr(tau), phase noise of -123 dBc (1Hz bw) at 100 Hz from carrier, fast warm-up of 4 minutes to 1E 9 and very small volume of 75cc.

The Servo-Loop block diagram is shown in figure 3. We start with 10MHz crystal, multiply by 12 to obtain 120MHz which, together with a synthesized 5.3125 MHz is injected into a Step Recovery-Diode (SRD) inside the Physics Package.
The SRD generates the 6.834...GHz resonance frequency which is phase modulated at 444Hz. The Physics Package produces an error signal which is phase detected, integrated and controls the Crystal Oscillator.

## 3. PERFORMANCE

Key performance levels have been measured on an engineering model and are listed in table 1. From these are notable, a very fast warm-up of 100 sec to 1.5E-10, excellent short-term-stability of 2.5E 12 at 500sec and good temperature coeficiant. The short-term-stability (Allan Variance) is also plotted in Figure 4.

## 4. IMPROVED PERFORMANCE UNDER VIBRATIONS AND ACCELERATIONS

Vibrations are known to be one of the worst enemy of precisions oscillators. Several schemes have been attemped in the past to reduce the vibration effect on frequency stability. One of the methods that has been

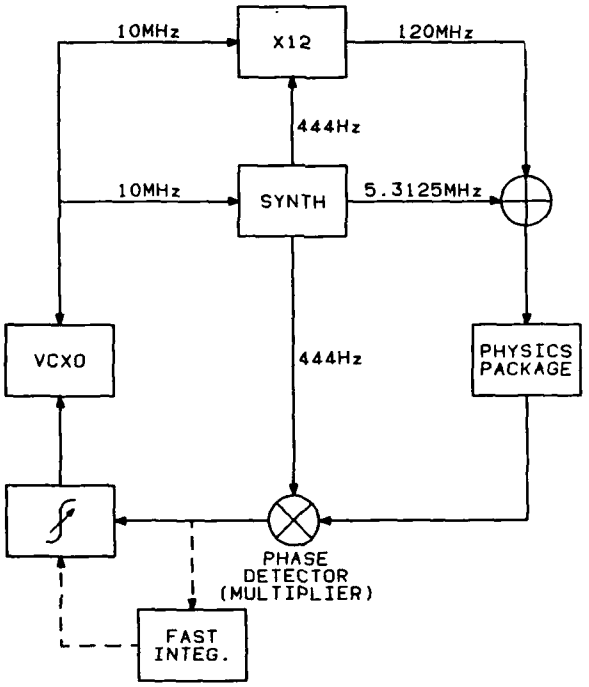



Figure 3: Conventional Servo-loop (solid line) and the fast integrator (dotted line) added for an adaptive loop-bandwidth (section 4).

| | |
|---|---|
| Output-Frequency: | 10MHz |
| Short-Term-Stability: | Sigma y(tau)<3E-11/sqr(tau) for tau≤1000sec (see Fig. 3) |
| SSB Phase Noise: | -93 dBc at 10Hz from carrier<br>-123dBc at 100Hz from carrier<br>-135dBc at 1KHz from carrier |
| Warm-up: | 100 sec to 1.5E-10 |
| Operating Temperature: | -55 to +76 deg C (base-plate)<br><3E-10 relative frequency change |
| Size: | 3.25 X 3.25 X 4.50 inch. |
| Vibrations: | Designed to perform in random vibrations 20-2000Hz, 5 grms with a frequency offset less than 5E-10 |

Table 1: Key performance levels Rubidium Frequency Standard Model TF4000A.

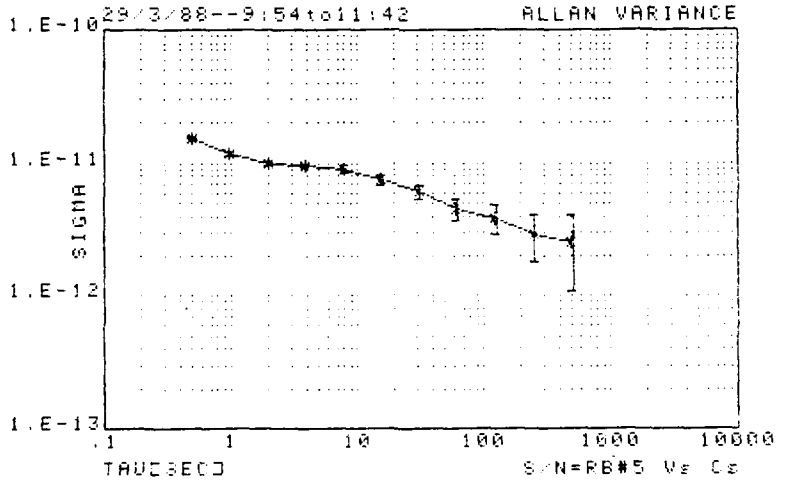



Figure 4: Short term stability ("Allan Variance") measured on an engineering model TF4000A.

suggested is that of accelerometer compensation[3]. In this method the crystal acceleration is measured and used to correct the frequency, according to a predetermined formula. A different approach has been taken at Litton [4].

The servo-loop bandwidth of a Rubidium standard was increased up to a 100Hz, resulting in a better performance under vibrations at frequencies ≤ 100Hz.

As is very well known [5], the frequency stability of a passive atomic frequency standard is a combination of the stabilities of the crystal oscillator and the atomic resonance system.
In the frequency domain we consider, for example, the SSB phase noise. For frequencies larger than the servo-loop bandwidth the phase noise is determined by the crystal oscillator, while at frequencies smaller than the bandwidth, the noise is related to the atomic reference system. The atomic reference, however, is less sensitive to vibrations and accelerations than the crystal. Hence, the increase of the loop bandwidth should result in an improved performance under vibrations as was shown in ref. 4.
However, since the oscillator stability is better than the Rubidium in the short-term, the stability of the combined system is degraded with the increase of the bandwidth in a static environment.

In order to deal with changing environmental conditions we introduce the adaptive servo-loop bandwidth.

In a quiet environment the bandwidth is set to an optimized low value (i.e. 0.1Hz). Under vibrations or change in acceleration the bandwidth is automatically increased, so that the corrective response of the atomic reference is faster, and the frequency stability is more related to the atomic reference. As a sensor for vibrations/change in acceleration we actually use the very same crystal in the crystal oscillator.
The scheme is depicted in figure 3, where a fast integrator is added to the conventional servo-loop. The fast integrator is connected to the output of the phase-detector and detects disturbances (noise) that are filtered by the loop integrator. As the unfiltered noise level is increased the fast integrator controls the loop integrator time-constant, and increases the loop bandwidth.

This method has been applied in a bread-board and the results are presented in figure 5.
The upper curve depicts the start of a simulated sine disturbance which is injected into the oscillator.
Middle curve gives the unfiltered signal at the output of the phase-detector. This signal

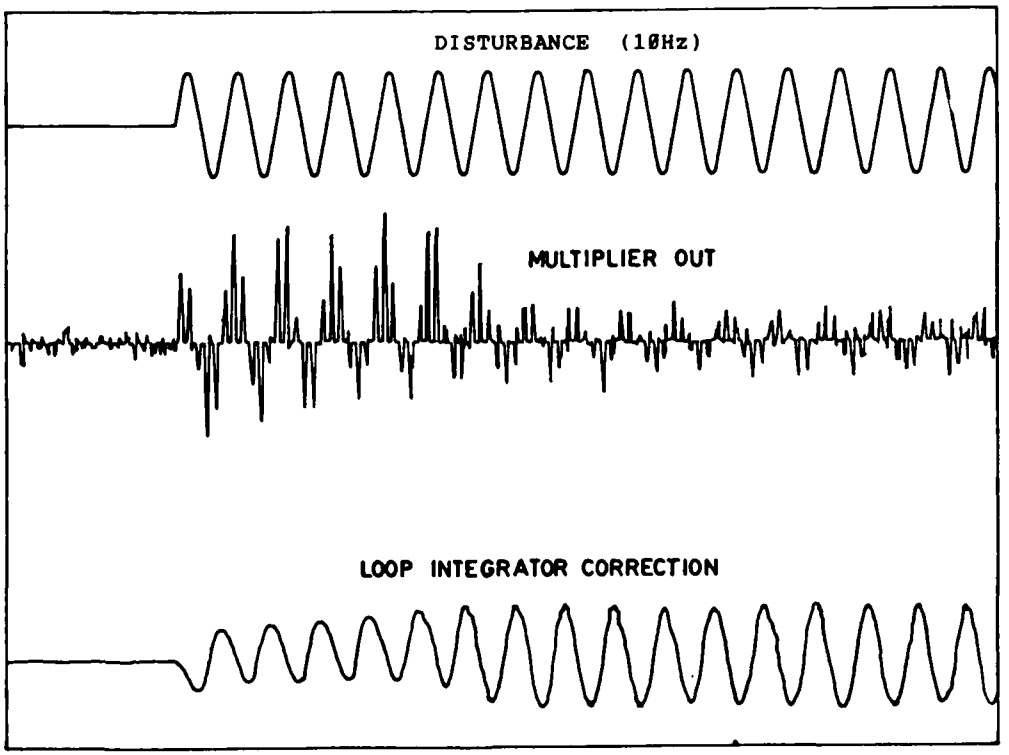



Figure 5: Adaptive serv-loop results.

measures the oscillator deviation relative to the reference.
Lower curve shows the loop intergrator output (or VCXO input). The curves exhibit a change in the loop-bandwidth from 3.62Hz to 7.96Hz, that occures ≈1 sec after the start of the disturbance. It is seen that the oscillator deviation, (middle curve) is being dampened due to an increase in the amplitude of the loop integrator corrective voltage (lower curve).

## 5. ELIMINATION OF MAGNETIC SENSITIVITY

The conventional approach to deal with magnetic disturbances is the use of one or several magnetic shields. These shields, however, increase size and weight and in addition can maintain a residual magnetization due to hysteresis. Moreover, they cannot shield against magnetic sources that may be found within the shielding envelope. Therefore we have thaught of an electronic scheme to cancel the effect of uncontrolled magnetic disturbances. In the following we describe the idea in its simplest form, where the magnetic effect is cancelled only to a first order approximation. However several extention schemes exist where, in theory, the effect of unwanted magnetic field is completely eliminated. These shall be described in another paper.

Consider the component H of the magnetic field, parrallel to the RF field, that is seen by a Rb (Cs or H) atom. This is composed of the "C" field, $H_c$, and an additional uncontrolled field, $H_n$ ,

$$H = H_C + H_n \qquad (1)$$

The clock transition frequency, $f$, has a quadradic dependence on $H$,

$$f = f_0 + aH^2 \qquad (2)$$

where $f_0 \approx 6.834...$GHz and $a \approx 573$ Hz/g$^2$ for the Rubidium clock transition.

Additional information can be gained once an alternating "C" field is introduced. Perhaps, the simplest scheme is to alternate the "C" field between $+H_c$ and $-H_c$ (i.e., square modulation) at a rate that is within the loop-bandwidth. Then, the clock transition frequency is alternating between $f_+$ and $f_-$

where,

$$f_\pm = f_0 + a\cdot(\pm H_c + H_n)^2 \qquad (3)$$

Since the alternating rate is smaller than the loop-bandwidth, the crystal oscillator shall be locked to the average frequency, $\bar{f}$, and

$$\bar{f} = \tfrac{1}{2}(f_+ + f_-) = f_0 + a\cdot H_c^2 + a\cdot H_n^2 \qquad (4)$$

Hence, the frequency is shifted from its nominal value by $a\cdot H_n^2$.
This is compared with the non-alternating conventional case where,

$$f = f_0 + aH_c^2 + 2H_c\cdot H_n + aH_n^2 \qquad (5)$$

The ratio of the frequency shifts due to Hr in both cases is,

$$\frac{f(\text{alternating "C" field})}{f(\text{non-alternating field})} = \frac{aH_n^2}{2aH_cH_n + aH_n^2} \qquad (6)$$

For $H_r \ll H_c$ this approximates to $H_r/H_c$

We have applied an alternating "C" field to our unit and measured the frequency shift in an external applied magnetic field. The results are plotted in figure 6. The upper curve depicts a frequency shift of 1.2E-9 that occurs in a conventional (non alternating) clock. The external field magnitude is 14.5 gauss and is attenuated by a factor of ≈800 by the shields.
The "C" field is set to ≈0.4 gauss.

The lower curve in the figure exhibits a frequency shift of ≈5E-11 measured under the same conditions, but with an alternating "C" field. The alternating frequency is 30Hz and the loop bandwidth is set to 1Hz. Hence, it is seen that the frequency shift is reduced by a factor of ≈24, in accordance with equation (6)

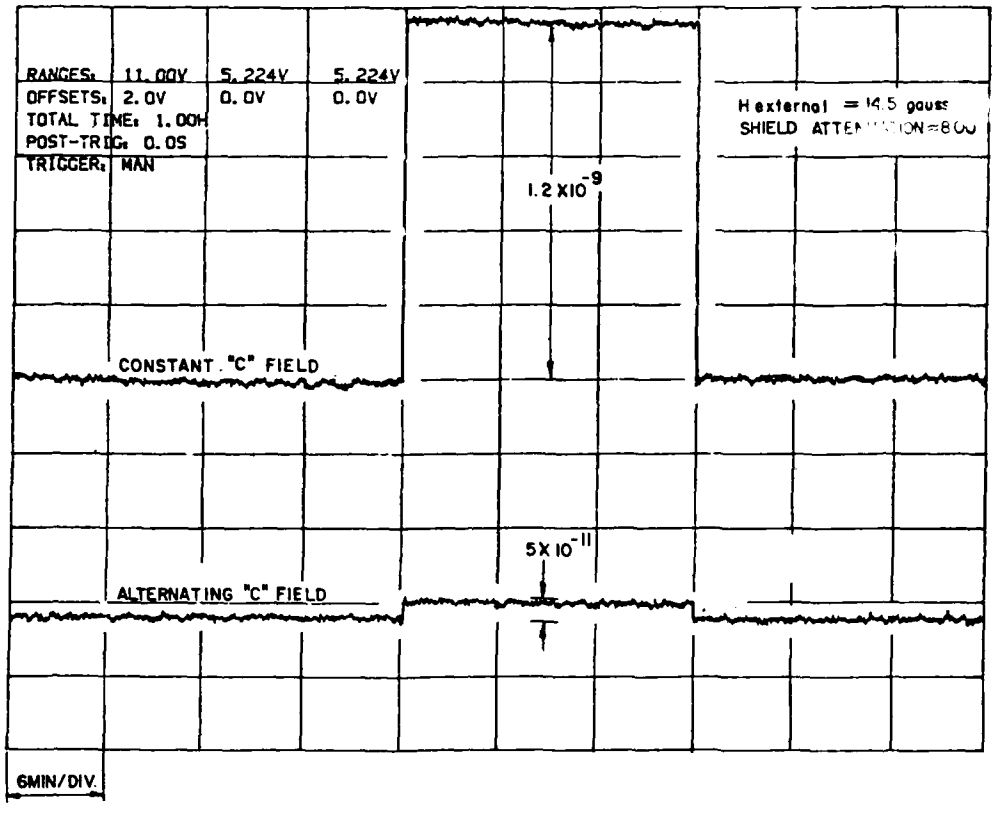



Figure 6: Frequency shift measured in an alternating "C" field. (lower curve) compared with the shift measured in the non-alternating "C" field (upper curve).

AKNOWLEDGMENTS:

To M. Bloch, T. McClelland, I. Pascaru and M. Meirs, from Frequency Electronics Inc., for their help and support in transfering the Rubidium technology to TFL.

REFERENCES

1. T.F.L. is a joint venture between Tadiran Ltd., Israel and F.E.I., Frequency Electronics Ltd., N.Y. The Rubidium technology at TFL was transfered from FEI.

2. H.E. Williams, T.M. Kwon and T. McClelland, Proc. 37th Ann. Freq. Control Symposium, 1983, page 12.

3. Vincent R. Rosati and Raymond L. Filler, Proc. 35th Ann. Freq. Control Symposium, 1981, page 117.

4. T.M. Kwon and T. Hahn, Proc. 37 Ann. Freq. Control Symposium, 1983, page 18.

5. J. Vanier and M. Tetu, IEEE trans. Inst. Meas. 28, 1979, page 188.

42nd Annual Frequency Control Symposium - 1988

# AN IMPROVED RUBIDIUM CONSUMPTION MODEL FOR DISCHARGE LAMPS USED IN RUBIDIUM FREQUENCY STANDARDS

R. A. Cook and R. P. Frueholz
Chemistry and Physics Laboratory
The Aerospace Corporation, P. O. Box 92957, Los Angeles, CA 90009-2957

## SUMMARY

The rubidium (Rb) discharge lamp is a critical component of the Rb gas cell atomic frequency standard. This frequency standard is flying on NAVSTAR/GPS satellites and will find applications on future satellites. The reliability of the lamp is a issue of concern, and aging of the lamp has been modelled in a satisfactory way for reliability determinations. As the lamp ages it slowly loses its Rb charge by diffusion into the lamp's glass walls. The model equation for Rb consumption C is $C = A + B\sqrt{T}$ where T is lamp operating time and B is an experimentally determined diffusion-related coefficient. A has been assumed to be loss resulting from a virtually instantaneous reactive process. In some of our data sets, consumption measurements indicate that the reactive cleanup is time dependent. We find this dependence well described by a term with two constants, A' and D, $A'(1-\exp(-DT))$, resulting from first order reaction kinetics. Including this term improves the ability to predict long term Rb consumption on the basis of a limited amount of experimental data.

## INTRODUCTION

Rubidium (Rb) frequency standards typically employ Rb electrodeless discharge lamps to supply the light needed for optical pumping, a process central to the operation of the standard. The standard's reliability has been shown in a number of instances to be limited by that of the discharge lamp. In satellite applications, such as the Global Positioning System (GPS), standard reliability is of the utmost importance. Consequently, Rb discharge lamp reliability is of great concern. One goal of The Aerospace Corporation's Atomic Clock's Research Program has been to insure the reliability of the lamps employed in satellite-compatible Rb frequency standards.

A typical Rb discharge lamp is illustrated on Fig. 1. Composed of an alkali resistant, glass envelope the lamp contains an initial Rb fill of hundreds of micrograms. An unlimited amount of Rb can not be placed in the lamp due to concerns over the stability of its output intensity. To allow the rf discharge several torr of a rare gas are also placed within the envelope. Lamp failure is indicated when the Rb resonance radiation required for optical pumping is no longer emitted by the discharge. In a previous publication[1], a very detailed analysis of the causes of Rb lamp failure was presented. Briefly, the metallic Rb initially placed in the lamp was found to disappear during operation. The major metallic Rb loss mechanism is diffusion into the lamp's glass walls. A less significant loss resulted from an initial, and assumed to be instantaneous, reaction between Rb and impurities within the lamp's envelope. The rate of Rb loss, or its consumption C was well modeled by an equation of the form,

$$C = A + B\sqrt{T}, \qquad (1)$$

with T the lamp's operating time and A and B empirical parameters determined by monitoring the loss of metallic Rb during lamp operation[1]. To extract A and B, consumption data versus operating time is fit to Eq. 1 using a least-squares procedure. The first term on the right hand side reflects the rapid reactive loss of Rb while the second term is due to the more gradual diffusive loss mechanism.

Once A and B have been determined for a particular lamp design, Eq. 1 can be employed to predict Rb consumption for any period of operation. Based on the consumption data available at the time of the prior publication Eq. 1 appeared to describe the consumption process very accurately. As more data has been obtained it appears that the reactive loss mechanism is not truely instanteous but occurs over a period of months. In this paper we analyze Rb consumption data treating A as a time dependent quantity. We find that long-term consumption predictions then become less sensitive to the amount of data analyzed and more accurate projections can often be made employing less data. Also, we find that when using limited amounts of consumption data Eq. 1 typically tends to over estimate long-term Rb consumption. It is then conservative and still adequate for lamp reliability predictions.

In Section II the observations that lead to a time-dependent A are discussed. The statistics of the modified consumption model are developed in Section III and applied to experimental data is Section IV. In Section V conclusions are presented.

## OBSERVATIONS THAT LED TO A RECONSIDERATION OF THE CONSUMPTION EQUATION

Many of the lamps being tested at Aerospace have several years of continuous operation, with interuptions only for the acquisition of consumption data. The method by which this consumption data is acquired has also been previously described[1], and for the purposes of this discussion is not important. As data was accumulated, and subsequently fit to the Eq. 1, certain trends became evident. The value for the intercept A obtained from the fitting procedure appeared to be decreasing monotonically. In addition, these changes over time appeared to be asymptotic with respect to some final values, for both the intercept A and the slope B. Presented in Table 1 are the intercepts and slopes of Eq. 1 for consumption data from four GPS-style Schott 8436 lamps produced by Efratom, a Division of Ball Corporation and six Corning 1720 lamps produced by EG&G, Electronic Components Division. For the EG&G lamps, A appears to be rising to some value around 19 or 20 micrograms, while B is falling to a value near 0.15 micrograms/square root hour. For the Schott lamps, it appears possible that A may eventually surpass 20 micrograms, and B fall below 0.4 micrograms/square root hour. Also note that the changes are generally smaller as more data is included into the fit, for both Corning 1720 and Schott 8436 lamp data. Thus the A's and B's appear to be asymptotically approaching some final values.

One way to readjust Eq. 1 to reflect the observed trends is to introduce a time dependence for both A and B, so that Rb consumption $C = A(T) + B(T)\sqrt{T}$. However, the model equation is directly related to simple diffusion theory, in which M, the amount of material consumed is given by $M = 2AC_o\sqrt{(\eta T/D)}$ where A is the surface area, $C_o$ is the density of the penetrating species at the glass surface, and D is the diffusion coefficient.[3] B, in principle, is equal to $2AC_o\sqrt{(\eta/D)}$. Neither the surface area or the

CH2588-2/88/0000-525 $1.00 C 1988 IEEE

diffusion coefficient have a clear time dependence. The density of the penetrating species, Rb metal in the vapor phase, should remain constant at the glass surface as long as there is Rb metal left in the liquid reservoir of the lamp to constantly supply the vapor. Only at failure, when all the Rb metal in the liquid reservoir is gone, would the constancy of penetrating species not be met. It would not be consistent with this model to assume a time dependence for B.

In contrast to B, it is reasonable that A would have some time dependence. A represents the amount of metallic Rb lost through a reaction of the form,

$$Rb + I \xrightarrow{k} RbI, \tag{2}$$

with I representing impurities within the lamp at its fabrication, RbI a nonvolatile reaction product, and k the reaction rate constant. While initially assumed to be instantaneous this is not necessarily the case. With the concentration of Rb within the lamp constant, the time dependence of the amount of I is given by

$$[I] = [I]_o \, e^{-DT} \tag{3}$$

with

$$D = k\,[Rb],$$

the brackets implying the amount of a particular species within the lamp and the subscript zero indicating initial amount. The time dependence of the amount of metallic Rb lost through Eq. 2 is then given by

$$A(t) = A' \, (1-e^{-DT}), \tag{4}$$

with A' representing the total amount of Rb lost through reaction. The expanded consumption equation is

$$C = A'(1-e^{-DT}) + B'\sqrt{T}. \tag{5}$$

Subsequently this equation will be used to analyze consumption data initially studied using Eq. 1. At this point we now need to introduce the statistical procedures by which we may determine the parameters and compare the performances of Eqs. 1 and 5.

## CONSUMPTION STATISTICS

Practical lamp reliability analyses use Rb consumption data obtained over limited periods of lamp operation to predicted long term consumption. We wish to obtain most probable Rb consumptions and their variances for any operating times. Eq. 1, which may be put in a linear form, can be analyzed using normal linear least squares techniques. However, Eq. 5 can not be put in linear form and requires a more sophisticated analysis. We employ a general approach applicable to both equations.

Let $C_i$ be a consumption measurement made after a lamp operating time $t_i$. It is assumed that p such measurements are made. The functional form to which the measurements will be fit is $\overline{C}(t)=f(a_k,t)$ with the $a_k$ the parameters of interest. Assuming the $C_i$ are gaussian distributed with standard deviations $\sigma_i$, a likelihood function, L, may be written,[4,5]

$$L = \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi}\,\sigma_i} \exp\left[-(C_i-\overline{C}(t_i))^2/2\sigma_i^2\right]. \tag{6}$$

Setting $W = \ln(L)$ we find,

$$W = -\frac{1}{2} M - \sum_{i=1}^{p} \ln\left(\sqrt{2\pi}\,\sigma_i\right)$$

with

$$M = \sum_{i=1}^{p} (C_i-\overline{C}(t_i))^2/\sigma_i^2. \tag{7}$$

The "best" parameter values, $a_k^*$ are solutions of the set of equations,

$$\frac{\partial M}{\partial a_k} = 0. \tag{8}$$

The least-squares variances for the $a_k^*$ are given by

$$\overline{(a_\ell - a_\ell^*)(a_m - a_m^*)} = (H^{-1})_{\ell m}, \tag{9}$$

with the bar implying averaging over many sets of consumption measurements each yielding a set of parameters $a_k$. $(H^{-1})_{\ell m}$ is the $\ell$-th row, m-th column element of the k-row by k-column matrix whose inverse has elements defined by,

$$H_{rs} = \frac{1}{2} \frac{\partial^2 M}{\partial a_r \, \partial a_s}. \tag{10}$$

Another quantity of interest is the variance for a predicted Rb consumption at a time $t_p$. The predicted consumption is $\overline{C}^*(t_p)$, equal to $f(a_k^*,t_p)$, while its variance at $t_p$ is specified by[4,5]

$$\sigma_{\overline{C}^*}(t_p) = \sum_{\ell} \sum_{m} \frac{\partial \overline{C}(t_p)}{\partial a_\ell} \, \frac{\partial \overline{C}(t_p)}{\partial a_m} \, (H^{-1})_{\ell m} \tag{11}$$

The quantity that is of most interest in making a lamp life prediction is the Rb consumption $C_p$ at $t_p$ specified with a certain confidence. Based on the normal distribution, a 90% confidence that a lamp will consume no more than $C_p$ results if

$$C_p = \overline{C}^*(t_p) + 1.282\,\sigma_{\overline{C}^*}(t_p). \tag{12}$$

All of the analysis done is this paper is on the collective body of consumption data for a particular type of lamp. That is, all consumption data for GPS-style Schott 8436 lamps are lumped into one data set, as are all the consumption data for the Corning 1720 lamps. While it is possible to analyze each lamp individually, we have chosen to maintain consistency with the previous analysis, reference 1.

## APPLICATION OF RUBIDIUM CONSUMPTION FITTING TECHNIQUES TO LAMP DATA

The EG&G 1720 glass lamp consumption data were analyzed using the techniques presented in Section III. In Table 2 the parameter values for Eqs. 1 and 5 for various lengths of analyzed data are presented. Additionally the most probable and the 90% confidence level consumptions have also been projected for eleven years of operation. While A and B display clear trends as lamp operating times are increased, no such trends are apparent for A', D and B'. This gives us confidence in our supposition of a time dependent reactive consumption mechanism. The most probable

11 year consumptions also show similiar behaviors. For these lamps, Eq. 5 using data based only on one half of a year's lamp operation, yields accurate long term projected consumption. For Eq. 1 to yield similiar most probable consumptions two years of lamp operation is required. In Fig. 2, fits to 0.29 and 3.20 years of consumption data for both equations are presented. As expected, fits resultings from Eq. 5 do not change significantly between the two operating times while Eq. 1 fits show clear variations. The knee in the Eq. 5 fits, occurring at about 500-1000 hours of operation, is due to the completion of reactive consumption and growth of the diffusive process. In terms of 90% confidence limits both equations yields similiar predictions. The reason for this apparent contradiction is that the uncertainties associated with the Eq. 5 analysis are typically greater than those resulting from an analysis based on Eq. 1. Including the additional parameter in Eq. 5 leads to a need for more data to obtain consumption predictions with levels of confidence similiar to those resulting from Eq. 1.

The same analyses were applied to the Schott 8436 glass lamp consumption data. The results are summarized in Table 3. The trends displayed by the parameters associated with Eq. 1 are the same as displayed by the Corning 1720 glass lamps. Using Eq. 5 meaningful fits did not result until after four years of data had been analyzed. The reason is apparent when Fig. 3, which displays all consumption data fit to both functional forms, is reviewed. The knee region in the Eq. 5 fit has moved out to between $10\text{-}40 \times 10^3$ hours of operation. This fit would say that only after about $5 \times 10^4$ hours of operation, nearly five years, is the long term diffusive process becoming significant. At times less than this, consumption is dominated by the reactive component and attempts to extract diffusive parameters are not successful. Viewing the D coefficient in Table 3, we see it is only about 6% of the size of that in Table 2, consistent with the apparently slow nature of the reactive process for this glass.

For the Corning 1720 glass lamps after three years of operation B and B' were very close in magnitude. After 4.7 years of operation of the Schott 8436 lamps' B and B' are still quite different. While B continues to decrease it may require many more years before it is apparent whether it will actually match B'. This is important as the two model equations yield significantly different relative consumption contributions from reactive and diffusive processes. As the extremely long term consumption will be dominated by diffusion, when comparing the merits of various glass types it is desirable to have accurate diffusive parameters. One means of resolving the question of the relative contributions of diffusive and reactive consumption in Schott 8436 glass would be to perform simple chemical analyses on the lamp envelopes. As we believe the reaction product resides primarily on the envelope's inner surface, washing with water and comparing the Rb content of the rinsings with that Rb remaining in the glass would be useful. However, we do not intend to do this as these lamps are valuable long-term test subjects. At this point, we lean toward the diffusive coefficient produced by Eq. 5 as most representative of Schott 8436 and hence feel, from the point of long-term Rb consumption that it is an excellent glass, probably superior to Corning 1720.

## CONCLUSIONS

We have observed that the parameters resulting from fitting Rb consumption data to the standard consumption model $C = A + B\sqrt{T}$ often depends upon the length of lamp operating time the data spans. Typically as operating time increases the value of A is seen to increase while that of B decreases. This dependence is attributed to a nonconstant A. When A is modified to include a simple time dependence resulting from first order chemical reaction kinetics ($A(T) = A'[1-\exp(-DT)]$) the parameters of the expanded model no long display this sensitivity to the length of lamp operation.

The underlying importance of fitting the consumption data to one model or another is the reliability we eventually ascribe to a lamp based on the fit used. Seeing a consistent consumption prediction enhances the respect one has for a particular model. On the other hand, one can argue for the simpler model, particularly as we know that its predictions will, in general, be overly pessimistic. Furthermore, when fitting consumption data obtained at short operating times to the expanded model, a larger standard deviation is associated with the eleven year consumption prediction. This translates into a not too significant difference in the eleven year consumption number with a 90% confidence. It is clear from Table 3 that the expanded model needs more data than the simpler model to make consumption predictions with equal precision. This is the result of having to fit data to an equation with three parameters instead of just two. The fact that the consumption predictions from the expanded model are so consistent, regardless of the additional complications of fitting to an equation with three parameters, leads us to believe that the expanded model makes consumption predictions that are ultimately more accurate and is based on a more physically correct model. Also, we feel that the expanded model allows more accurate assessment of a glass's long-term diffusive behavior. At this point, though, we don't recommend discarding the standard model. Its simplicity and conservative approach to lamp reliability cause it to remain quite attractive. As additional long-term consumption data is generated a more definitive statement will be possible.

## REFERENCES


1. C. H. Volk et al., "Lifetime and Reliability of Rubidium Discharge Lamps for use in Atomic Frequency Standards," Proceedings of the 38th Annual Symposium on Frequency Control (IEEE Press, New York, 1984), pp 387-400.

2. C. H. Volk and R. P. Frueholz, "The Role of Long-Term Lamp Fluctuations in the Random Walk of Frequency Behavior of the Rubidium Frequency Standard: A Case Study," J. Appl. Phys. 57, 980, (1985).

3. R. H. Doremus, "Diffusion in Non-Crystalline Silicates," in Modern Aspects of the Vitreous States, Vol. 2, J. D. MacKenzie, Ed. (Butterworth Inc., Washington, DC., 1962), pp. 1-77.

4. J. Orear, "Notes on Statistics for Physicists", University of California Radiation Laboratory report No. UCRL-8417, 1958.

5. M. Annis et al., "On Statistical Estimation in Physics", Rev. Mod. Phys. 25, 818 (1953).

Table 1. Consumption Parameters, A and B, as Functions of Increasing Periods of Lamp Operation

EFRATOM Schott 8436 Glass Lamps

| For data taken to: (yr). | A (μg) | B (μg/$\sqrt{hr}$) |
|---|---|---|
| 0.80 | -2.6±5.9 | .864±.122 |
| 1.14 | -5.0±5.6 | .930±.100 |
| 1.71 | -3.8±5.5 | .902±.086 |
| 2.28 | 8.4±6.1 | .638±.081 |
| 3.43 | 12.8±5.9 | .550±.069 |
| 4.00 | 15.2±5.5 | .506±.057 |
| 4.57 | 18.1±5.4 | .456±.050 |

EG&G Corning 1720 Glass Lamps

| For data taken to: (yr) | A (μg) | B (μg/$\sqrt{hr}$) |
|---|---|---|
| 0.29 years | 13.3±2.9 | .341±.093 |
| 0.46 | 14.1±2.5 | .304±.069 |
| 1.37 | 15.3±2.0 | .261±.047 |
| 2.06 | 16.7±1.7 | .181±.030 |
| 2.28 | 18.3±1.4 | .176±.020 |
| 2.85 | 18.2±1.4 | .181±.019 |
| 3.20 | 18.7±1.3 | .168±.016 |

Table 2. Analysis of Corning 1720 Glass Lamps' Consumption Data

| Lamp Data Analyzed (yr) | A (μg) | B (μg/$\sqrt{hr}$) | $\bar{C}$*(11 yr) (μg) | $C_p$(11 yr) (μg) | A' (μg) | D ($10^{-3}$/hr) | B' (μg/$\sqrt{hr}$) | $\bar{C}$*'(11 yr) (μg) | $C_p$' (11 yr) (μg) |
|---|---|---|---|---|---|---|---|---|---|
| 0.29 | 13.3±2.9 | .341±.093 | 119±26 | 153 | 21.5±9.6 | 4.8±3.0 | .137±.234 | 64±63 | 145 |
| 0.46 | 14.1±2.5 | .304±.069 | 109±19 | 133 | 20.3±6.4 | 5.2±2.9 | .168±.144 | 73±38 | 122 |
| 1.37 | 15.3±2.0 | .261±.047 | 96±13 | 113 | 19.5±3.7 | 5.6±2.5 | .184±.073 | 77±19 | 101 |
| 2.06 | 16.7±1.7 | .181±.030 | 73±8 | 83 | 19.6±2.3 | 5.5±2.0 | .181±.035 | 76±9 | 87 |
| 2.28 | 18.3±1.4 | .176±.020 | 73±5 | 80 | 21.6±2.0 | 4.7±1.4 | .142±.024 | 66±6 | 73 |
| 2.85 | 18.2±1.4 | .181±.019 | 75±5 | 81 | 21.1±1.9 | 4.8±1.5 | .151±.022 | 68±5 | 75 |
| 3.20 | 18.7±1.3 | .168±.016 | 71±4 | 76 | 21.8±1.8 | 4.6±1.4 | .140±.020 | 65±5 | 71 |

Table 3. Analysis of Schott 8436 Glass Lamps' Consumption Data

| Lamp Data Analyzed (yr) | A (μg) | B (μg/$\sqrt{hr}$) | $\bar{C}$*(11 yr) (μg) | $C_p$(11 yr) (μg) | A' (μg) | D ($10^{-4}$/hr) | B' (μg/$\sqrt{hr}$) | $\bar{C}$*'(11 yr) (μg) | $C_p$' (11 yr) (μg) |
|---|---|---|---|---|---|---|---|---|---|
| 2.28 | 8.4±6.1 | .638±.081 | 206±20 | 232 | - | - | - | - | - |
| 3.43 | 12.8±5.9 | .550±.069 | 184±17 | 205 | - | - | - | - | - |
| 4.00 | 15.2±5.5 | .506±.057 | 172±13 | 189 | 73.3±22.0 | 2.9±.8 | .100±.149 | 104±25 | 137 |
| 4.57 | 18.1±5.4 | .456±.050 | 160±11 | 174 | 71.1±18.7 | 2.9±.9 | .117±.118 | 107±19 | 132 |

# Typical Rubidium Discharge Lamp

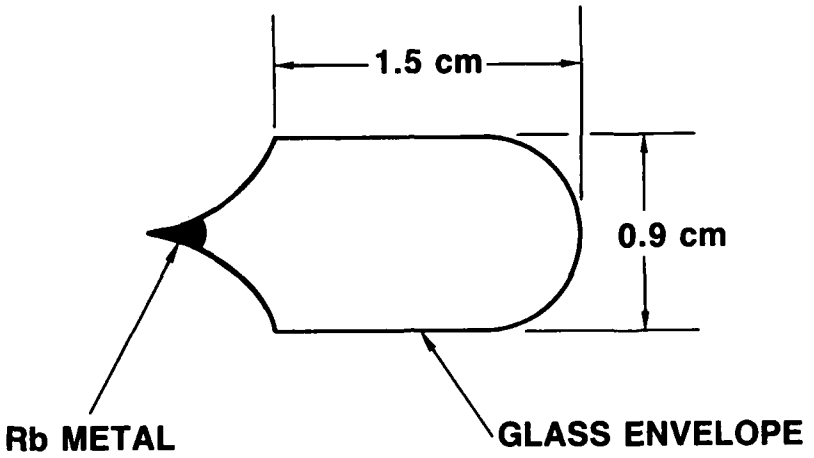



Figure 1. Schematic drawing of a rubidium discharge lamp. The internal surface areas of the Corning 1720 glass lamps studied are approximately 60% those of the Schott 8436 glass lamps investigated.

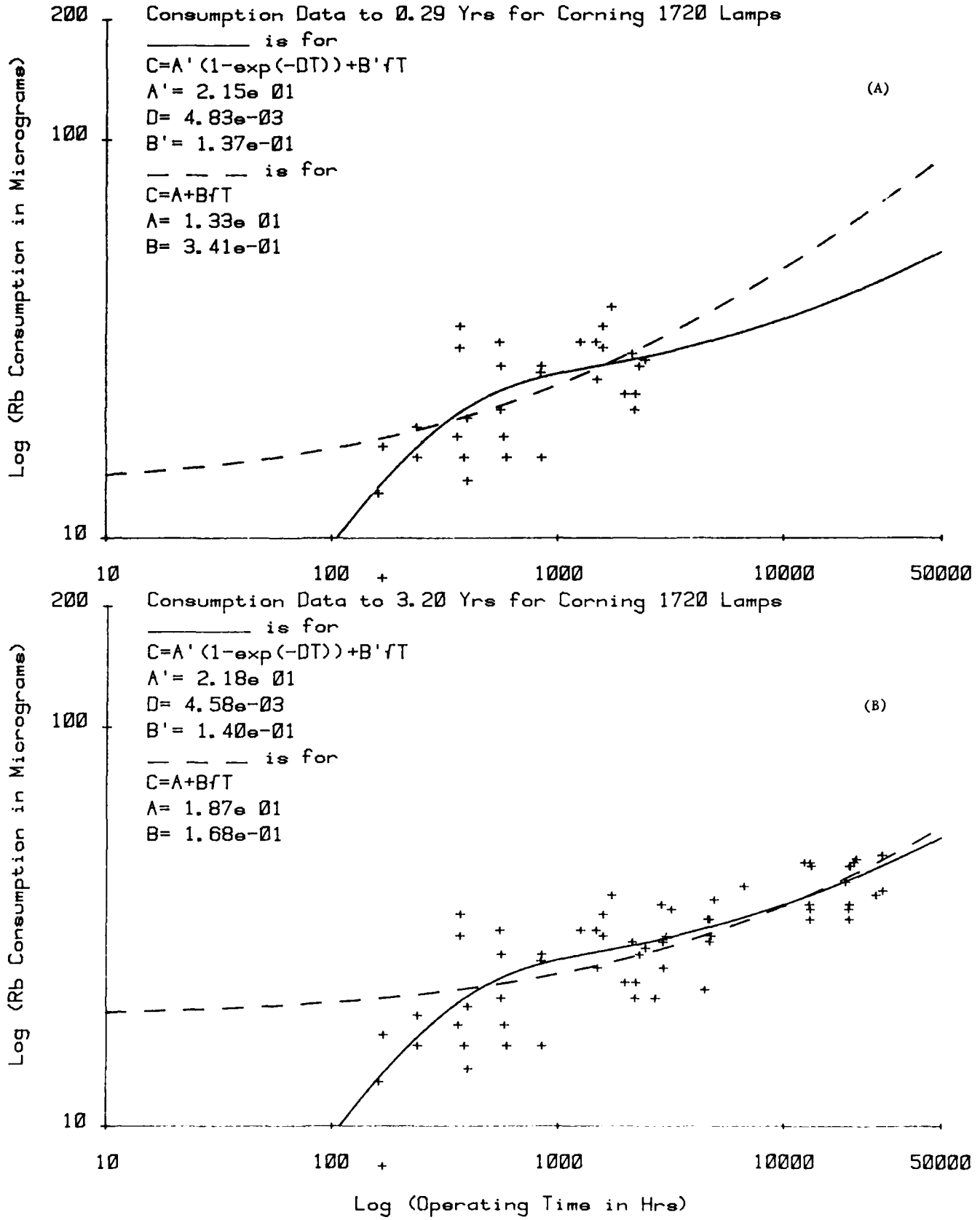



Figure 2. Most probable rubidium consumptions versus lamp operating time for Corning 1720 glass lamps based on the standard model, Eq. 1 (dashed line) and the expanded model Eq. 5 (solid line). Years of consumption data used in fitting procedures a) 0.29, b)3.20.

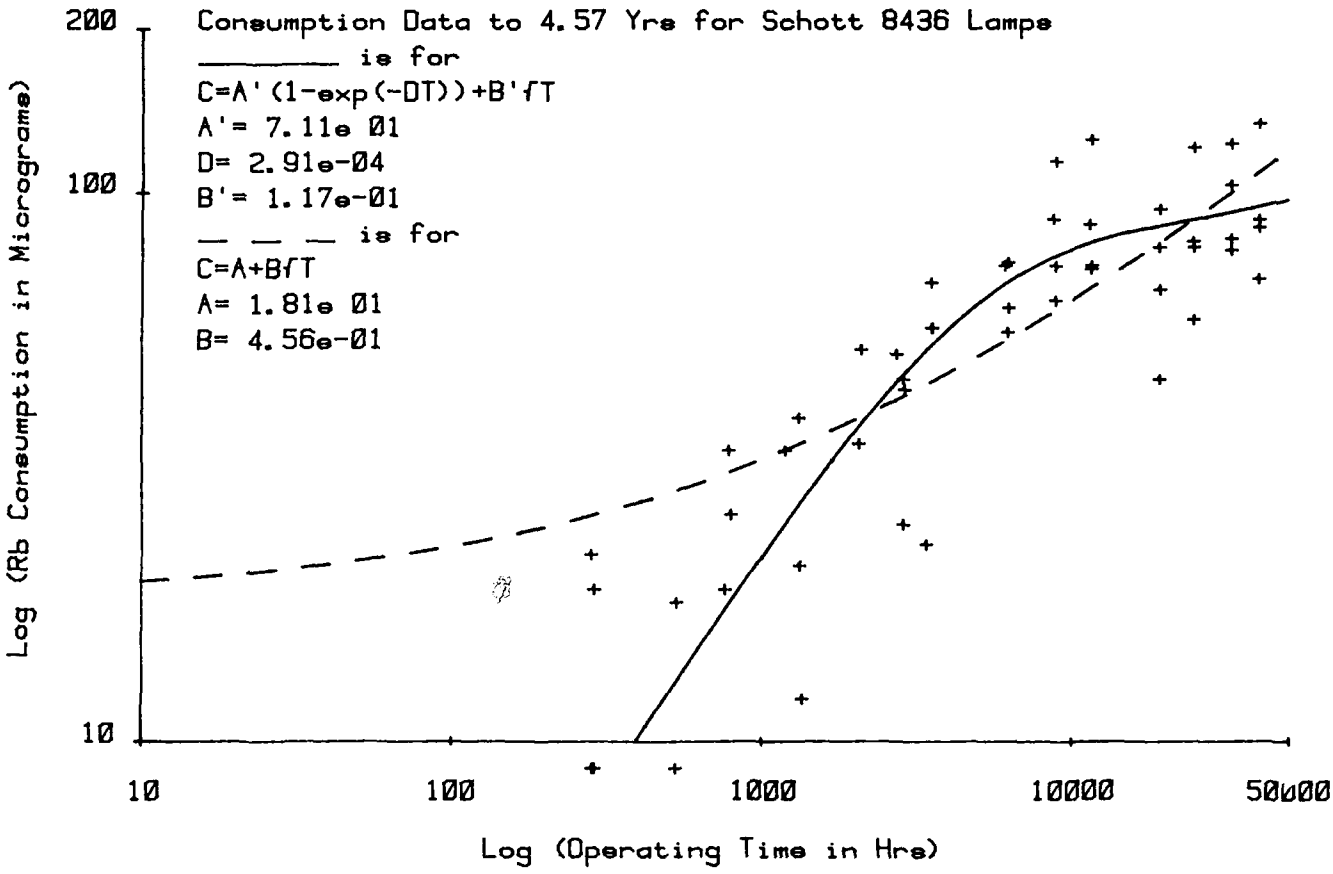



Figure 3. Most probable rubidium consumption versus lamp operating time for Schott 8436 glass lamps based on the standard model, Eq. 1 (dashed line) and the expanded model, Eq. 5 (solid line). 4.57 years of consumption data were used in fitting procedures.

42nd Annual Frequency Control Symposium - 1988

# NEUTRON HARDNESS OF PHOTODIODES FOR USE IN PASSIVE RUBIDIUM FREQUENCY STANDARDS

Thomas C. English & George L. Malley
Ball Corporation / Efratom Division
Irvine, CA 92718

Raj Korde
United Detector Technology
Hawthorne, CA 90250

## SUMMARY

The photodiode used for optical detection of the atomic resonance in a passive rubidium frequency standard (RFS) is a critical component that can limit the radiation hardness of the standard at high neutron-fluence levels. With the goal of obtaining a more neutron-resistant photodiode than is normally used in an RFS, experiments were carried out to evaluate the neutron hardness of several different types of planar silicon photodiodes. The diodes tested had approximately 1 $cm^2$ of active area, and were maintained at 100 C during neutron irradiation and testing. They were exposed to neutron fluences ranging from $7 \times 10^{11}$ to $1.4 \times 10^{13}$ $n/cm^2$. For each of the diodes, the dc short-circuit current, the ac short-circuit current, and the small-signal ac shunt resistance were measured as a function of neutron fluence.

Of the various types of diodes tested, the n/p types exhibited better performance than the p/n types normally used in RFSs. Superior performance for all of the measured parameters was obtained for n/p diodes fabricated on boron-doped p-p+ epitaxial substrates. Excellent performance was also obtained for conventional n/p diodes having a boron-doped substrate resistivity of 5 ohm-cm.

Detailed experimental results are presented, and are compared with previously published data, where possible. The tradeoffs inherent in using the different types of diodes for RFS operation are also discussed.

## INTRODUCTION

In recent years there has been an increasing trend toward the use of atomic frequency standards, especially passive rubidium frequency standards (RFSs), in military systems where operation and survival in a nuclear radiation environment is mandatory. As a result, significant effort, most of it unpublished and/or classified, has been directed toward improving the radiation hardness of RFSs. In the past, one area of concern has been the performance degradation of the RFS photodiode produced by neutron irradiation. This is the topic addressed in this paper.

Photodiodes of the p-on-n type have commonly been used as photodetectors in most RFSs. For neutron fluences above about $1 \times 10^{11}$ $n/cm^2$, the short-circuit current (for constant illumination) of this type of photodiode starts to drop rapidly, and although a well-designed RFS is relatively insensitive to even fairly large changes in the short-circuit current, the margin of safety for allowable photocurrent degradation decreases as the fluence level is increased.[1] In the present investigation, six different types of photodiodes were tested by measuring the changes in various parameters of importance for RFS operation as a function of neutron irradiation (fluence). In addition to this application-specific testing, tests of a fundamental nature were also conducted on an additional 40 photodiodes of different types. It is our intent to submit the results of this latter investigation for publication elsewhere.

## TEST DESCRIPTION

### Photodiodes

Six different types of planar silicon photodiodes, each having approximately 1 $cm^2$ of active area, were tested: n-on-p type fabricated on boron-doped p-p+ epitaxial wafers ("epitaxial" type); n-on-p type with boron-doped substrate resistivities of 5, 25, and 100 ohm-cm ("n/p" type); and conventional p-on-n type with substrate resistivities of 10 and 400 ohm-cm ("p/n" type). All of the diodes were square with 1 cm long sides. Leads were attached with solder to the metal contacts on the top and bottom of each device.

Two photodiodes of each type were tested to provide at least a crude measure of reproducibility (time and resources did not permit testing more than two of each type). All of the photodiodes used in these experiments were manufactured by United Detector Technology.

### Test Facility

Neutron irradiation and testing were carried out on December 15-16, 1987, at IRT Corporation's new Neutron/Gamma Range Facility[2] located in San Diego. Neutrons were provided by several californium-252 sources having a neutron-to-gamma ratio of $1 \times 10^9$ neutrons per rad (Si) of gamma radiation. The californium neutron spectrum has been extensively studied, and is well documented, thereby allowing 1-MeV (Si) damage-equivalent fluences to be accurately calculated. Comparison tests made between the IRT facility and the Sandia Fast-Burst Reactor confirm the results of these calculations.[2]

For neutron testing using this type of facility, the

CH2588-2/88/0000-532 $1.00 C 1988 IEEE

irradiation times needed to realize the higher fluence levels can be quite long. For the present experiments, irradiation times of several hours were required to attain fluences above $1 \times 10^{13}$ n/cm$^2$. Because of the long irradiation times and also possibly because of the 100 C temperature of the photodiodes during irradiation, annealing effects were not observed in these experiments (any annealing takes place rapidly during the irradiation, not after it).

Test Fixtures

The 12 photodiodes were placed, in groups of four each, in three separate test fixtures. Each test fixture consisted of an insulated, cylindrical copper oven that was electrically heated and thermostatically controlled to operate at 100 C during neutron irradiation and electro-optical testing.[3] The photodiodes were epoxied to the inside, flat, end surface of the oven. Care was taken to ensure that heat leaks through the electrical leads did not lower the temperature of a photodiode significantly. Each oven had a hole at the far end to admit rubidium light during electro-optical testing, thereby allowing all four photodiodes to be illuminated and tested simultaneously.

Light Source

The light source used for electro-optical testing was a rubidium lamp assembly from an Efratom Model M-100 rubidium frequency standard. The lamp itself contained both isotopes of rubidium and approximately 1.5 torr of xenon gas. The test fixtures were configured for three-point kinematic mounting on top of this light source so as to ensure precise positioning of each fixture before and after each of the four neutron-irradiation levels. Reference measurements of the light intensity of this source that were made several times during the two-day testing period, using an unirradiated reference photodiode, showed the maximum (peak-to-peak) variation to be <0.5%. This confirms both the previously documented[4] stability of this type of light source and the effectiveness of the kinematic mount used in these experiments.

The light source was amplitude modulated at 127 Hz for the ac measurements described below. The amplitude was very small compared to the dc level so that the modulation did not affect the dc measurements. Reference measurements of the modulation amplitude, taken at the same time as the dc measurements described in the previous paragraph, showed a maximum variation of 1.3% over the two days of testing.

Electro-Optical Measurements

The following electro-optical characteristics of the tested photodiodes were measured before and after each irradiation level for each photodiode: dc short-circuit current, ac short-circuit current, and small-signal ac shunt resistance. All of these quantities were measured in the photovoltaic mode under conditions of constant illumination using the light source described above.

Dc Short-Circuit Current The dc short-circuit current was measured using a μA739C op amp configured as a current-to-voltage converter. The input of this type of circuit is a virtual short so that the dc output voltage of the op amp is proportional to the short-circuit current for the photodiode. Four of these converters were used to provide simultaneous measurements of the dc short-circuit currents of all four photodiodes in a test fixture.

Ac Short-Circuit Current The 127 Hz intensity modulation of the light source (see above) produces a 127 Hz short-circuit current at the output of the photodiode. This short-circuit current is transformed into an ac voltage by the current-to-voltage converter, and then synchronously demodulated and filtered to produce a dc voltage that is proportional to the amplitude of the ac short-circuit current. Using four identical circuits that perform this function, measurements of the ac short-circuit currents from the four photodiodes in a test fixture were made simultaneously. In an RFS, the servo that locks the frequency of the voltage-controlled crystal oscillator to the rubidium reference frequency uses a low-frequency ac error signal. This signal is present as an intensity modulation of the light beam that is converted by the photodiode to an ac short-circuit current. Any degradation in this conversion process due to neutron bombardment of the photodiode is therefore of concern.

Shunt Resistance The term "shunt resistance" as used in this paper, unless stated otherwise, refers to small-signal ac shunt resistance. A definition of this quantity can be found in the Appendix. It is of importance for RFS operation because its value, if too small, can increase amplifier noise and significantly degrade the short-term frequency stability of the RFS (see the Appendix for details). A method of measuring it is also given in the Appendix.

Irradiation-Test Sequence

The iterative method of radiation testing was used whereby the component to be tested is first characterized, then exposed to nuclear radiation, then removed and recharacterized. This process is repeated incrementally to obtain data showing how the characteristics of the component vary with the accumulated nuclear-radiation dose. In these experiments, the three test fixtures were not always simultaneously irradiated with neutrons. Consequently, the three fixtures did not receive identical incremental radiation doses, as can be seen from the graphs of the results.

It should be noted that the light source and its electronics, the photodiode preamplifier and associated electronics, and the test fixture thermostat electronics (excluding the oven transistor heaters and several metal-film resistors) were not exposed to nuclear radiation during these experiments.

Dosimetry

Sulfur dosimeters were used to determine the incremental neutron fluences. The tolerance[2] of such dosimeters is usually taken to be ±15%, but cross correlation with NBS standards indicate the tolerance of IRT sulfur dosimetry to be less than 10%. The spatial uniformity of the irradiation was within the tolerance of the dosimeters. The fluences quoted in this report are those at the locations of the photodiodes (corrections have been made for

attenuation by the test fixture; the correction factor was .73 for all fluences).

## TEST RESULTS & DISCUSSION

### Dc Short-Circuit Current

Fig. 1 shows the variation of the normalized dc short-circuit current as a function of cumulative neutron fluence for each of the six different types of photodiodes. The "normalized short-circuit current" for a given fluence is defined as the short-circuit current of a photodiode at that fluence divided by the value of short-circuit current for the same photodiode before any exposure to nuclear radiation. This definition assumes constant illumination of the photodiode during testing which was the case here. Each data point plotted in Fig. 1 is the average of the normalized dc short-circuit currents for the two samples of that type at the indicated fluence. For the epitaxial and n/p photodiodes, the differences between the values for the two samples were less than 3% in all cases. For the p/n, 10 ohm-cm photodiodes the differences were less than 10%, whereas for the p/n, 400 ohm-cm photodiodes the differences were as great as 55% at the highest fluences. Thus, with the exception of the high resistivity p/n devices, the results for each pair of samples exhibit a high degree of consistency which gives added confidence in the results.

The dc short-circuit currents of the p/n photodiodes exhibited a much greater degradation due to neutron bombardment than those of the other photodiodes tested. Clearly, the p/n devices are less suitable for use in a neutron-radiation environment than the other types. Of the two types of p/n devices, those with the higher substrate resistivity showed greater susceptibility to neutron-induced dc short-circuit-current degradation at the higher fluence levels. As regards the low resistivity p/n devices, their performance is consistent with, but somewhat better than, the previously reported performances of large-area, p/n solar cells.[1]

### Ac Short-Circuit Current

Fig. 2 shows the results for the normalized ac (127 Hz) short-circuit currents. These results are very similar to those obtained for the dc short-circuit currents. For the epitaxial diodes, the ac results are nearly identical to the dc results, but for the other diodes, the neutron-induced degradation is slightly greater for the ac short-circuit currents than for the dc short-circuit currents, with the ac degradation exceeding the dc degradation by about 10-40% at the higher fluence levels. This same effect was observed previously when studying p/n solar cells[1]. Although it is not known with certainty what is causing this effect, the decrease in small-signal ac shunt resistance with increasing neutron fluence appears to be a

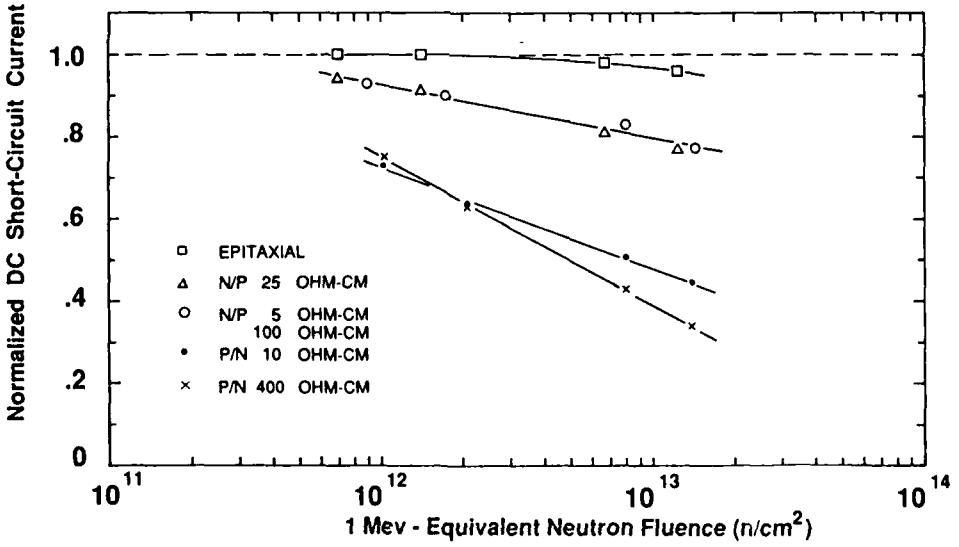



**Figure 1. Photodiode DC Short-Circuit Current Versus Neutron Fluence**

It is clear from Fig. 1 that the epitaxial photodiodes are the most resistant to neutron degradation of the dc short-circuit current. Even at the highest fluences, the dc short-circuit currents were still in excess of 95% of the preirradiation values. The n/p photodiodes also performed very well with the dc short-circuit currents being in excess of 75% of the preirradiation values at the highest fluences. These results for the n/p photodiodes appear to agree quite well with previously published results.[5,6] Fig. 1 also shows that the dc short-circuit current behavior for these latter photodiodes is nearly independent of substrate resistivity. To the best of our knowledge, the results for epitaxial devices are presented here for the first time.

contributor. As the ac shunt resistance approaches the smaller values (several kohms), significant ac photocurrent is shunted through it, thereby reducing the ac short-circuit current by a small amount. If the dc shunt resistance remains large enough that very little dc current is shunted through it, then the ac short-circuit current will decrease somewhat more rapidly than the dc short-circuit current.

### Shunt Resistance

A method of measuring the small-signal ac shunt resistance is given in the Appendix. Because of the way this method was implemented in these experiments, the

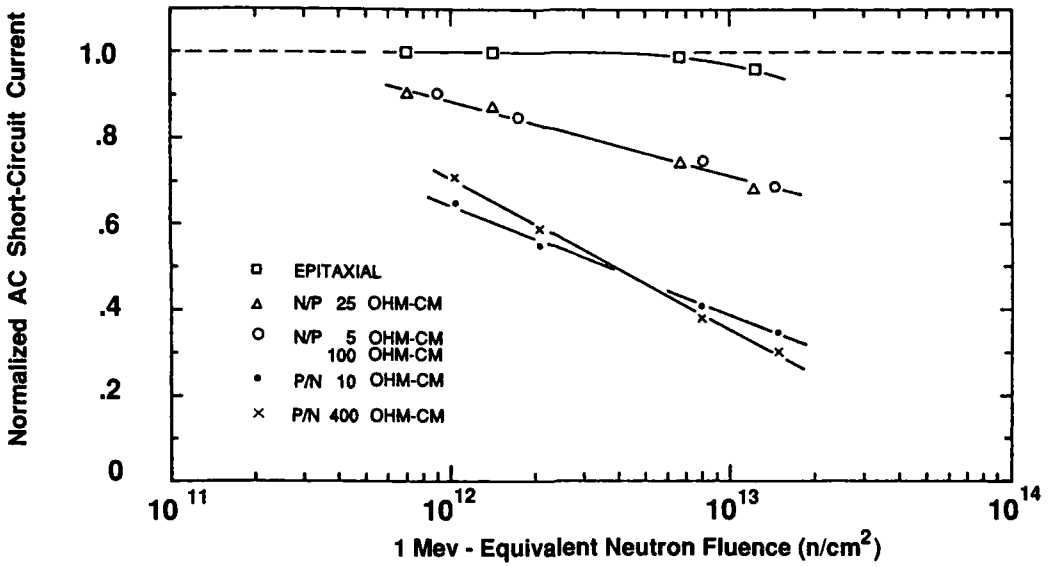



**Figure 2. Photodiode AC Short-Circuit Current Versus Neutron Fluence**

measured shunt resistance values are only approximate, and are expected to be somewhat smaller than the true values.

The small-signal ac shunt resistance results at 100 C are shown in Figs. 3 and 4. In most cases, the ac shunt resistances of the two samples of each type were sufficiently different that they were plotted as separate points. (The ac shunt resistance, like the "beta" of a transistor, is a parameter that is difficult to control tightly.) Without exception, the effect of neutron irradiation was to reduce the shunt resistance. The minimum acceptable shunt resistance for satisfactory RFS operation depends on light intensity and also on what type of op amp is used for the photodiode preamplifier; nevertheless, it is still possible to draw some general conclusions. The high resistivity devices (100 ohm-cm and above) are not suitable for RFS use because the shunt resistance of these devices drops to very low values, even at very low neutron fluences. The epitaxial and low resistivity n/p photodiodes exhibited the highest shunt resistances and would give satisfactory performance in most RFS's even at the highest neutron fluences ( $>10^{13}$ n/cm$^2$).

## DAMAGE MECHANISMS

Previous work[5,7] has established that the decrease in photocurrent of solar cells that occurs due to bombardment with nuclear radiation is due to displacement damage within the silicon crystal lattice. This type of damage reduces the diffusion length for minority carriers, thereby lowering the collection efficiency for the photo-generated electron-hole pairs in the substrate (base) material where a significant portion of the electron-hole production occurs at large distances from the p-n junction (depletion region). (The photo-produced minority carriers must diffuse from the base to the p-n junction before recombining in order to be collected.)

The improved radiation resistance of n/p photodiodes compared to p/n photodiodes is a direct result of the increased resistance of p-type silicon to displacement damage by neutrons; see Ref. 8 for details.

The additional radiation resistance of the epitaxial photodiodes results from a somewhat different effect. In these diodes, the substrate consists of a layer of very-heavily-doped p-type material (p+ layer) with a layer of "ordinary" p-type material epitaxially grown on the former. The rear metallization of the photodiode is then placed on the p+ layer. Because of the very high doping of the p+ layer (much higher hole density), photo-generated minority carriers (electrons) produced in this material are much more likely to recombine there than in the "ordinary" p-type material. Consequently, the effective collection volume of the base region is reduced, and unirradiated photodiodes of this type produce smaller photocurrents than the other types. However, they are also less affected by any reduction in the minority carrier diffusion length due to displacement damage because the rear portion of the effective collection volume is closer to the p-n junction.

At the highest neutron fluences used in this study ( $>10^{13}$ n/cm$^2$), the neutron irradiation is accompanied by a total gamma dose of approximately 10 krad(Si). However, this is not expected to have any significant effect upon the test results because gamma radiation produces very little permanent damage in silicon photodiodes.[9]

## CONCLUSIONS & DISCUSSION

The epitaxial photodiodes exhibited the highest neutron hardness in all categories tested. The neutron hardness of the 5 ohm-cm n/p photodiodes was also very good in all respects, and both types offer a considerable hardness improvement, with no major disadvantages, compared to the p/n photodiodes normally used in an RFS.

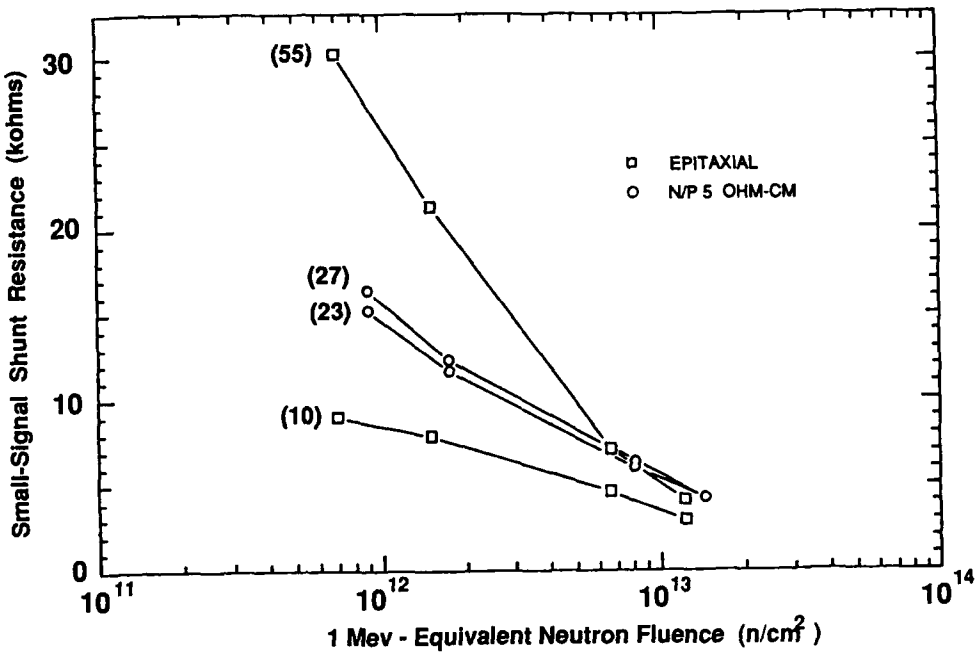



**Figure 3. Small - Signal AC Shunt Resistance Versus Neutron Fluence.** (The numbers in parentheses are the shunt resistances in kohms before neutron irradiation.)

There are, however, a few minor tradeoffs involved in the use of these diodes. The first is the higher unit cost of manufacturing small quantities of less-commonly-made devices. This amounts to roughly a factor of two for the n/p photodiodes in quantities of 100. For the epitaxial photodiodes, the need to use epitaxial wafers as the starting (substrate) material adds about an additional 10% to the unit cost. The second minor tradeoff is the somewhat lower responsivity (light sensitivity) of these devices. As an example of the latter for the low-substrate-resistivity devices tested in this investigation, the dc short-circuit currents under conditions of nearly identical illumination prior to neutron irradiation were 57, 70 and 88 μA, respectively, for the epitaxial, the 5 ohm-cm n/p, and the 10 ohm-cm p/n types. Thus, the epitaxial and n/p light sensitivities are about 65% and 80% of "normal", respectively. For most RFS designs this reduced light sensitivity can be compensated for by increasing preamplifier gain at the expense of a small decrease in signal-to-noise ratio.

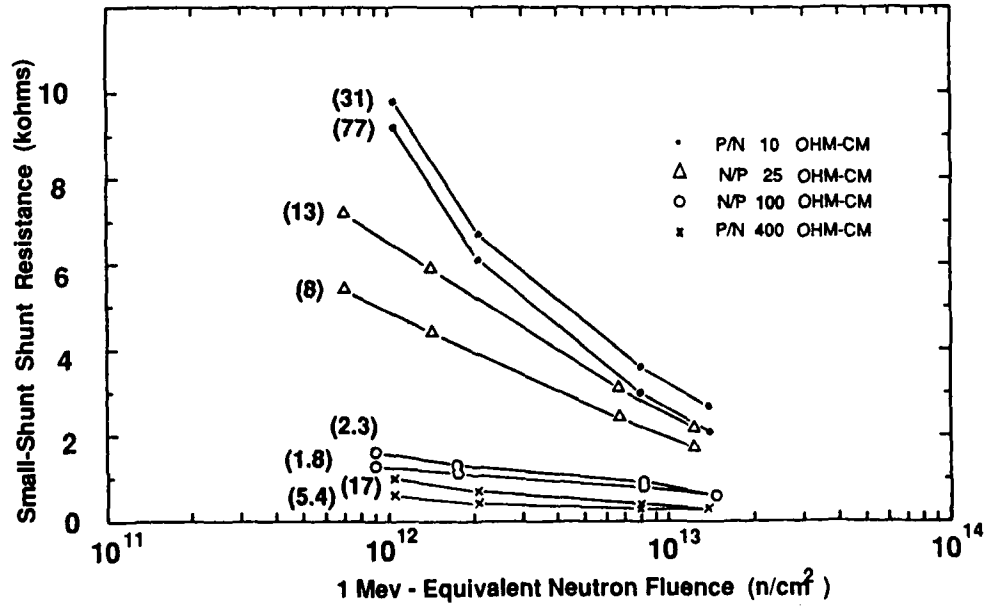



**Figure 4. Small - Signal AC Shunt Resistance Versus Neutron Fluence.** (The numbers in parentheses are the shunt resistances in kohms before neutron irradiation.)

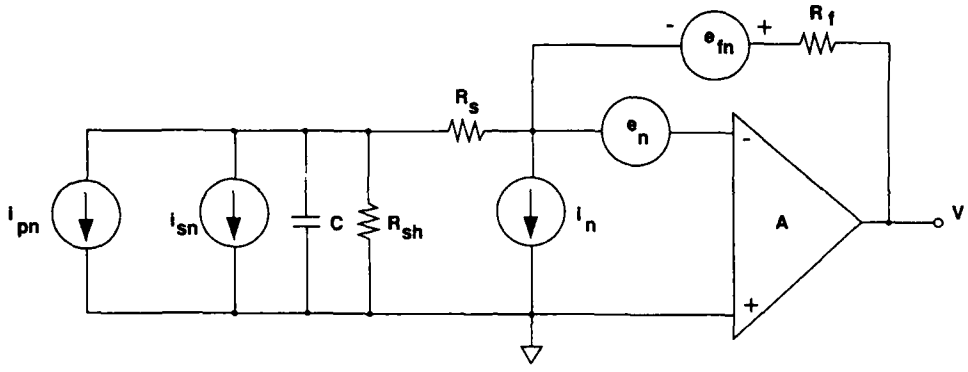



**Figure 5. Photodiode and Photodiode Preamplifier AC Equivalent Circuit with Noise Sources**

In closing, it should be noted that use of planar construction produces photodiodes having high shunt resistances. For the applications considered in this paper, this is an essential device feature. In addition, the use of p-type substrate material is essential for the manufacture of photodiodes having improved neutron hardness.

## ACKNOWLEDGEMENT

The authors would like to express their appreciation to Robert Braasch of IRT Corporation for his expert and helpful contribution to the success of this work in his role as the manager of the IRT Neutron/Gamma Test Range Facility.

## APPENDIX

In this appendix a noise model of the photodiode and its associated preamplifier is presented. Expressions for the various noise contributions are also presented, and special attention is given to the effect of small-signal ac shunt resistance on photodiode-preamplifier noise performance. Finally, the method used to measure photodiode small-signal ac shunt resistance is described.

### Noise Model

Fig. 5 shows the photodiode ac equivalent circuit along with a schematic diagram of the associated preamplifier. The shunt capacitance (C) represents the total parallel capacitance of the device, including the capacitance of the p-n junction. It was less than 10 nanofarads for all of the photodiodes tested, and is not of interest here, primarily because the modulation frequencies used in RFSs are typically very low (several hundred Hertz).

The schematic of Fig. 5 also includes all known noise sources. In this diagram, $i_{pn} = \sqrt{(2eI_P)}$ is the shot noise associated with the dc photocurrent $I_P$, $i_{sn} = \sqrt{(4kT/R_{sh})}$ is the Johnson noise associated with the small-signal ac shunt resistance, $e_{fn} = \sqrt{(4kTR_f)}$ is the Johnson noise associated with the feedback resistance $R_f$, $e_n$ is the op amp input-equivalent voltage noise source, $i_n$ is the op amp input-equivalent current noise source, A is the op amp open loop voltage gain, e is the magnitude of the electronic charge ($1.60 \times 10^{-19}$ C), k is Boltzmann's constant ($1.38 \times 10^{-23}$ J/K), and T is the absolute temperature in degrees Kelvin. As long as the noise fluctuations are small compared to the quiescent values, which is the case for a good op amp, the noise contribution of each noise source to the output voltage of the op amp may be found by using the principle of superposition, considering each noise source to

**Table 1. Noise Contributions to Output Voltage of Photodiode Preamplifier***

| Noise Source | Output Voltage Contribution |
|---|---|
| $i_{pn}$ | $i_{pn} R_f (1 + R_s / R_{sh})$ |
| $i_{sn}$ | $i_{sn} R_f (1 + R_s / R_{sh})$ |
| $i_n$ | $i_n R_f$ |
| $e_n$ | $[ 1 + R_f / (R_{sh} + R_s) ] e_n$ |
| $e_{nf}$ | $e_{nf}$ |

* Valid for $A >> 1 + R_f / (R_{sh} + R_s)$

be acting alone. The various noise contributions obtained in this way are summarized in Table 1. For a well-designed RFS, the contribution due to the shot noise of the dc photocurrent is dominant. However, as can be seen from the table, if the photodiode shunt resistance is too small,[10] the contribution from the op amp input-equivalent voltage noise source can become important and even overwhelm the shot-noise contribution. The other noise contributions are normally small by comparison, and can usually be neglected.

The values of $R_L$ used in these experiments were different for each photodiode tested, but once selected were kept constant throughout the experiments. The values were selected to produce a DV before irradiation of approximately 50 mV. The corresponding $R_L$ values were in the range of 15 - 500 ohms, depending on the value of $R_{sh}$ for the diode.

As the shunt resistances dropped with neutron irradiation, the DV's increased, typically being around 200 mV after irradiation at the highest levels. It has been our experience

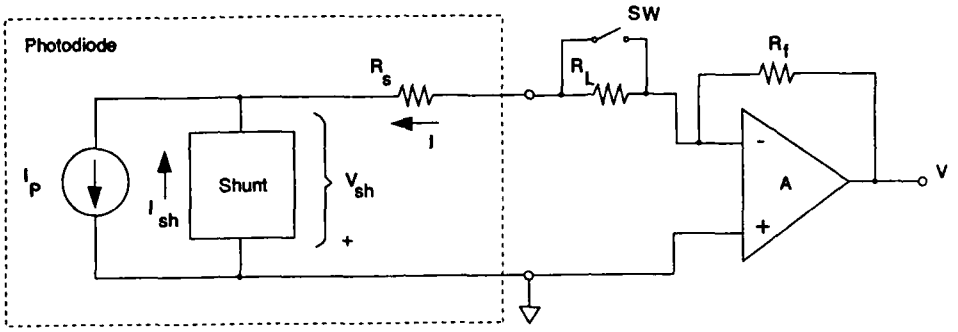



**Figure 6. Test Configuration Used for Measurement of Photodiode Ac Shunt Resistance**

Shunt Resistance Measurement

The overall I-V characteristic of a photodiode is a nonlinear function. As in other nonlinear electronic devices, the parameters that are useful for the description of small-signal and noise behavior are differential, or "small-signal ac" parameters. Parameters defined in this way are themselves nonlinear functions of the operating ("quiescent") conditions of the device (e.g., light intensity).

The test configuration used to measure the small-signal ac shunt resistance is shown in Fig. 6. The box labelled "shunt" represents all shunt current paths internal to the photodiode, and includes the p-n junction and any leakage resistances. This box is assumed to have an unknown, nonlinear I-V characteristic. The input resistance of the current-to-voltage converter is estimated from $R_f$ ( <100 kohm) and A (20 k) to have been about 5 ohms, and can be taken as a short in what follows.[10] The shunt resistance was measured operationally by taking readings of the op amp output voltage with the switch closed (short-circuit condition), and with the switch open (load condition) under conditions of constant illumination ($I_P$ = constant).

The small-signal ac shunt resistance $R_{sh}$ can be computed from the above readings by starting with the definition, $R_{sh} = dV_{sh}/dI_{sh}$. The derivative can be approximated by finite differences, $DI_{sh} = -DI = -(I_{sc} - I_L)$ and $DV_{sh} = I_{sc}R_s - I_L(R_s + R_L)$ to obtain $(R_{sh} + R_s) = \text{limit}\ (I_LR_L/DI)$ as DI --> 0. Since what is actually measured are the op amp output voltages, $V_{sc} = I_{sc}R_f$ and $V_L = I_LR_f$, with $DV = V_{sc} - V_L$, the final result is that $(R_{sh} + R_s) = V_LR_L/DV$, valid for small DV. Although this expression involves the series resistance, it is the expression of interest since the same quantity also appears in the noise equation for the $e_n$ contribution (Table 1).

that as DV is made smaller, the measured shunt resistance tends to increase, approaching the true value which is higher than the measured value. Ideally, the DV values should have been kept on the order of 10 mV or less for all shunt resistance measurements, but this would have meant changing $R_L$ for each photodiode after each irradiation, and this was not practical.

## REFERENCES


[1] Thomas C. English, Henry Vorwerk, and Norman J. Rudie, "Radiation Hardness of Efratom M-100 Rubidium Frequency Standard", Proceedings of the Fourteenth Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, NASA Conference Publication 2265, pp.547-574, Dec. 1982.

[2] Robert Braasch, Joseph Azarewicz and Douglas Willis, "(U) Testing with Californium-252 Neutrons", paper presented at the Instrumentation for Nuclear Weapons Effects Testing (INWET) conference, Washington, D.C., 6, October 1987 (to be published).

[3] 100 C is an upper limit on the operating temperature of photodiodes used in Efratom RFS's. Since small-signal ac shunt resistance degrades with increasing operating temperature, it is important to conduct the testing at the highest temperature likely to be encountered in an actual operating environment.

[4] The light source used for the present tests is the same type as used previously in the transparency test fixture of reference 1 (p. 556).

[5] E. J. Stofel, T. B. Stewart and J. R. Ornelas, "Neutron Damage to Silicon Solar Cells", IEEE Trans. on Nuc. Sci. **NS-16** (5), 97 (1969).

[6] There are several differences between the test conditions of Ref. 5 and the present work. The neutron fluences quoted in Ref. 5 are not specified there as 1-MeV-equivalent (Si) fluences. Also, short-circuit currents were measured using illumination that was intended to simulate space sunlight.

[7] H. Y. Tada, J. R. Carter, Jr., B. E. Anspaugh and R. G. Downing, Solar Cell Radiation Handbook, Third Edition, JPL Publication 82-69, 1982, pp. 3-13 to 3-16.

[8] H. W. Kraner, "Radiation Damage in Silicon Detectors", Nuc. Inst. & Methods in Phys. Res. **225**, 615 (1984).

[9] Ref. 7, pp. 3-40, 3-23.

[10] For the photodiodes tested here, the series resistances $R_s$ were of order 100 ohms or less, and are not expected to change with neutron bombardment (Ref. 5, p.100).

42nd Annual Frequency Control Symposium - 1988

# HIGH TEMPERATURE SUPERCONDUCTIVITY: NOVEL CONCEPTS, FUNDAMENTAL ISSUES, AND FREQUENCY CONTROL ASPECTS

Gerald J. Iafrate, U.S. Army ET&D Laboratory, Fort Monmouth, NJ 07703-5000
and Thomas E. Parker, Raytheon Research Division, 131 Spring Street,
Lexington, MA 02173

## ABSTRACT

A review of the basic technical issues relevant to the effective utilization of high temperature superconductivity in electronic device and component applications is presented. Many of the key issues pertain to material science; in this discourse, however, other more fundamental consequences of high temperature superconductivity are addressed, which inherently influence, and in fact, may limit the role of high temperature superconductivity in many conventional technology applications. BCS theory is used to estimate the characteristics of high temperature superconductivity and to provide insight into the use as well as the potential limitations of high temperature superconductivity in high-impact technology areas.

As a potential frequency control application, the predicted phase noise performance of a 10 GHz high $T_c$ superconducting cavity oscillator is analyzed. This performance is based on estimated values for Q and power handling capability and is compared to that of conventional frequency control technologies such as BAW and SAW oscillators and dielectric resonator oscillators. A significant reduction in phase noise can be obtained over existing technologies if a loaded Q on the order $10^7$ can be obtained. However, current high $T_c$ materials are not yet capable of providing such high Q values.

## 1. INTRODUCTION

The recent discovery of high temperature superconductivity rivals the invention of the transistor, both as a scientific novelty and in terms of potential technological impact. In the scientific community, this discovery is viewed as a remarkable manifestation of nature; much excitement and speculation has been generated concerning the impact of high temperature superconductivity on a wide variety of commercial and military applications. Although there is ample cause for enthusiasm, there are two major issues yet to be resolved before viable applications can be realized. The first, the materials issue, concerns the synthesis of stable, reproducible materials with useable electromagnetic properties and current densities. The second, the concepts issue, involves the elucidation of fundamental physical limitations inherent in the phenomena of high temperature superconductivity.

In this paper, a synopsis of the discovery is presented, and an elucidation of the fundamental technical issues embodying high temperature superconductivity is given. Some of the key issues are concerned with material science; in this article however, other more fundamental consequences of high temperature superconductivity are addressed, which inherently influence, and in fact, may limit the role of superconductivity in technological applications. As a specific example of a potential frequency control application, the predicted phase noise performance of a 10 GHz high $T_c$ superconducting cavity oscillator is presented. The basic microscopic theory of Bardeen, Cooper and Schrieffer (BCS) is used to estimate the characteristics of high temperature superconductors - - - the critical distances over which the superconductivity occurs, the coherence length, are quite small (~15 angstroms), and the critical magnetic fields needed to destroy the superconductivity have extreme type-II values, with lower critical fields of less than 1 kG and upper critical fields of well over 150 kG. The magnitude of the estimated characteristics has led us to identify the possible use as well as the potential limitations of high-temperature superconductors in high-impact technology areas. In addition, the magnitude of the coherence lengths are comparable to the size of a unit cell thereby suggesting that the high temperature superconductivity is localized to within lattice dimensions; this microscopic localization could have a profound influence on the superconductive transport and magnetic properties. The importance of including a temperature dependence in the electron-electron coupling strength is also highlighted.

## 2. PRESENT MATERIAL STATUS AND ISSUES

High temperature superconductivity is observed in the class of rare-earth compounds $(RE)Ba_2Cu_3O_{9-x}$ (RE refers to rare-earth) at temperatures of ~90K, well above the temperature of liquid nitrogen. The discoverers of this observation, K. A. Müller and J. G. Bednorz of IBM (Zürich), were recently awarded the 1987 Nobel Prize in physics for their efforts. To date, this material has been synthesized in polycrystalline bulk and oriented thin film forms; the high temperature superconductivity has been observed in these materials only when the oxygen content is seven oxygen atoms per molecule (x=2). The major material problem, commonly referred to as the oxygen problem, is to control chemically and maintain this specific oxygen content under a variety of different environmental and temperature cycling conditions.

To prepare these materials as superconductors, in both thin film or bulk forms, with the appropriate oxygen content, it is necessary to heat the materials at a temperature as high as 900°C in oxygen. When the materials are subsequently heated in vacuum, or in gasses, the materials are observed to lose oxygen. With loss of oxygen, these materials not only lose their superconductivity but they become semiconducting. The oxygen content, and therefore the superconductivity, can be restored by heating the materials to 550°C in oxygen. Thus the oxygen content must be maintained during all processing steps or must be restored as a final step; it is noted that most recently synthesis techniques have been developed to deposit high-$T_c$ thin films in oxygen containing atmosphere with excellent properties. This experimental fact has serious implications for surface and junction technologies. It is also noted that materials displaying good superconducting properties also degrade after several days of exposure to room ambients, especially humidity.

The need to expose the materials to 900°C annealing cycles in oxygen is very detrimental to useful application in microelectronics; 900°C is a very high

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

temperature compared to those used to process silicon and gallium arsenide circuits, which can tolerate temperatures of only 450°C and 350°C respectively after final processing. It therefore remains to determine the thermal and chemical stability, radiation damage sensitivity, and reaction kinetics with substrates and passivating overlayers of these materials as well as their compatability with manufacturing processing techniques.

Thin films of high temperature superconducting materials have been synthesized by painting, plasma spraying, evaporation, sputtering, and chemical vapor deposition onto various substrates. To date, the only process which has produced useful critical current densities of thin film material has been its evaporation on the substrate strontium titanate. All other methods have produced material with subcritical current density values. It is also noted that these ceramic materials have limited formability and therefore require special production techniques to prepare the materials in wire and sheet forms.

## 3.CHARACTERISTICS OF HIGH TEMPERATURE SUPERCONDUCTIVITY

The microscopic theory of Bardeen, Cooper and Schrieffer (BCS) is used to estimate the characteristics of high temperature superconductors. Within the BCS theory, the critical temperature, $T_c$, and the zero temperature energy gap, $\Delta_o$, are given by

$$\frac{1}{g} = \int_0^{\beta_c \hbar\omega_o/2} \frac{\tanh y}{y}\, dy \qquad (1)$$

and

$$\frac{1}{g} = \int_0^{\hbar\omega_o/\Delta_o} \frac{1}{(y^2+1)^{\frac{1}{2}}}\, dy \qquad (2)$$

where $g = D(E_F)V$ and $\beta_c = 1/kT_c$. As observed from equations (1,2), the critical temperature and the energy gap depend on two basic model dependent parameters: $g = D(E_F)V$ where $D(E_F)$ is the single particle density of states at the Fermi surface and V is the electron-phonon coupling constant, and $\hbar\omega_o$, the cut-off energy for the electron-electron attractive interaction. Other relevant parameters that characterize the superconducting state are the coherence length $\xi = \hbar V_F/\pi\Delta_o$, ($V_F$ is the Fermi velocity), the thermodynamic critical magnetic field $H_c=(4\pi D(E_F)\Delta_o^2)^{\frac{1}{2}}$, and the London penetration depth $\lambda_L = (mc^2/4\,ne^2)^{\frac{1}{2}}$ (m is twice the electron mass and n is one-half the conduction electron density).

BCS theory was developed originally to explain the occurence of low temperature superconductivity through phonon mediated electron-electron interactions. As such, V was identified as the electron-phonon coupling constant and, consequently, it has been shown that[1] $g \leq 1/2$; in addition, $\hbar\omega_o$ was approximately equal to the phonon cut-off energy, the Debye energy, so that $T_c \leq 35K$ in accordance with Eq.(1).

In this analysis, BCS theory is utilized as a generic superconductive electron-electron pairing theory, where V is considered to be a generalized electron-charge deformation coupling constant, and $\hbar\omega_o$ is assumed to be a cut-off energy for high frequency electron collective excitations.

For arbitrary values of "g" and $\hbar\omega_o$, the zero temperature energy gap is found by direct integration of Eq.(2) to be

$$\Delta_o = \hbar\omega_o / \sinh(1/g) \qquad (3)$$

Solving for $\hbar\omega_o$ in Eq(3) and substituting this result into the upper limit of Eq.(1), we get

$$\frac{1}{g} = \int_0^{[R(g)/4]\sinh(1/g)} \frac{\tanh y}{y}\, dy \qquad (4)$$

where R(g) is given by

$$R(g) = 2\Delta_o / kT_c \qquad (5)$$

Although $2\Delta_o$ and $kT_c$ depend on both "g" and $\hbar\omega_o$, it is clear that $R(g) = 2\Delta_o/kT_c$ depends only on "g". A straightforward analysis of Eq(4) shows that R is bounded, with a range $3.51 < R < 4$ as $\infty > g > \infty$. A numerical integration of Eq(4) shows the explicit dependence of R upon "g"; the resulting numerical integration is displayed in Fig. I.

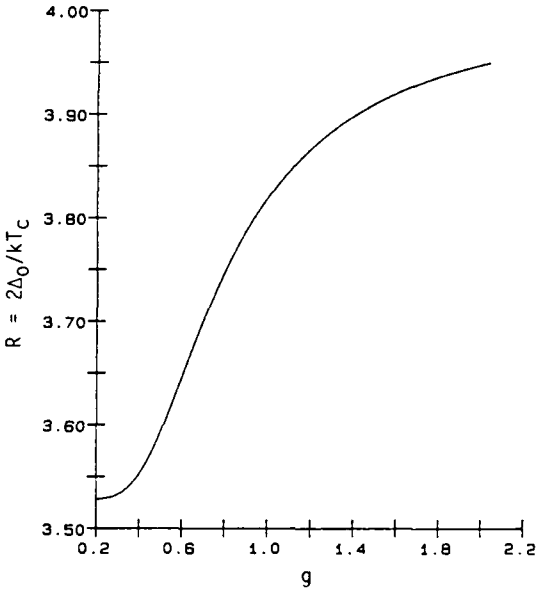



FIG.I Numerical Integration of Eq.(4) to obtain $R(g) = 2\Delta_o/kT_c$

Interestingly, there have been many reported experimental observations[2] of $R>4$ in high temperature superconducting materials. Actually, values of $R>4$ can be achieved within BCS theory if the generalized electron-charge deformation coupling constant V used in "g" is assumed to be temperature dependent. Letting $g_c = D(E_F)\,V(T_c)$ and $g_o = D(E_F)\,V(T=0)$, it can be shown from the BCS theory that (assuming that the cut-off energy for the interaction is independent of temperature)

$$R = 4\frac{g_o}{g_c} \qquad (6)$$

in the strong coupling limit ($g_c >> 1$) and

$$R = 3.51\, e^{-\frac{1}{g_o}\left(1-\frac{g_o}{g_c}\right)} \qquad (7)$$

in the weak coupling limit ($g_c, g_o < \frac{1}{2}$). For values[3] of

$$\frac{g_o}{g_c} > 1$$

it is clear from Eq. (6) that R can be greater than four in the strong coupling limit, and less than "3.51" (as seen from Eq. (7)) in the weak coupling limit.

It then follows from this analysis that the coherence length can be written explicitly as

$$\xi = 2\hbar V_F / \pi R(g) k T_c \qquad (8)$$

where R(g) is bounded as shown in Fig. I. The inverse dependence of coherence length on critical temperature raises a major point concerning the spatial extent of the high temperature superconductive pairing interaction. The coherence lengths are quite small at $T_c \simeq 95K$, approaching the size of a unit cell. For example, using values of $2\Delta_o / kT_c = 3.9$ and $\Delta_o = 20$ mev from recently reported[2] results on single crystal thin films of $YBa_2Cu_3O_7$, we find that $g \simeq 1.42$ from Fig. I and $\hbar\omega_o \simeq 11.48$ mev from Eq.(3). Using these parameters and rough estimators for $V_F$ and $D(E_F)$, the coherence length, the distance over which two electrons attract in the superconductive pairing interaction, is quite small, about 15 angstroms, and the London penetration depth, the penetration distance of a magnetic field into the superconductor, is quite large, about 1400 angstroms. For these magnitudes of coherence length and penetration depth, the superconductor is categorized as an extreme type-II superconductor with an extremely small lower critical magnetic field and huge upper critical magnetic field (the lower critical magnetic field is the field below which the superconductor is perfectly diamagnetic; the upper critical magnetic field is the field above which the superconductivity is totally quenched); lower critical fields are very low giving rise to limited effectiveness in shielding large magnetic fields (see Table I).

It is interesting to note that in high temperature superconductors, the coherence lengths are about the size of a molecular unit cell so that the superconductivity is confined to a very small microscopic region of the sample. In contrast (see Table I), in a low temperature superconductor such as niobium, the coherence lengths are many hundred times larger than a unit cell thereby rendering the superconductivity truly macroscopic. The localized nature of the superconductivity makes the high temperature superconductors more sensitive to radiation and very unattractive for technologies that require ultrathin, laterally homogeneous, spatial tolerances such as Josephson junctions.

Because high temperature superconductivity is confined to an atomically microscopic region, there will be many limitations to conventional applications of this material even if it were fabricated in a perfect form. The atomic confinement gives rise to very anisotropic superconducting properties and suggests the need for a transport theory based on electron pair hopping; the anisotropy may in fact be helpful in exploiting many unconventional applications not thought possible with relatively isotropic, low temperature superconductors - this conjecture remains as a challenge and opportunity to explore novel applications of high temperature superconductivity heretofore not possible with low temperature superconductivity.

TABLE I. COMPARISON OF LOW/HIGH TEMPERATURE SUPERCONDUCTOR PROPERTIES

| | CONVENTIONAL LOW $T_c$ | NEW HIGH $T_c$ |
|---|---|---|
| CRITICAL TEMP | ~1-25°K (HELIUM TEMP) | ~95°K(LIQ N TEMP) |
| CURRENT DENSITY | ~ $10^6$ Amp/cm$^2$ | ~ $10^3$ Amp/cm$^2$<br>$10^5$ Amp/cm$^2$ on $SrTiO_3$ |
| CRITICAL MAGNETIC FIELDS | | |
| $H_c$ | ~0.6 kGauss | ~10 kGauss |
| $H_{c1}$ | ~0.4 kGauss | ~0.7 kGauss |
| $H_{c2}$ | ~ 3 kGauss | $\gtrsim$ 150 kGauss |
| PENETRATION LENGTH | ~400 Å | ~1400 Å |
| CRITICAL LENGTH | ~1000 Å | ~ 15 Å |
| ENERGY GAP | ~ 0.1 mev | ~50. mev |
| MECHANISM | -LATTICE DEFORMATION | -ELECTRONIC |

The atomic nature of high temperature superconductivity inferred by such small coherence lengths suggests that the macroscopic electromagnetic and transport properties may well be described by electron-pair dynamics and dispersion through a large array of identical superconducting cells or weak links. The author is currently pursuing a high temperature phenomenological theory based on such a model; detailed results of the theoretical analysis as well as comparison with available experimental parameters will be forthcoming shortly.

Current experimental evidence indicates that the superconductivity arises from the interaction of electrons from neighboring atomically thin sheets of dense electric charge through interconnected electron channels; these channels are believed to arise from bridged oxygen atoms present in the $(RE)Ba_2Cu_3O_7$ configuration. While the presence of oxygen is vital to the occurrence of high temperature superconductivity, its volatility in the perovskite system makes the material somewhat unstable to environmental and external chemical activity. As an alternative it might be possible to develop man-made structures that provide an interaction similar to the one described above thereby yielding controllable, high temperature superconductivity under more desirable environmental and device processing conditions.

In summary, the fundamental technological issues described herein must be addressed and resolved if high temperature superconductivity is to be realized in many commercial and military applications. The primary challenge is to identify opportunities for possible near term technology insertion, while looking

for new applications heretofore not possible with low temperature superconductors, and exploring alternatives to ceramics, perhaps artifically created structures, to exploit this remarkable phenomenon for future technology applications.

## 4. ANALYSIS OF HIGH $T_C$ SUPERCONDUCTING CAVITY OSCILLATOR

One potential application for high $T_C$ superconducting materials that is important to the frequency control community is for use in high Q microwave resonators. In this section, estimates will be made of the phase noise performance of a microwave oscillator using a hypothetical high Q superconducting resonator. These phase noise levels will be compared to that which is achieved with conventional oscillator technologies, and this will provide guidance as to what level of performance is required from the superconducting resonator. Conventional, low temperature superconducting cavities have demonstrated Q's greater than $10^9$[4], but it will be shown that Q's of this order are not necessary to achieve a significant improvement in the performance of an X-band microwave frequency source.

The estimates of phase noise performance will be made for a simple feedback oscillator using a superconducting resonator as shown in Fig. II.

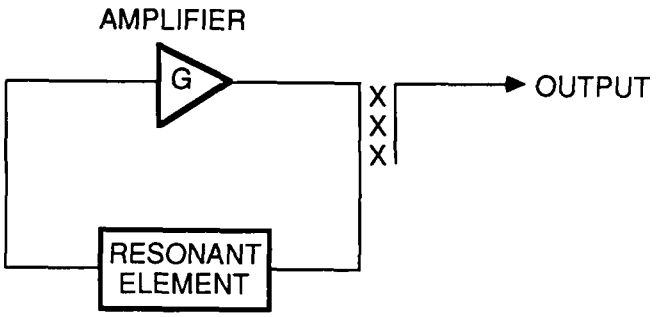



FIG. II Diagram of a simple feedback oscillator

A modified version[5] of Leeson's equations[6] will be used to make these calculations using estimated performance parameters for the superconducting resonator and realistic parameters for the amplifier. These estimates will be made for a resonator operating temperature of 77°K (or above) since the likelihood of large scale military or commercial applications is small if liquid He temperatures are required. The resonator is assumed to operate at 10 GHz and could be an empty or dielectrically loaded cavity, or any other resonant structure that can be constructed with superconducting materials. A two-port resonator with approximately 6 dB of insertion loss will be assumed, along with an incident power level on the resonator of +10 dBm. The flicker noise level of the resonator at 1 Hz is assumed to be much lower than the flicker noise of the amplifier ($\mathcal{L}(1) << -115$ dBc/Hz), though there is no justification for this assumption at this time. If the resonator has a flicker noise level larger than the amplifier noise level, there will be a dB for dB degradation in the close-to-carrier phase noise level of the oscillator.

The phase noise calculations have been made using loaded Q's for the resonator in the range of $10^6$ to $10^7$. It is possible that these high values may only be achieved in cavity resonators, as opposed to transmission line resonators. The Q values above correspond to surface resistance values in a microwave cavity in the range of $10^{-3}$ to $10^{-4}$ ohms/square. These values for loaded Q are low compared to those achieved in superconducting cavities using conventional, low temperature materials but are high compared to what has currently been achieved with the new high $T_C$ materials[7]. (This will be discussed in more detail later.) For a loaded Q of $10^7$ and an incident power of +10 dBm, the stored energy in the resonator is approximately 10 microjoules and the average magnetic field strength in an empty cavity would be about 35 gauss. These do not appear to be unrealistically large values.

The amplifier used in this oscillator is assumed to be a GaAs FET based amplifier with a gain (at 3 dB of compression) of 10 dB, a noise figure of approximately 4 dB and, as mentioned above, a flicker noise level at 1 Hz of $\mathcal{L}(1) = -115$ dBc/Hz. If a 3 dB power divider is used to couple out the oscillator signal, the amplifier output power level in 3 dB of compression would have to be +13 to +14 dBm to achieve the +10 dBm of incident power assumed above. All of these parameters are realistic values for a GaAs FET amplifier operating at room temperature (the flicker noise level may be a little optimistic), and some improvement may be achieved if the amplifier were operated at the reduced temperature of the resonator.

One other assumption that was made in making these calculations was to neglect environmental effects on phase noise, such as sidebands induced by vibration and temperature fluctuations. In real applications these factors could not be ignored and may turn out to be a major limitation at noise frequencies less than a few kHz for superconducting resonator based oscillators. In general, to take full advantage of the high Q of superconducting resonators, the temperature coefficient and vibration sensitivity of the resonator would have to be comparable to that of high quality bulk acoustic wave (BAW) resonators. At this time it is not clear that this is possible with the new high $T_C$ materials.

Figure III shows the calculated phase noise of the superconducting resonator oscillator (SRO) described above for two loaded Q values of $10^6$ and $10^7$. (Note that the close-to-carrier noise varies as $1/Q_L{}^2$.)

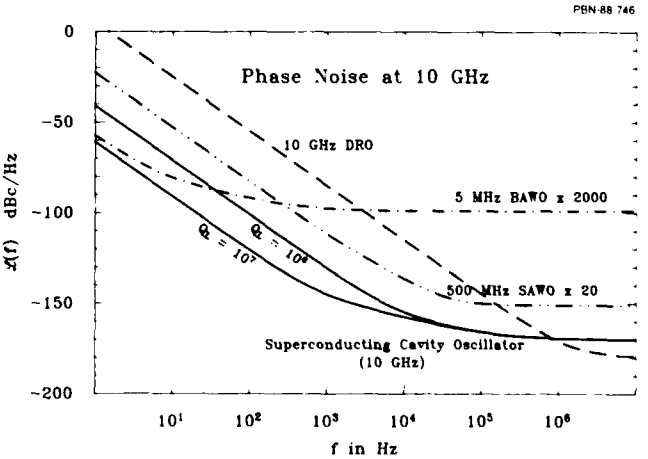



FIG. III Estimated phase noise of a 10 GHz superconducting resonator oscillator for loaded Qs of $10^6$ and $10^7$. Also shown are phase noise levels for three alternative 10 GHz sources.

Also shown are phase noise levels for three other alternative 10 GHz frequency sources. These include a 10 GHz room temperature dielectric resonator oscillator (DRO), a 500 MHz surface acoustic wave oscillator (SAWO) with a times 20 multiplier, and a 5 MHz BAW oscillator with a times 2000 multiplier. The curve for the BAWO source is based on an ideal multiplier chain and measured data for a state-of-the-art 5 MHz oscillator[8]. The curve for the 500 MHz SAWO is based on a state-of-the-art SAWO and actual measured data at X-band that demonstrated ideal multiplication[9]. The curve for the DRO is based on a hypothetical, but realistic, oscillator.

One important feature of the curves in Fig. III is that the DRO has the lowest noise floor of all of the various types of oscillators. Though the specific values of the parameters used in these calculations may be open to adjustment, it is reasonably safe to assume that the power handling capability of a room temperature X-band dielectric resonator (or empty cavity) will likely always be higher than that of a cryogenically cooled superconducting resonator. It, therefore, follows that the DRO will probably always have a lower noise floor than the superconducting cavity oscillator. A second important feature of the curves in Fig. III is that the close-to-carrier phase noise of the SRO drops below that of the multiplied BAWO only if the loaded Q of the resonator is at or above approximately $10^7$. For a loaded Q of $10^6$, the SRO is noisier than the BAWO source at noise frequencies less than about 50 Hz but the SRO is about 20 dB better than the SAWO source at all noise frequencies.

If one is going to resort to the added complexity, weight and size of a cryogenically cooled frequency source, a substantial improvement in performance over existing technologies should be expected. Therefore, as a standard for comparison, a complex, and large system should be considered which consists of a 10 GHz DRO phase locked to a multiplied SAWO, which is phase locked to a multiplied BAWO. With the proper choice of bandwidths for the phase lock loops, a system can be assembled that results in a frequency source that has the phase noise characteristics of the 5 MHz BAWO x 2000 source in Fig. III out to a noise frequency of about 200 Hz and then transitions to the SAWO x 20 source. At approximately 100 kHz, the system would then transition to the DRO noise characteristic. With this as a comparison, it is clear that the superconducting resonator would have to have a loaded Q larger than approximately $10^5$ before the superconducting resonator oscillator even began to compete with the system based on existing oscillator technology. With a loaded Q of $10^6$, significant improvement is achieved in parts of the noise spectrum by the SRO, but not at all frequencies. Only when the loaded Q reaches $10^7$ does the superconducting resonator oscillator offer across the board improvement in close-to-carrier noise.

At this time a loaded Q of only about $10^4$ has been demonstrated in a high $T_c$ material (YBaCuO) at 77°K and a Q approaching $10^5$ was only achieved at temperatures near 10°K[7]. It is very likely that improvements in the quality of the high $T_c$ materials will result in higher Q's, however, it is not at all clear that materials with $T_c$ less than 100°K will ever produce Q's in the range of $10^6$ to $10^7$ at X-band at 77°K. The best possibility for achieving useful Q's at 77°K is the development of materials with $T_c$ near 150°K. Progress is being made in this direction, but it is still too soon to speculate as to whether this is a realistic goal.

If room temperature superconductors are ever achieved, it would be fair to use the DRO alone as a comparison. In this case, loaded Q's of $10^5$ would offer significant improvement in close-to-carrier phase noise. However, high microwave Q's at room temperature would require a $T_c$ near 600°K, and this is beyond the point of legitimate speculation at this time. Another possibility is a significant advance in low temperature technology that would provide temperatures below 100°K at more acceptable limits in size, weight and power consumption. Again, the DRO alone could be used as a comparison and loaded Q's on the order of $10^5$ would become the threshold for usefulness. However, until such advances occur, high $T_c$ superconductor resonators will have to demonstrate loaded Q's of at least $10^6$ at 77°K before they should be considered for use in viable low noise x-band sources.

## 5. ACKNOWLEDGEMENTS

The authors acknowledge assistance from E. Potenziani for the numerical integration of Eq.(4), and from C.S. Kavina for the preparation of the manuscript.

## 6. REFERENCES

1. J. M. Blatt, Theory of Superconductivity, p.206, Academic Press, New York (1964).

2. J. Kirtley, R.T. Collins, Z. Schlesinger, W. Gallagher, R. Sandstrom, T. Dinger, and D. Chance, Phys Rev B., V35, 8846 (1987).

3. A. Rothwarf, Phys. Lett., V26A, 43 (1967).

4. S.R. Stein and J.P. Turneaure, "Superconducting-Cavity Stabilized Oscillators with Improved Frequency Stability," Proc. of the IEEE, Vol. 63, pp. 1249-1250, Aug 1975.

5. T.E. Parker, "Characteristics and Sources of Phase Noise in Stable Oscillators", Proc. of the 41st Annual Symp. on Frequency Control, 1987, pp. 99-110.

6. D.B. Leeson, "A Simple Model of Feedback Oscillator Noise Spectrum," Proc. of the IEEE, Vol. 54, No. 2, pp. 329-330, Feb 1966.

7. C. Zahopoulous, W.L. Kennedy and S. Sridhar, "Performance of a Fully Superconducting Microwave Cavity Made of the High $T_c$ Superconductor $Y_1Ba_2Cu_3O_y$" (to be published).

8. Brightline Corp. Cedar Park, Texas, Model 1039.

9. G.K. Montress, T.E. Parker and M.J. Loboda, "Extremely Low Phase Noise SAW Resonator Oscillator Design and Performance," Proc. of the 1987 IEEE Ultrasonics Symposium, 1987, pp. 47-52.

42nd Annual Frequency Control Symposium - 1988

# SUPERCONDUCTING RESONATORS AND HIGH-$T_c$ MATERIALS*

D. E. Oates, A. C. Anderson and J. Steinbeck**
M.I.T. Lincoln Laboratory, Lexington, Massachusetts 02173

## ABSTRACT

The application of the new metal-oxide high-transition-temperature superconducting materials to low-noise oscillators operating between 1 and 10 GHz is discussed. Stripline resonator structures are analyzed and values of Q and oscillator power are calculated. These parameters are used to calculate the expected values of phase noise, which are better than those of other technologies. Measurements of residual noise in niobium resonators have been made to identify possible sources of flicker noise in superconducting resonators. Results show good performance using niobium.

## I. INTRODUCTION

The recent discovery of the new metal-oxide, high-transition-temperature superconducting materials[1] such as $YBa_2Cu_3O_7$, BiSrCaCuO and TlBaCuO has stimulated interest in many applications, including high-quality-factor resonators which could be operated at liquid-nitrogen temperatures (77 K). Such resonators would be the basis for construction of low-phase-noise oscillators for application to radar and communication systems. The use of superconductors is of interest because extremely high quality factors (Qs) are possible, and because of the high power-handling capabilities of superconductors. Qs of $10^6$ for thin-film stripline resonators operating at 1 GHz and $10^5$ at 10 GHz are achievable with materials such as niobium at liquid-helium temperatures.[2] Even higher Qs have been achieved using superconducting cavities, but for reasons discussed below we feel that the stripline structure is preferred.

The resonator Q is determined by the radio-frequency surface resistance $R_s$ of the superconducting material which, although low, is not identically zero as it is at DC. The finite RF resistance can be easily understood from a two-fluid model. According to this model, at any nonzero temperature less than $T_c$, superconducting electron pairs and normal electrons exist intermingled in the superconductor. For DC current flow, the superconducting electrons shield the normal electrons from electric fields. At non-zero frequencies the inertia of the electron pairs prevents them from following the applied field, resulting in a non-zero time-varying electric field which drives the normal electrons, causing losses in the superconductors. A circuit-model representation for the two-fluid model is shown in Fig. 1. The inductor $L_s$ represents the inertia of the superconducting electrons. The inductor $L_n$ represents the inertia of the normal electrons and the resistor $R_n$ the losses associated with these electrons.

Although the new high-$T_c$ materials exhibit zero DC resistance at temperatures as high as 125 K,[3] few measurements of the RF surface resistance have been made and as yet there is no theory on which to base calculations. As discussed below, estimates of resonator Q presently must rely upon measured surface resistance which can depend strongly upon the methods of material fabrication. Thus, results now are by no means to be considered definitive. If, however, low-loss characteristics can be achieved reproducibly at frequencies between 1 - 100 GHz with the new materials, then devices of practical importance can be considered.

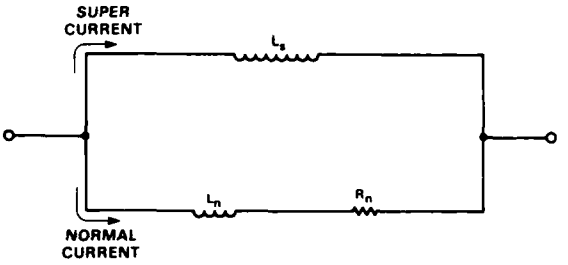



Fig. 1 Two-fluid, AC circuit model of a superconductor. $L_s$ represents the inductance of the superconducting electrons. $L_n$ and $R_n$ represent the inductance and resistance of that portion of the electrons that are in the normal state. $R_n$ results in AC losses in the superconductor.

Furthermore, the RF surface resistance is strongly temperature-dependent, decreasing with temperature. In general, a superconductor must be operated well below the transition temperature to obtain low losses. For most superconductors a generally accepted criterion is that the operating temperature must be 50 - 70% of the transition temperature before the RF losses are adequately low.

Detailed calculations of oscillator phase noise will be presented, including such factors as amplifier flicker noise and power-handling capability of the superconductor.

## II. MATERIALS

A low-phase-noise oscillator requires resonators with a very high Q, low insertion loss, and the ability to handle high input power. In terms of superconducting material parameters, one needs very low surface resistance and very high current-carrying capability. Presently, in the new high-$T_c$ materials, both surface resistance and critical current are orders of magnitude better in thin films than in bulk materials, dictating the use of thin-film to meet the low-phase-noise requirements. We thus concentrate on the properties of deposited thin films materials. In addition to low-RF-loss thin films, low-dielectric-loss substrates compatible with the thin-film-deposition process must also be employed.

The Q of a resonator depends upon the RF surface resistance of the conductor, the losses of the dielectric material, the distribution of currents in the conductors and the coupling to the external circuit. In general, we may write

$$Q^{-1} = Q_L^{-1} + Q_C^{-1} + Q_D^{-1} \qquad (1)$$

where $Q_L$ represents the loading of the resonator and can be adjusted by varying the coupling to the external

*This work supported by the Dept. of the Air Force.
**Present Address: RADC/EEAL, Hanscom AFB, MA 01731

CH2588-2/88/0000-545 $1.00 C 1988 IEEE

circuit, $Q_C$ reflects the conductor loss and, for a given resonator configuration, is inversely proportional to the RF surface resistance of the conductors, and $Q_D$ results from losses in the dielectric and is related to the loss tangent (tan $\delta$) by,

$$Q_D = 1/\tan \delta. \qquad (2)$$

The surface resistance of the new oxide superconductors has been measured only very recently by our group at Lincoln Laboratory[4] and by groups at Stanford/UCLA,[5] Cornell University,[6] and Northeastern University.[7] The Lincoln Laboratory measurements on YBaCuO show a $T_c$ of about 85 K but surface resistance several orders of magnitude larger than that of the best reported films. We are developing deposition methods to lower $R_s$. The lowest values of surface resistance have been obtained by the Stanford/UCLA group for oriented thin films deposited on $SrTiO_3$ by RF magnetron sputtering. These values are at least two orders of magnitude lower than those obtained in polycrystalline bulk samples.

The high Qs and high input power necessary for low phase noise in the oscillators translate into very high current densities for the conductors. As discussed in the next section, current densities as high as $10^6$ A/cm$^2$ may be required. The highest critical current values have also been obtained in oriented thin films deposited either on $SrTiO_3$ or on MgO substrates. Thus both surface-resistance and critical-current-density measurements indicate that the use of oriented thin films is required in practical low-noise oscillators.

As stated earlier, besides films with low RF loss, the resonator structure requires low-loss substrates. As can be seen from expressions (1) and (2), tan $\delta$ lower than $5 \times 10^{-6}$ is necessary to achieve Qs larger than 200,000. Many of the well-known low-loss materials such as sapphire react with the films during the high-temperature (800$^0$ C) annealing required after film deposition and, as a result, the $T_c$ of the film is drastically lowered, rendering the film unusable. To date, the films with highest $T_c$, highest critical currents, and lowest RF surface resistance are $YBa_2Cu_3O_7$ deposited epitaxially on $SrTiO_3$, which is unfortunately a very lossy substrate and cannot be used for resonators. Good results have also been reported on yttria-stabilized zirconia. However, we have measured the dielectric losses in these substrates and they are too high, limiting the attainable Qs of the resonators to about 3000. Sumitomo recently reported[8] that oriented films capable of very high critical currents can be deposited on MgO, which is expected to be an adequate dielectric at low temperatures, based on measured room-temperature values of the dielectric loss. Measurements of the dielectric loss at low temperature of single crystals of MgO have been performed in our laboratory and imply a tan $\delta$ of $10^{-4}$. This value is disappointingly large, but we believe it can be improved by reducing the amount of water absorbed by the hygroscopic MgO.

Thus there are material problems to be solved before a practical high-temperature superconducting resonator can be realized. The superconducting material, dielectric substrate, and method of thin-film fabrication have not yet been determined. Not only must the substrate be compatible with good superconducting material but it must also have low RF losses.

## III. RESONATOR STRUCTURES

The resonators used for high-performance oscillators can be built using either cavities or transmission-line structures. As discussed in the previous section, the requirements of high Q and high power force the use of thin films rather than bulk materials. Cavity structures of either rectangular or cylindrical geometry that use thin films would require very low-resistance joining of planar structures or deposition of films in the interior of the cavity, both of which are extremely difficult to accomplish using the new high-$T_c$ oxides. Alternatively, we could think of depositing thin films on the external surface of a dielectric; indeed, superconductive dielectric-loaded cavities have been produced using Nb deposited on sapphire.[9] For the new oxides, however, the need for oriented films makes the fabrication exceedingly difficult because of the nonplanar nature of the cavity structure.

We propose a stripline resonator structure in which energy is coupled into an isolated length of line. The superconducting stripline technology has been developed over the past few years at MIT Lincoln Laboratory for the fabrication of analog signal processing devices. The resonant frequency is determined by the length of line which is an integral number of half-wavelengths long. Stripline has a center conducting line with top and bottom ground planes, separated by two dielectric layers, as shown in Fig. 2. The stripline strucutre eliminates the radiation losses of microstrip. The stripline resonator has the advantage of planar fabrication techniques; the coupling and resonant frequency are determined photolithographically and the structure can be made compact and environmentally rugged. This stripline resonator structure has been used in our laboratory to measure RF losses in superconducting thin films of materials such as niobium, niobium nitride, and niobium tin and is now in use to evaluate the new high-$T_c$ materials. A complete theory relating the Q of the resonators to the RF surface resistance of the films has been developed and confirmed by cross reference with other surface-resistance-measurement methods. This theory is used in section IV to calculate the Qs for the resonators fabricated using the new materials.

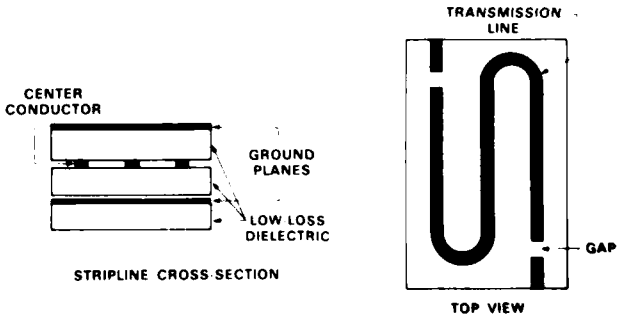



Fig. 2 Schematic view of the stripline resonator. On the left is cross-section showing dielectrics and conductors. On the right is a top view of the center section showing the center conductor which has been patterned photolithographically. The gap shown determines the coupling to the resonant section of line.

## IV. PREDICTIONS OF PHASE NOISE

Before calculating the phase noise of an oscillator stabilized with a superconducting resonator, we must first estimate the Q expected for the stripline structure and the maximum oscillator power which the resonator can support. To estimate the Q, we first determine the attenuation constant $\alpha$ for the line using a perturbation method. We assume perfect conductors and calculate the current distribution in the ground planes and central conductors using an incremental inductance method developed by Wheeler.[10] Using this current distribution and measured values of surface resistance, $R_S$, we can calculate the power loss per unit length of line and $\alpha$. Once $\alpha$ is known in nepers/m, the conductor Q is given by,[11]

$$Q_C = \pi/\alpha\lambda \qquad (3)$$

To calculate the maximum allowable power incident on the resonator, we use the equivalent circuit model shown in Fig 3. The components are derived from the Q and other parameters of the resonator.[12] From the input impedance of the resonator the voltage across the input of the stripline may be calculated for a given input power and source and load impedance. Knowing this voltage, the stripline impedance, and the resonator Q, one can then calculate the peak current in the stripline resonator. The maximum power is that which develops a peak current equal to the superconductor critical current.

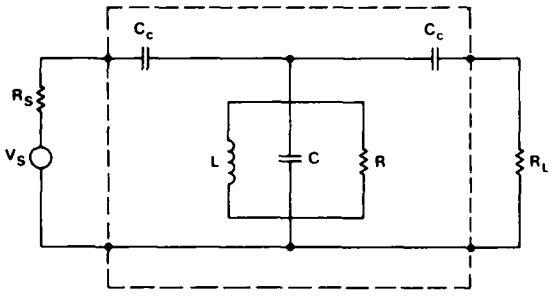



Fig. 3 Equivalent circuit of the transmission line resonator. $R_S$ and $R_L$ are source and load resistances. $C_C$ represents the coupling capacitance and is determined by the gap shown in Fig. 2.

In Figure 2 we show the cross section of the proposed stripline resonator. The two substrates are MgO, 2-mm thick. The center conductor is 2.54-mm wide. For this configuration, the $Q_C$ obtainable with a material having surface resistance $R_S$ in ohms at a frequency f

$$Q_C = 3.73 \times 10^{-9} \ f/R_S \qquad (4)$$

for f in Hz. The maximum input power P is

$$P = \frac{n\pi Z_0 I_c^2}{8 I_L (1-I_L) Q} \qquad (5)$$

where n is the mode number (the number of half wavelengths within the resonator), $Z_0$ is the impedance of the stripline, $I_C$ the maximum current for the center conductor, $I_L$ the insertion loss and Q the loaded Q. For the chosen dimensions, $Z_0$ is 28 Ω.

To calculate the Q, we assume $R_S = 4 \times 10^{-4}$ Ω at 10 GHz and 50 K. This value of $R_S$ is obtained by extrapolating the results from the Stanford/UCLA[5] measurements of oriented YBCO deposited on $SrTiO_3$, done at 100 GHz, to lower frequencies by assuming that $R_S$ is proportional to $f^2$. To calculate P, we assume a critical current density of $4 \times 10^6$ A/cm$^2$ at 77 K and extrapolate to 50 K assuming $J_C(T)$ is proportional to $(T_C-T)/T_C$. Values of critical current density of approximately this magnitude have been reported by several groups. For these values, and assuming an insertion loss of 15 dB, we obtain an unloaded Q of $32 \times 10^5$ at 5 GHz and a maximum input power of 1.5 W. The corresponding values for a center frequency of 10 GHz are $Q = 7.5 \times 10^4$ and P = 64 W.

Having calculated the Q and oscillator power which does not exceed the superconductor critical current, one can predict the values of phase noise to be expected with superconducting resonators using the standard Leesom model[12] of the phase noise of a feedback oscillator. The single sideband phase noise $L(\omega)$ is given by the following expression:

$$L(\omega) = 10 \log \left[ N^2 \left( 1 + \frac{\omega_0^2}{4Q^2\omega^2} \right) \left( \frac{GFKT}{P} + \frac{\alpha}{\omega} \right) \right] \qquad (6)$$

where $L(\omega)$ is the noise power relative to the carrier in a one-hertz bandwidth given in dBc/Hz, P is the power at the output of the amplifier, G is the loop gain, F is the amplifier noise figure, $\omega$ the offset frequency in radian/s, Q the loaded resonator quality factor, $\omega_0$ the oscillator fundamental frequency, $\alpha$ the empirically determined flicker-noise constant, k is Boltzmann's constant, T is the absolute temperature, and N the frequency-multiplication factor.

The flicker (1/f) noise can originate in either the amplifier or tne resonator or both. Since the flicker noise is not well understood theoretically, it must be empirically determined. We can make reasonable estimates for the amplifier, but for the resonator we must rely on experimentally determined values. Our projections of phase noise that follow assume that the contribution of the resonator to the flicker noise is small compared to the amplifier contribution, but the phase noise at offsets of less than about 5 kHz is dominated by the flicker noise. If the flicker noise contribution of the superconducting resonator is larger than that of the amplifier, then the actual phase noise will be greater than the following predictions. The importance of the resonator flicker noise has led us to undertake the measurements described below in Sec. V.

Figure 4 shows the projected phase noise at 10 GHz for a superconducting-resonator-stabilized oscillator for two different sets of parameters. One curve is for an oscillator operating at 5 GHz, frequency-doubled to obtain 10 GHz. The parameters are $Q = 1.5 \times 10^5$, G = 15B, F = 6 dB, P = +25 dBm (0.3 W), and $\alpha = 4 \times 10^{-12}$. This value of alpha is that normally assumed for a GaAs FET amplifier.[13] The second curve is for an oscillator operating at 10 GHz with a $Q = 7.5 \times 10^4$ and the other parameters unchanged from the first curve. The reduction in Q when raising the frequency from 5 GHz to 10 GHz comes from the increase in the surface resistance of superconductors. The

parameter values chosen for these curves are for illustrative purposes and should not be considered final. We have, for instance, assumed room-temperature operation for the amplifier, but when the amplifier is operated at the cryogenic temperatures, noise performance would be improved. Also shown in Fig. 4 are the phase-noise curves for quartz crystal[14] oscillators and for SAW oscillators.[15] Clearly the superconducting oscillator provides better performance than other technologies.

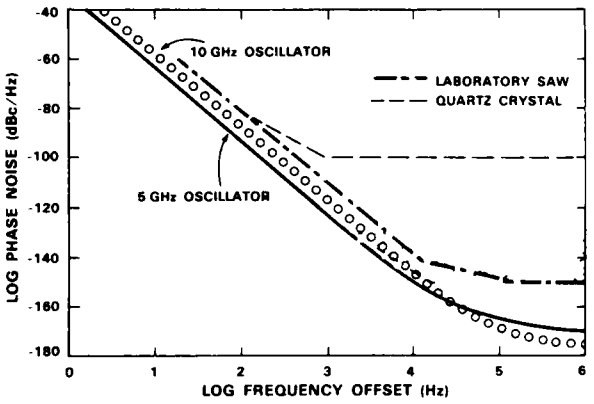



Fig. 4 Projected values of the single-sideband phase noise at 10 GHz for oscillators stabilized with superconducting resonators. Two examples are shown: one operating at a fundamental of 5 GHz and multiplied up to 10 GHz and one operating at a fundamental of 10 GHz. See text for the detailed parameters of these oscillators. Also shown are the phase noise for the best SAW oscillator and a production quartz crystal oscillator.

## V. RESIDUAL NOISE MEASUREMENTS

Because, as discussed above, the flicker noise of the superconducting resonator is unknown and because even for the standard superconducting materials there is no experimental data to draw upon, we have made measurements of the residual noise of an all-niobium superconducting resonator with a sapphire dielectric fabricated with the stripline structure described above. The devices have a loaded Q of approximately $2.5 \times 10^4$ and a fundamental resonant frequency of 1.15 GHz. The measurements, at 4 K with the resonator in a liquid-helium bath, were made in the standard way where the output of an RF source is split and one path is directed through the device under test and the other path is shifted in phase by $90^0$ to produce the reference signal. The signals are mixed, low-pass filtered and measured in a spectrum analyzer to obtain the phase noise. The schematic of the measurement system is shown in Fig. 5. In this manner the noise of the RF signal source is correlated out of the noise measurement except for offset frequencies greater than the bandwidth of the resonator.

The results are shown in Fig 6. The bandwidth of this resonator is 40 kHz, so one sees in the data the synthesizer phase noise for offset frequencies greater than 40 kHz. For offset frequencies less than 1 kHz, the increase in noise shown in the data is the result of the resonator and the liquid helium dewar probe. It is not yet clear how much of the noise is due to fundamental flicker noise in the resonator and how much is due to vibration or temperature fluctuations. Between 10 Hz and 1 kHz the curve follows closely a 1/f dependence. The noise below 10 Hz offset no longer follows the 1/f dependence so that it is probably the result of either thermal of mechanical instabilities in the resonator. This resonator design was optimized for material evaluation and not for oscillator application, so it is not surprising that it is vibration sensitive. We are investigating designs of resonator which will reduce the vibration sensitivity.

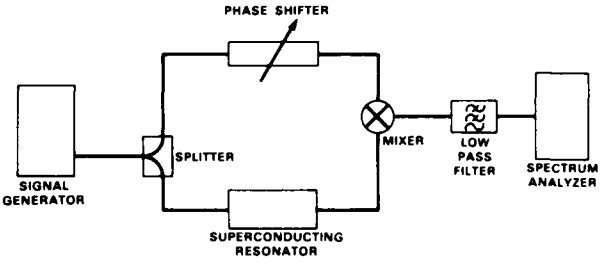



Fig. 5 Block diagram of the residual noise measurement system.

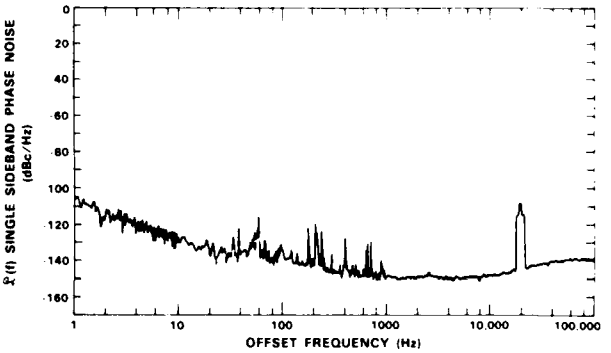



Fig. 6 Measured residual noise of a Nb stripline resonator at 4 K.

The 1/f behavior shown between 10 Hz and 1 kHz would imply a resonator flicker-noise constant, alpha, of approximately $4 \times 10^{-12}$ which is the same value assumed in the calculations of phase noise in sec IV for the amplifier contribution, a value consistent with GaAs amplifiers. This flicker-noise contribution would not seriously degrade the phase noise projected in section IV. These results are preliminary and nothing definitive can be said about the noise properties of the high-$T_c$ materials at this time.

## VI. CONCLUSIONS

The new high-transition-temperature superconductors hold great promise for low-phase-noise oscillators in the frequency range from 1 - 10 GHz. Doppler radar and communication systems are the most important applications of the technology. We have shown projections of the phase noise that are more than an order of magnitude better than other technologies. This is because superconducting resonators possess two attributes which lead directly to lower phase noise, namely low loss and high current-carrying capacity. The superconducting resonators can be implemented in a compact and rugged form with all critical dimensions determined by photolithography.

Material-related problems must, however, be solved before these devices become a reality. Better superconducting material processing must be developed and compatible low-loss substrate materials must be found.

We have begun flicker-noise measurements on superconducting Nb resonators and found values which are consistent with the phase-noise projections.

## VII. ACKNOWLEDGEMENTS

We would like to thank P. Murphy for help with device fabrication, G. Fitch for programming, and K. Poe, G. O'Malley and W. Walker for help with the measurements. We would also like to thank R. W. Ralston and R. S. Withers for support and encouragement.

## REFERENCES

1. See for example, High-Temperature Superconductors II: Extended Abstracts, D. W. Capone II, et al. editors Pittsburgh, PA: 1988 (Materials Research Society Pittsburgh PA).

2. A. C. Anderson, unpublished results.

3. A. Khurana, "The $T_c$ to Beat is 125 K," Physics Today, vol. 41, pp. 21-25 (1988).

4. M. S. DiIorio, A. C. Anderson and B. Y. Tsaur, "RF Surface Resistance of YBaCuO Thin Films," Phys. Rev. to be published.

5. M. Beasley, private communication.

6. D. L. Moffat, K. Green, J. Gruschus, J. Kirchgessner, H. Padamsee, D. L. Rubin, J. Sears, Q-S. Shu, T-W. Noh, R. Buhrman, S. Russek and D. Lathrop, "Temperature Dependence of the Surface Resistivity at 1.5 GHz and 6 GHz of Single Crystals, Thin Films and Sintered Pellets of $YBa_2Cu_3O_{7-x}$," in High-Temperature Superconductors II: Extended Abstracts, D. W. Capone II, et al. editors Pittsburgh, PA: 1988 (Materials Research Society Pittsburgh PA), pp. 347-350.

7. S. Sridhar and W. L. Kennedy, "A Novel Technique to Measure the Microwave Response of High $T_c$ Superconductors Between 4.2 K and 200 K," submitted for publication in Rev. Sci. Inst.

8. H. Itozaki, S Tanaka, K. Higaki and S. Yazu, "High Tc and $J_c$ Superconducting This Films Prepared by Sputtering," High-Temperature Superconductors II: Extended Abstracts, D. W. Capone II, et al. editors Pittsburgh, PA: 1988 (Materials Research Society Pittsburgh PA), pp. 205-208.

9. V. B. Braginsky, V. P. Mitofanov and V. I. Panov, Systems with Small Dissipation, (translated by E. Gliner) Chicago: University of Chicago Press, p. 69, 1985.

10. H. A. Wheeler, "Formulas for the Skin Effect," Proc. IRE, vol. 30, pp. 412-424, 1942.

11. G. L. Mattaei, L. Young and E. M. T. Jones, Microwave Filters, Impedance-Matching Networks, and Coupling Structures, New York: McGraw-Hill, 1964, p. 215.

12. D. B. Leeson, "Short Term Stable Microwave Sources," Microwave Journal vol. 13, pp. 59-69, June 1970.

13. T. Parker, "Characteristics and Sources of Phase Noise in Stable Oscillators," in Proc. of the 41st Ann. Freq. Control Symp. 1987, pp. 99-110.

14. E. A. Gerber, T. Lukaszek and A. Ballato, "Advances in microwave acoustic frequency sources," IEEE Trans. Microwave Theory Tech., vol. MTT-34, pp. 1002-1016, 1986.

15. G. K. Montress, T. E. Parker and M. J. Loboda, "Extremely Low Phase Noise SAW Resonator Oscillator Design and Performance," in Proc. of the 1987 Freq. Control Symp., pp. 47-52.

42nd Annual Frequency Control Symposium - 1988

# HIGH TEMPERATURE SUPERCONDUCTORS AND THEIR PROMISE FOR BETTER MICROWAVE CIRCUITS

Paul H. Carr and John S. Derov

Electromagnetics Directorate, Rome Air Development Center

Hanscom Air Force Base, Massachusetts 01731-5000

## ABSTRACT

The startling discovery of superconductors with transition temperatures about 90 K, which is above the 77 K temperature of liquid nitrogen, has created great interest. Superconductors can produce microwave circuits with lower conductor loss, higher-Q, lower noise, and wider bandwidths. The paper will first give a tutorial on the microwave resistance of superconductors and the status of high-temperature-superconductor materials technology in achieving low resistances.

The resistance R of superconductors at microwave frequencies is unfortunately not zero, as it is for d.c., but is equal to R = R(T) + R(res), where R(res) is a residual resistance due to such mechanisms as impurity scattering and R(T), is due to the thermal excitation of normal (as opposed to paired) electrons. For a well-developed superconductor like niobium, R(T) at 4.2 K is about three-orders of magnitude lower than that of room temperature copper at 10 GHz and about two-orders of magnitude lower at 100 GHz. Initial microwave measurements of ceramic Yttrium Barium Copper Oxide yielded R(res) values ranging from 0.1 to that of copper. Lower resistance values observed for thin films and single crystals will be reported.

Applications of the new superconductors may come at higher temperatures than those based on present metal superconductors. The latter include wideband signal processors and very low noise mixers. New hybrid applications may also arise at 20-70K where solid state amplifiers also have lower noise and higher gain. High temperature superconductors have the potential of important applications in frequency control and signal processing.

## I. INTRODUCTION

The 70 degree increase in the critical temperature of superconductors in early 1987 sparked renewed interest, because the 77-K-temperature of liquid nitrogen was now exceeded. This temperature increase has made it possible to consider superconductivity an economically viable technology. Since a considerable amount of device technology has already been developed at liquid helium temperature (4.2 K), this technology should be transferable to the new higher temperature superconductors. Low temperature superconductors have shown that microwave circuits can be produced with lower conductor loss, higher-Q, lower noise, and wider bandwidths. The paper will first give some insight to the microwave resistance and the status of the rf. loss measurements on high-temperature-superconductors and, second, discuss application for the new high temperature superconductors.

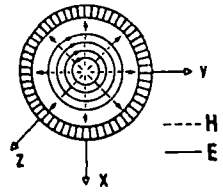
## II. MICROWAVE MEASUREMENTS

In the dc case, the superconducting surface resistance is zero due to the exclusion of the electric field. The zero value for the surface resistance is in agreement with ohm's law, $J = \sigma E$, where J is the current density, $\sigma$ is the conductivity and E is the electric field. Because the conductivity of the superconductor is infinite, the dc electric field must be zero to have a finite current density.

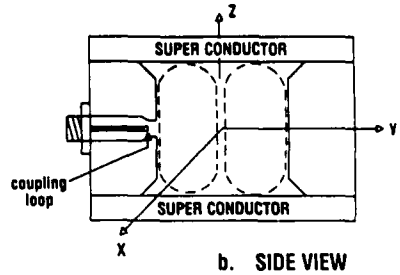
The ac surface resistance of a superconductor is non zero. London(1) used a two fluid model to describe ac surface resistance and this model yields reasonable first order results. It uses two terms to describe the current density J. Thus, $J = J_n + J_s$ where $J_n$ is the current density for the normal conduction electrons (which have resistance) and $J_s$ is the current density of the superconducting super-superfluid. The relationship for the surface resistance for a homogenous superconductor can be derived from the two fluid model and the theory of Bardeen, Cooper and Schrieffer (BCS)(2,3)

$$R = \text{const} \times f^2 \times \exp(-2\Delta/kT) + R(\text{res}) \qquad (1)$$

where $\exp(-2\Delta/kT)$ is the temperature dependent term used in the BCS theory, $\Delta$ is the superconducting energy gap and R(res) is the residual resistance. The residual resistance in a superconductor may be explained by trapped flux or impurities in the materials.




a. TOP VIEW




b. SIDE VIEW

Fig. 1 Cylindrical Waveguide Cavity Resonator (a.) shows the Electric (E) and Magnetic (H) field pattern for the $TE_{011}$ mode (b.) shows the Y-Z cross section of the cylindrical cavity.

Waveguide cavity type resonators and thin film stripline resonators are most often used for making rf measurements of the surface resistance. We will confine our discussion to cylindrical waveguide cavity resonators. These cylindrical cavities can have many transverse electric (TE) and transverse magnetic (TM) modes. A $TE_{011}$ mode cavity is typically used for small signal characterization. There are two reasons for choosing a $TE_{011}$ mode cavity: first, this mode has the lowest loss characteristic; second, due to symmetry of the electric and magnetic fields as shown in figure 1. it is least effected by discontinuties between the end walls and side walls of the cavity.

US GOVERNMENT WORK IS NOT PROTECTED BY US COPYRIGHT

The disadvantage of this mode is that it is degenerate with the $TM_{111}$ mode. The degeneracy can be handled in two ways: by altering the shape of the cavity so that it affects only the $TM_{111}$ mode and not the $TE_{011}$ mode (this can be done by cutting a groove in the cavity) and by the method used to excite the cavity. A groove or grooves can be cut into the cavity so that the length of the cavity looks longer to the TM mode than to the TE mode. The excitation method is to use a coupling loop. Figure 1. shows the cavity configuration being used at RADC to make the surface resistance measurements on the bulk ceramics and thin films.

Using the waveguide resonant cavity allows one to easily predict the quality factor (Q) of the cavity from classical electromagnetics. The cavity Q is defined as Q = THE ENERGY STORED/POWER DISSIPATED PER CYCLE. From this definition of the cavity Q, one can show that

$$1/Q \propto R_s \qquad (2)$$

where $R_s$ is the surface resistance.

Since the 90 K breakthrough in 1987, microwave and millimeter wave measurements have been made on the bulk ceramic, thin films and single crystals of the yttrium barium copper oxide (YBaCuO) superconducting compounds. We shall discuss measurements made by S. Sridhar and W.L. Kennedy (4), Derov et.al (5) and H. Padamsee et. al. on bulk ceramic and by Padamsee et.al on bulk ceramic and by Padamsee et.al. (6,7) on thin films and single crystals. The surface resistance for the bulk ceramic superconductor have been measured to be equal to or about 1/2 the surface resistance of copper. The surface resistance measurements made on thin film and single crystal samples by Padamsee et.al. were done using a niobium cavity resonator. The results of the surface resistance measurements are shown in Table 1 and for single crystals the surface resistance values range from 1/2 to 0.005 the surface resistance of copper.

CORNELL UNIVERSITY DATA AT 6 GHz

(PADAMSEE ET.AL.)

| SPECIMEN | RF $T_c$ (K) | $R_s$ (77 K) MILLIOHM | $R_s$ (4.2K) MILLIOHM | $R_s$ (4.2K) / R (Cu) |
|---|---|---|---|---|
| COPPER | - | 5.1 | 3.8 | - |
| PELLET | 96 | 10.5 | 4 | 1.0 |
| FILM #2/$ZrO_2$ | 90 | 171 | 20 | 5.3 |
| FILM #3/MgO | 77-87 | 200 | 7 | 1.8 |
| CRYSTALS-1 | 85 | 100 | <0.02* | 0.005 |
| CRYSTALS-2 | 85 | 150 | 2 | .53 |

*SEPARATE MEASUREMENT IN AN 8.6 GHz Nb CAVITY AT 1.5K

Table 1 is a list of YBaCuO samples measured at Cornell University. The surface resistance of the samples are given and the samples are compared with copper.

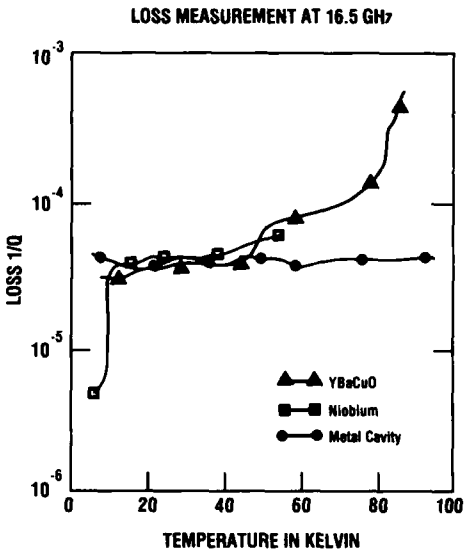



Fig. 2 The Loss Measurement of the $TE_{011}$ Brass cavity, YBaCuO bulk ceramic, and Niobium Thin Film.

Measurements have been made by Derov et.al at RADC using a 16.5 GHz cavity (as shown Figure 1) on bulk ceramic samples and on niobium thin films. A plot of the results is shown in Figure 2. The inverse of the unloaded Q is plotted as a function of temperature. The plot consists of the brass $TE_{011}$ cavity alone, a YBaCuO ceramic sample and niobium thin films. The results show the loss of the ceramic superconductor to be the same order of magnitude as the metal cavity itself.

Sridhar and Kennedy's measuements were made in a 10 GHz lead coated superconducting cavity as shown in Figure 3. Sridhar and Kennedy made the measurements on lanthanum strontium copper oxide (LSCO) and yttrium barium copper oxide (YBCO) superconductor. Their results are shown in Figure 4 as plot of inverse unloaded Q as function of temperature T. This plot shows the temperature dependence of the resistivity of the LSCO and YBCO ceramic, which have critical temperatures of 40K and 90K, respectively. The base line loss for a lead superconducting $TE_{011}$ cavity is also plotted. These results indicate that the YBCO ceramic has a surface resistance approximately 0.1 that of copper.

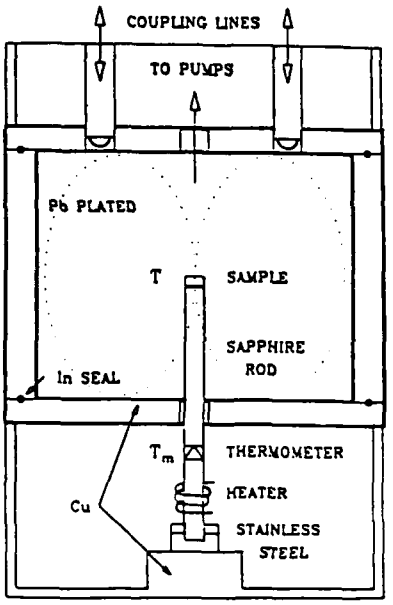



Fig. 3 Schematic of the superconducting cavities used at Northeastern University (Pb) (ref 4) and Cornell University (Nb)

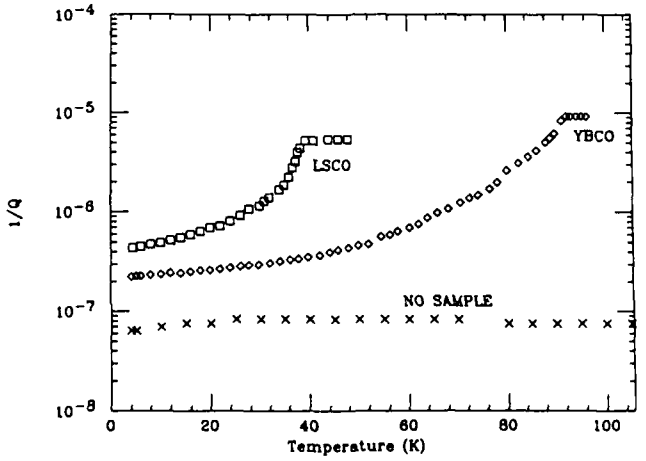



Fig. 4 1/Q Loss Measurement at 10 GHz as a function of temperature on YBCO and LSCO bulk ceramics. (ref 4)

A niobium cavity similar to the one shown in Figure 3 was used in the work done by Padamsee et. al at Cornell University. In Figure 5, the result for a single crystal YBaCuO superconductor is shown and the inverse of unloaded Q is plotted as a function of inverse T of the crystal. The temperature dependence of the resistivity of the crystal appears to be exponential between 20K to 2K. The data was then fit using a five parameter function that defines the residual Q and two distinct energy gaps. The two values used in the exponentiual were $2\Delta/kT_c$ = 4.0 and 0.5, respectively. The first is close to the BCS value of 3.5 and dominates below the transition temperature, but at about 20K the crystal resistivity takes on the second exponential value of 0.5. This result is plotted in Figure 6.

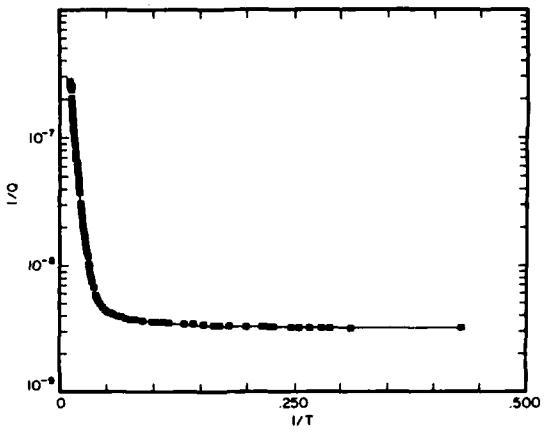



Fig. 5 1/Q Loss Measurement as a function of temperature for a YBaCuO single crystal. (ref 7)

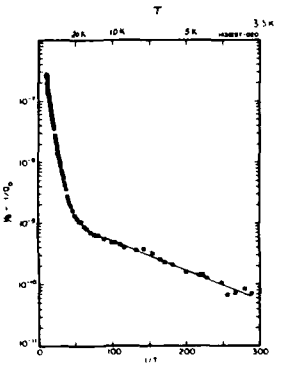


Fig. 6 Double Expotential fit of the unloaded A for a YBaCuO single crystal as a function of temperature. (ref 7)

A plot of surface resistance as a function of frequency is shown in Figure 7. This plot shows a theoretical prediction for room temperature copper, and gold at 50K. It also plots a theoretical prediction using the two fluid model of superconductivity for YBaCuO. The plot for niobium (Nb) is experimental. Points have been added to show where the actual measured values fall. The thin film measurement point at 100 GHz was done by J.P. Cavini et.al. (8).

## RF SURFACE RESISTANCE OF SUPERCONDUCTORS

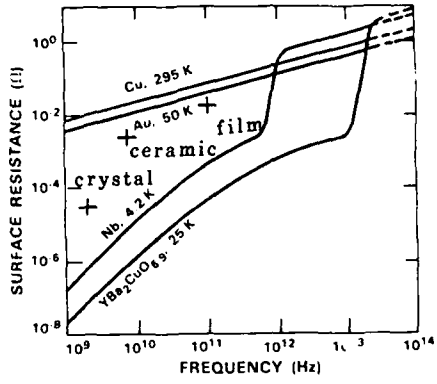



Fig. 7 Plot of surface resistance in ohms as a function of frequency for copper, gold, niobium with measured points or added for bulk ceramic thin film and single crystal YBaCuO. The smooth YBaCuO curve is a theoretical plot using the two fluid model. (ref 9)

Most of the results for the surface resistance of the new high temperature superconductors are near or equal to that of a good conductor. However, new high temperature superconductors are anisotropic in their superconducting phase. This anisotropy makes the high temperature superconductors much more sensitive to defects and orientation in their crystal structure. *Thus the best results for YBaCuO should come for* oriented thin films and single crystals and the rf surface resistance should improve as more is learned about the thin film processing and crystal growth of anisotropic superconductors.

### III. MICROWAVE APPLICATIONS

We shall first describe applications of present liquid-helium-refrigerated (4.2K) superconductors. These should be more pervasive with liquid-nitrogen-refrigeration (77K), which has a 3000-times cost advantage. In addition, new applications inherent in the unique properties of the high temperature superconductors may emerge.

#### Lower Loss and Noise

Microwave circuits using niobium superconductors refrigerated with liquid helium are being used to achieve low conductor loss. These include niobium stripline circuits and cavities. Stripline circuits deposited on silicon substrates are used as gigahertz

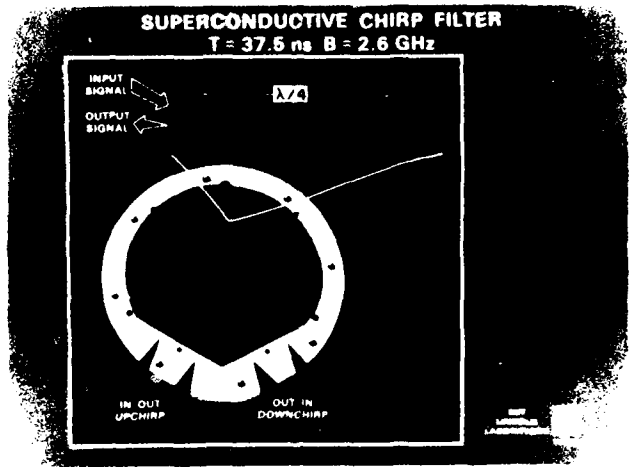



Fig. 8 Photograph of spiral coupled microstrip delay line on 5-cm-diameter wafer. (ref 9)

bandwidth pulse expanders and compressors. Fig 8 illustrates a spiral coupled microstrip circuit capable of generating a dispersive "chirped" pulse having a differential time delay of 37.5 nsec over a 2.6 GHz band centered at 4 GHz (9). The 3-m-long delay line, which is approximately 1-mm-wide, is folded to fit on a 5-cm-diameter wafer, 125 micrometers thick. This circuit packing density, which has an insertion loss of about 10 dB, is much higher than that of room-temperature conductors.

### SUPERCONDUCTIVE TAPPED DELAY LINES COMPRESSED PULSE RESPONSE

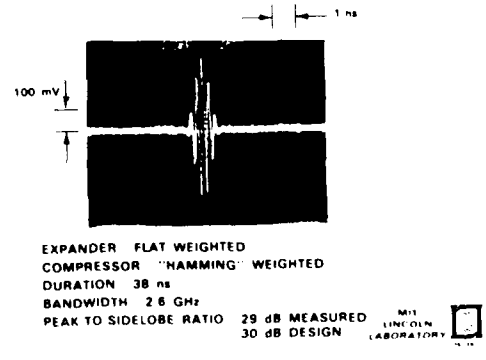



Fig. 9 Oscilloscope trace of compressed pulse after being expanded in flat weighted delay line shown in Fig. 8. (ref 9)

Fig. 9 illustrates the compressed pulse after being expanded in an unweighted or "flat" delay line. The compressing delay line is "Hamming Weighted". The measured peak-to-sidelobe-ratio of 29 dB is close to the 30 dB design ratio. The 2.6 GHz bandwidth achieved with this compact, low power analog technology would be impossible to achieve with comparably sized digital technology. Wideband pulse compressors are used in high resolution radar and also in fast, real-time spectrum analyzers.

High-Q cavities are important for low-noise, stable oscillators in frequency control applications. Since these are discussed in other papers, they will not be covered further here. However we will mention that superconducting cavities are important for accelerating the charged particles in high energy

particle machines and "super-colliders". Whereas superconducting magnets provide the high magnetic fields to guide protons or electrons in circular orbits, rf cavities provide the high accelerating voltages needed for electron rings, where electrons suffer severe energy losses. For this application, the cavities must of course be able to withstand the rf high fields needed to accelerate the particles. The niobium cavities currently used for this application may someday be replaced by the high temperature superconductors, which have large energy gaps and high critical fields. (6)

High temperature superconductors have the potential to substantially improve the performance of monolithic microwave integrated circuits. The lower circuit loss leads to higher Q lumped circuit elements and transmission lines. Lower circuit loss requires less amplifier gain. Most important is the fact that solid state transistors work better at temperatures in the vicinity of 77 K and below. Semiconductor mobilities increase, resistive losses decrease, resulting in transistors with higher gain and lower noise. Aging effects are slower at low temperatures, resulting in high reliability. The ultimate result will be monolithic microwave circuits with higher performance and lower cost.

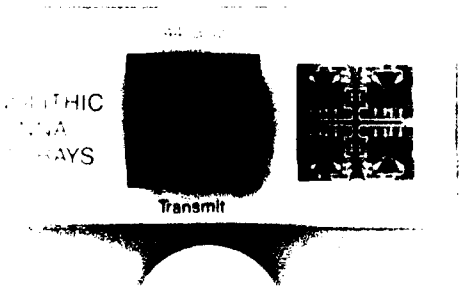



Fig. 10 Photograph of monolithic antenna subarrays on roughly 1.4 cm square GaAs chips. (ref 10)

Figures 10 and 11 illustrate the positive impact of high temperature superconductors on 20-60 GHz monolithic arrays (10). Figure 10 is a photograph of roughly 1.4 cm. square GaAs chips for 44 GHz transmit and a 20 GHz receive phased array antenna. The 44 GHz chip contains rectangular-shaped patch antennas monolithically integrated with GaAs field effect transistor amplifiers and phase shifters. The monolithic chip also has a considerable length of microstrip transmission lines, whose loss could be substantially reduced with the use of a superconductor. Transmission loss is particularly important for large phased array antennas (i.e. thousands of radiating elements).

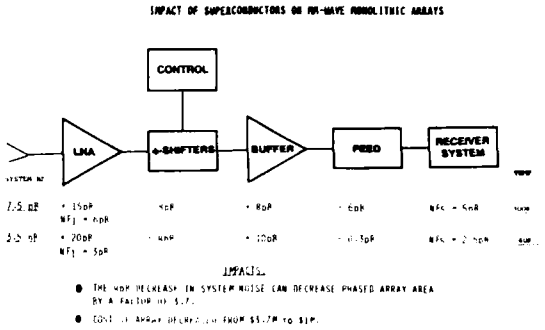



Fig. 11 Impact of high temperature superconductors on decreasing the noise of monolithic receiving arrays.

The block diagram shown in Figure 11 is for each element of a receiving phased array. It consists of a monolithic low noise amplifier (LNA), a phase shifter, a buffer amplifier, a transmission line distribution or feed network, and a receiving/processing system. Under each block is listed the gain or loss estimated for each element at 300 K and at 40 K. The system noise figure at 300 K of 7.5 dB is determined primarily by the 6 dB noise figure of the LNA. The amplifier gains increase at 40 K, the circuit losses decrease; the most marked decrease is in the superconducting feed network. Again the overall, system noise figure at 40 K is dominated by the LNA. The overall system noise decreases by 4 dB at 40 K. For the same level of detectability, this noise decrease can be used to decrease the area of the phased array. Since the number of circuit elements is proportional to the area, the cost of the elements and the phased array itself can decrease from and estimated $3.7M to $1M.

Substantial advances in fabricating the high-temperature superconductors will be required to make their processing compatible with monolithic microwave integrated circuits. At present, a final step in processing Yttrium Barium Copper Oxide thin films is an anneal in oxygen at about 900C. This high temperature will destroy circuits already fabricated on silicon or gallium arsenide. Lower processing temperatures need to be developed. Recently, Adachi, et. al. (11), have developed a technique for sputtering Er-Ba-Cu-O films directly at 650C. The as deposited films on MgO exhibited a sharp superconductive transition with zero resistance at 86 K.

Many mm-wave receivers use a low-noise mixer instead of the low-noise amplifier shown in Fig. 11. Figure 12 shows that the lowest noise mm-wave mixers available today are made with superconductor-insulator-superconductor (SIS) diodes. Noise figures of room temperature Schottky diodes are about 5 dB up to 100 GHz. The 4-dB-lower noise increases the range of communications and radar systems.

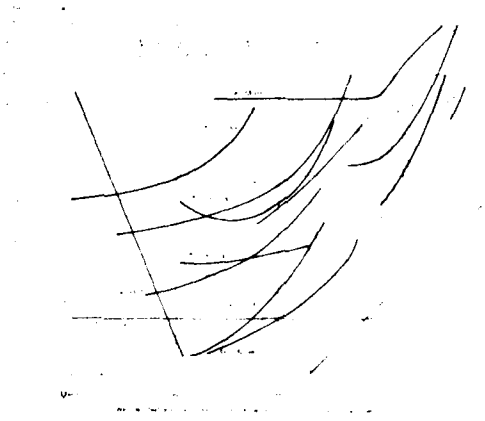


Fig. 12 Noise Figure (F) and Noise Temperature for SIS (Superconductor-Insulator-Superconductor) mixers and other devices. (By S. Weinreb, National Radio Astronomy Observatory).

High Speed and Wider Bandwidth

The higher speeds and wider bandwidths achievable with superconductors have important implications for instruments, computers, and systems. An important example is the picosecond signal processor recently marketed by Hypres. This instrument can operate as a digital sampling oscilloscope having an internal rise time of 5 psec, a sensitivity of 50 microvolts, and a bandwidth of 70 GHz. Only the Josephson junctions on the "front end" or input need to be cooled. This is done by efficently spraying the niobium circuit chip with liquid helium. This instrument for microwave/mm-wave circuit measurements has the shortest rise time and bandwidth. (13)

Present supercomputers, such as the Cray, are capable of about a billion operations per second and require a substantial cooling system. High temperature superconductors have the potential of increasing this to a trillion operations per second. Such speeds are already achievable with specialized analog processors. (13)

## CONCLUSION

Superconducting technology should make possible radar and communication systems operating in the 100-1000 GHz range. These systems are uniquely suited for applications in space. RADC had programs underway to develop superconducting niobium nitride before the discovery of high temperature superconductors, which will make these systems easier to refrigerate.

Substantial advances in materials processing and fabrication of high temperature superconductors are required for actualization of their potential. These advances should occur due to the great interest and activity.

The 90K transition temperature of Y-Ba-Cu-O was first announced in March 1987. Recently Bi-Ca-Sr-Cu-O and Tl-Ca-Ba-CU-O have been discovered to have transitions temperatures as high as 110 and 125 K respectively.

## REFERENCES


1. J.M. Blatt, (1964) THEORY OF SUPERCONDUCTIVITY, Academic Press, New York, pp 59-62

2. Van Duzer, and C.W. Turner, (1981) PRINCIPLES OF SUPERCONDUCTIVE DEVICES AND CIRCUITS, Elsevier, New York, pp 53-68, 131

3. A. Philipp, and J. Halbritter, (1981) ON THE TEMPERATURE AND FREQUENCY DEPENDENCE OF THE SURFACE RESISTANCE OF SUPERCONDUCTING Nb, IEEE Trans. on Magnets, Mag-17, (No. 1), pp 951-954

4. S. Sridhar, and W.L. Kennedy, (1988) A NOVEL TECHNIQUE TO MEASURE THE MICROWAVE RESPONSE OF HIGH Tc SUPERCONDUCTORS BETWEEN 4.2K and 200K, Review of Scientific Instruments, Vol. 59, p 513

5. J.S. Derov, W.D. Cowan, and C. VonBenken, Measurements made at RADC/EEAC, Hanscom AFB, MA to be published.

6. H. Padamsee, K. Green, J. Gruschas, et al, Microwave Superconductivity for Particle Accelerators - HOW THE HIGH TC SUPERCONDUCTORS MEASURE UP. Conference on Superconductivity Applications, Institute on Superconductivity, Buffalo, NY 14260, (April 18-20) 1988

7. D.L. Rubin, K. Green, J. Gruschus, J. Kirchgessmen, D. Moffat, H. Padamsee, J. Senn, Q.S. Shu, RF CHARACTERIZATION OF SINGLE CRYSTAL SAMPLES OF YaBa-Cu-O. CLNS 87/115 private communication, (April 1988)

8. J.P. Carini, A. Awasthi, W. Beyermann, G. Crunner, T. Hylton, K. Char, and M.R. Beasley, M.R., MILLIMETER WAVE SURFACE RESISTANCE MEASUREMENTS IN HIGHLY ORIENTED $YBa_2 Cu_3 O_{y-8}$ THIN FILM, 1988, Phys Rev B Vol. 37, (No.16), pp 9726-9729

9. R.S. Withers, A.C. Anderson, J.B. Green, and S.A. Reible, (1985) IEEE Trans on Magnetics, vol. MAG-21, pp 186-192

10. John F. McIlvenna, (March 1988) Microwave Journal, Vol. 31, pp 113-125

11. H. Adachi, K. Hirochi, K. Setsune, M. Kitabatake, and K. Wasa, (Dec 1987) Appl. Phys. Lett. Vol. 51, pp 2263-2265

12. S.K. Pan, M.J. Feldman, A.R. Kerr, P. Timbie, (1983) Applied Physics Letters, Vol 43, pg. 786

13. P.H. Carr, (Dec 1987), Microwave Journal, Vol 30, pp 91-94

# EVALUATING SUPERCONDUCTING RESONATOR MATERIALS

B. R. McAvoy, J. D. Adam, and G. R. Wagner
Westinghouse R&D Center
Pittsburgh, PA 15235

## Summary

Reliable techniques for the microwave evaluation of superconductors are essential to continued materials development. Calibrated data together with the confidence of repeatability are required for exchange between laboratories and for developing useful design data. In an effort to meet these requirements we have undertaken an examination of techniques using microwave stripline resonators. These resonators provide for the rapid measurement of microwave parameters with minimal constraints on processing. Sandwiched microstrip line resonators are used to compare the performance at 4.2°K of oxygen-free copper, lead films, niobium films and ceramic high $T_c$ films, measurements at C-band (2.8 GHz) and X-band (7.6 GHz) are described. typical results for a niobium film mic.ostrip resonator show a loaded Q value of about 80,000 with a transmission insertion loss of 5 dB at 3 GHz.

In addition high Q microwave resonators have important application in low phase noise reference oscillator. Preliminary results on the phase noise performance of a Nb stripline resonator at 2.8 GHz are presented.

## Stripline Design

The various approaches to planar microwave line geometry are depicted in Figure 1. In this figure A and B show microstrip configurations and C and D, with a ground plane at top and bottom, are referred to as stripline. Slot line and coplanar line are shown in E and F, respectively. Configurations A and B may be put aside immediately because of high radiation losses and thus lower attainable Q than that for options C and D. In the suspended substrate configuration (D) an advantage is that higher loss substrates such as $SrTiO_3$ and MgO may be used to facilitate epitaxial film growth of $YBa_2Cu_3O_7$ for the strip conductor without causing substantial dielectric loss. A potential difficulty with this approach is the means of suspension to provide for reproducible mechanical positioning and for vibration immunity. If a very low loss dielectric such as sapphire were employed to fill the void top and bottom then the option is essentially that of configuration C. The thin (suspended) layer might then be a buffer layer on the sapphire to support ordered growth of the high $T_c$ layer. The best option then, appears to be configuration C which we have implemented as shown in figure 2. A brass mounting block, shown to the

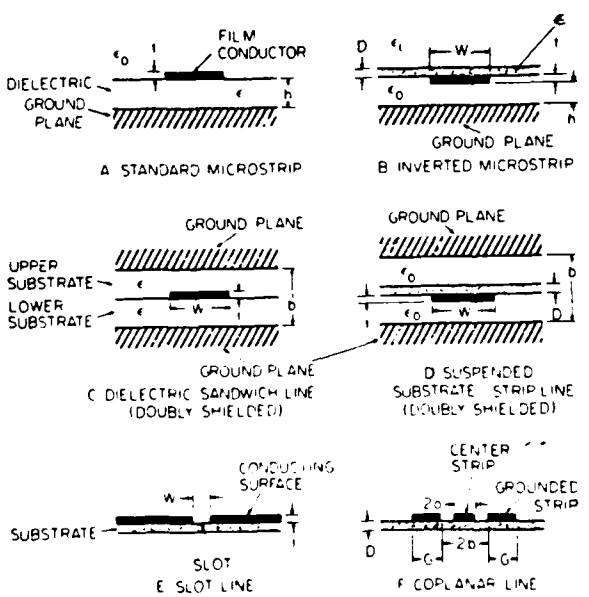



Figure 1. Planar microwave line configurations.

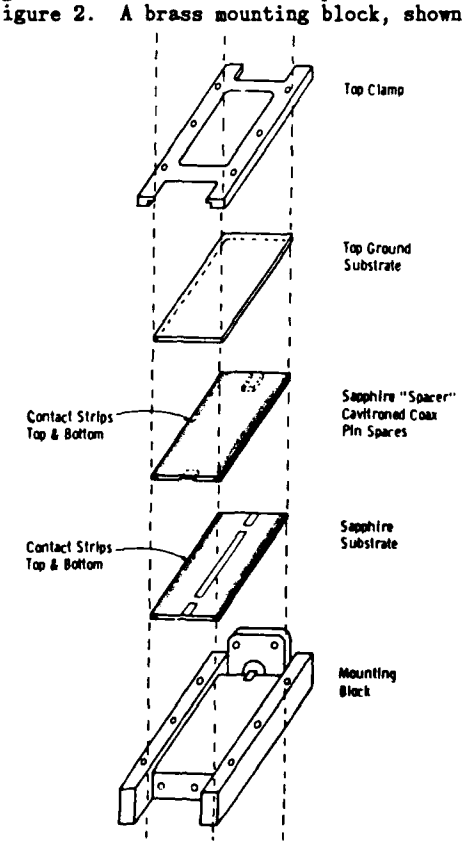



Figure 2. Dielectric sandwich line (stripline) showing placement of sapphire spacer and top ground substrate.

CH2588-2/88/0000-556 $1.00 C 1988 IEEE

right in Figure 3, receives a sapphire substrate which has a superconducting ground plane and a superconducting stripline resonator (half wavelength) pattern on the top surface as shown in Figure 2.[1] The spacer or upper substrate has small portions of each end removed by ultrasonic machining to provide space for the flat pins of the 3 mm coaxial connectors. Figure 4 shows this feature in detail together with evaporated gold contact strips along the sides to facilitate ground current return and subsequent reduction of radiation losses. The goal of such a design is the achievement of consistent results with continual assembly and disassembly of identical parts.

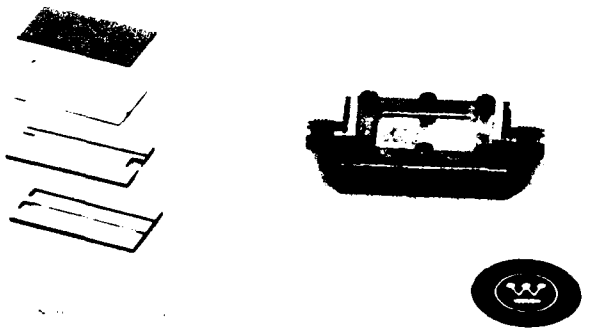


Figure 3. At left bottom is a niobium stripline resonator with a spacer sapphire substrate above it. At left top is a polished copper top ground and a lead plated copper plate. Assembled resonator is to the right.

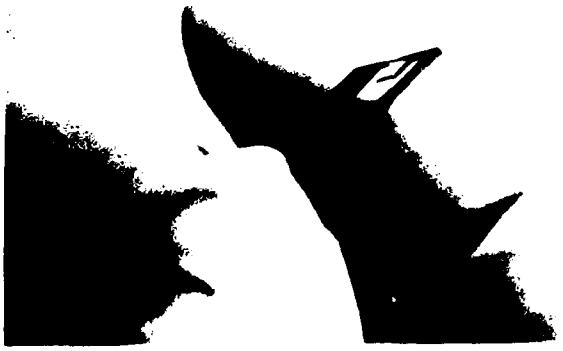


Figure 4. End of sapphire spacer substrate showing accommodation for rf connecting pins.

## Experimental Results

The dewar probe currently used consists of mechanically stiffened 3 mm, flexible coaxial cable (Sucoflex 104) providing input and output lines for transmission measurements. This assembly will fit into a 1.25 inch diameter opening in a cryostat. Figure 5 shows a typical result obtained on a network analyzer. As mentioned earlier transmission measurements are used, affording an expedient evaluation of loaded Q from the frequency width at the half power points. This together with an accurate measure of the insertion loss at resonance provides the unloaded Q value.[2] The state of the coupling to the resonator is provided by the reflection measurements ($S_{11}$ and $S_{22}$). the input power level is monitored so that, for niobium, critical rf current density is not exceeded with attendant Q spoiling.

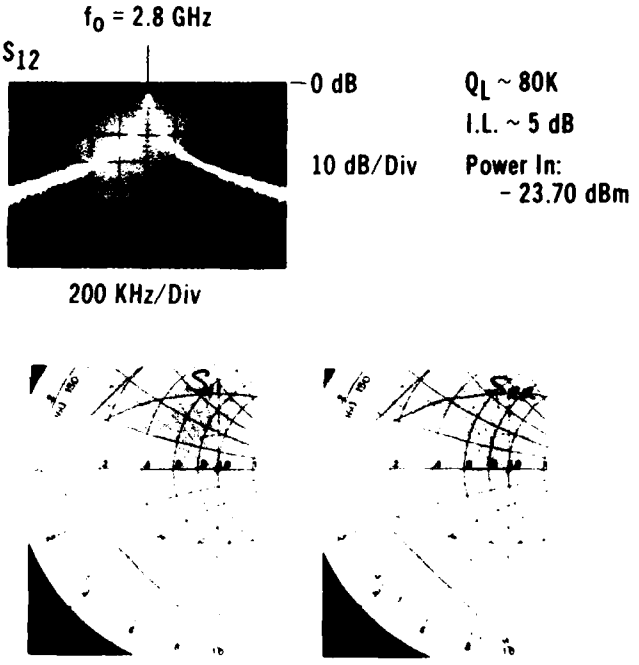



Figure 5. Network analyzer results for a Nb stripline resonator film on sapphire at 4.2°K.

Figure 6 summarizes the result of a series of measurements. In block No. 1 oxygen free copper which has been highly polished is used as the top ground plane (shown at the top left in Figure 3). Three runs at the niobium stripline resonant frequency of 2.8 GHz are listed. The data for the four runs in block No. 2 characterize a lead coated polished copper top ground plane with a niobium stripline resonator at 2.8 GHz. In each case, as shown, the values for loaded Q ($Q_L$ x 1000) and unloaded Q ($Q_U$ x 1000) are fairly consistent. As mentioned in the introduction each run represents an essentially different resonator, each having been reassembled before the run. Block No. 3 shows the results for a niobium resonator with a niobium film on sapphire forming the top ground plane. The importance of returning the top ground plane current, i.e., sealing, is evident in the nearly factor of two difference in measured Qs. Block No. 4 shows a similar resonator but at higher frequency and using lead. The lead coatings were prepared by

thermal evaporation and the niobium films by sputter deposition in a UHV system. Block No. 5 suggests our current approach to the microwave evaluation of high $T_c$ films. The film on $SrTiO_3$ is clamped against a sapphire spacer in similar fashion to the method No. 1 and No. 2. The high $T_c$ material forms the top ground plane for a Nb or other stripline resonator. We believe that work done to this point on will yield confidence in the results obtained on high $T_c$ materials.

## Resonator Noise

High Q superconducting resonators have potential for application in low phase noise microwave reference sources. Preliminary results have been obtained for noise performance of a Nb resonator of the type shown in Figure 6, block No. 3 suspended on a flexible probe as described in the previous section. The result is shown in Figure 7. With a loaded Q of about 50,000 and insertion loss of 11 dB the self noise was measured using standard methods and HP test equipment.[3] The excessive noise near the carrier is currently attributed to mechanical vibrations and helium bubbling but this could be due to resonator self noise as in quartz piezoelectric resonators and high overtone bulk acoustic resonators.[4] This will be confirmed in future measurements.

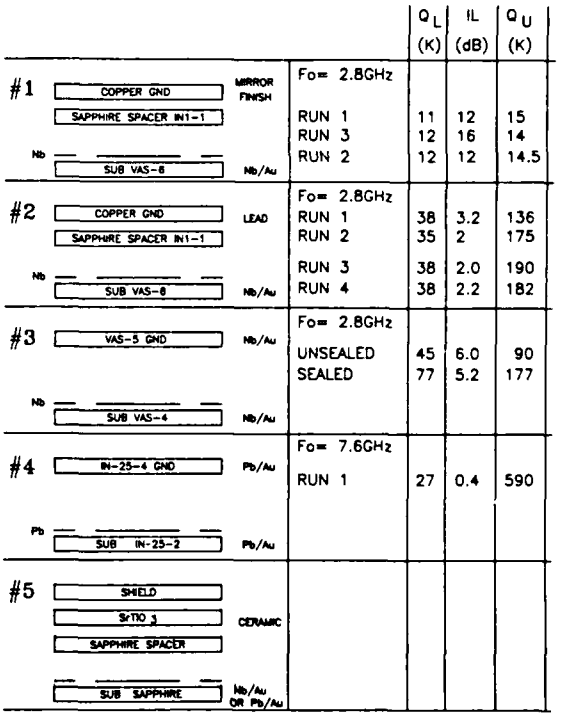



Figure 6. Summary of loaded Q ($Q_L$), insertion loss (IL) and unloaded Q ($Q_u$) values for niobium and lead stripline resonators.

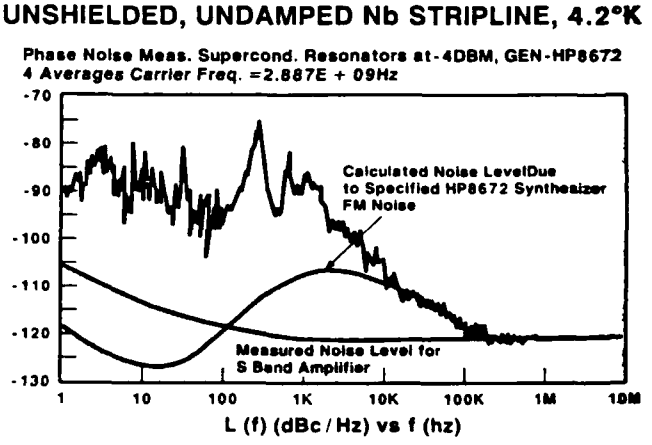



Figure 7. Preliminary self noise measurement of niobium stripline resonator at 2.9 GHz.

## Acknowledgements

The authors wish to acknowledge the assistance of Don Watt and Stan Pieseski in fabrication and Gary Draper in the microwave measurements. The invaluable assistance of M. Driscoll is detailed in Reference 3.

## References

1. A. J. DiNardo et al., "Superconducting Microstrip High-Q Microwave Resonators," J. Appl. Phys., 42, 1 (January 1971), p. 186.
2. E. L. Ginzton, Microwave Measurements, McGraw-Hill, New York, 1957.
3. We are grateful to Mr. M. Driscoll of the Signal Generation and Control Section of the Electronic Systems Group, Baltimore, MD for lending his expertise and performing these measurements.
4. H. L. Salvo, Jr. et al., "Shear Mode Transducers for High Q Bulk Microwave Resonators," 41st Annual Frequency Control Symposium Proceedings," IEEE 87CH2427-3 (1987), p. 388.

# AUTHOR INDEX


| AUTHOR | PAGE |
|---|---|
| Adam, J. | 556 |
| Albert, W. | 78 |
| Allan, D. | 419,465,485,514 |
| Anderson, A. | 545 |
| Angove, R. | 73 |
| Armington, A. | 162 |
| Aubry, J. | 101 |
| Audoin, C. | 496 |
| Avinens, C. | 138 |
| Balascio, J. | 162 |
| Ballato, A. | 6,53 |
| Barnes, J. | 426 |
| Barns, R. | 116 |
| Bauch, A. | 490 |
| Bhaskar, N. | 510 |
| Bigler, E. | 224 |
| Brajovic, V. | 404 |
| Brown, H. | 116 |
| Burns, S. | 272 |
| Capelle, B. | 19,85,208 |
| Carr, P. | 550 |
| Carru, H. | 19,101 |
| Cerez, P. | 496 |
| Chase, C. | 378 |
| Chuang, S. | 29 |
| Citrinovitch, A. | 485 |
| Clarke, R. | 472 |
| Clements, A. | 432 |
| Cook, R. | 525 |
| Coquerel, R. | 224 |
| Coyle, L. | 352 |
| Cracknell, M. | 284 |
| Curie, M. | 85 |
| de Labachelerie, M. | 496 |
| DeBoer, H. | 490 |
| Derov, J. | 550 |
| Detaint, J. | 19,85,138 |
| Dimarcq, N. | 496 |
| Dobnikar, B. | 456 |
| Dowsett, J. | 205 |
| Driscoll, M. | 369 |
| EerNisse, E. | 442 |
| Embree, D. | 276 |
| Emmons, D. | 334 |
| English, T. | 532 |
| Euler, F. | 169 |
| Felton, C. | 432 |
| Filler, R. | 380,447 |
| Fischer, B. | 490 |
| Foise, J. | 127 |
| Frueholz, R. | 525 |
| Gee, C. | 453 |
| Gehrke, J. | 412 |
| Giordano, V. | 496 |
| Goiffon, A. | 138 |
| Goldberg, J. | 304 |
| Gray, J. | 514 |
| Greer, J. | 239 |
| Groves, J. | 288 |
| Gualtieri, J. | 155 |
| Guillon, P. | 259 |
| Halperin, A. | 184 |
| Hamel, A. | 496 |
| Hanna, S. | 112 |
| Hanson, W. | 472 |
| Harrison, A. | 284 |
| Hauden, D. | 224 |
| Heindorff, T. | 490 |
| Hellwig, H. | 419 |
| Hertz, A. | 519 |
| Hoenig, S. | 189 |
| Hou, J. | 217 |
| Howe, D. | 472 |
| Hwang, H. | 176 |
| Iafrate, G. | 540 |
| Ikeda, Y. | 327 |
| Irvine, R. | 127 |
| Jacques, C. | 505 |
| James, B. | 146 |
| Jeffries, J. | 472 |
| Jespersen, J. | 472 |
| Johnson, G. | 127 |
| Jumas, J. | 138 |
| Kahla, C. | 510 |
| Kandpal, P. | 364 |
| Kartaschoff, P. | 419 |
| Katz, S. | 184 |
| Kawashima, H. | 45,65 |
| Ketcham, R. | 106 |
| Kim, B. | 357 |
| Klawitter, R. | 412 |
| Klepczynski, W. | 472 |
| Kline, G. | 106,272 |
| Klohn, K. | 53 |
| Korde, R. | 532 |
| Kosinski, J. | 53 |
| Lakin, K. | 106 |
| Laton, R. | 378 |
| Laudise, R. | 116 |
| LeChopier, S. | 101 |
| Lee, P. | 14,29 |
| Leeson, E. | 127 |
| Lefevre, R. | 101 |
| Lepek, A. | 485,519 |
| Lipson, H. | 169 |
| Litman, W. | 485 |
| Loboda, M. | 263 |
| Logan, S. | 276 |
| Lombardi, M. | 432 |
| Lopez, A. | 162,176 |
| Lukaszek, T. | 53 |
| Lutes, G. | 478 |
| Machlan, H. | 514 |
| Malley, G. | 532 |
| Martin, D. | 334 |
| Martin, J. | 162,176 |
| Matsuyama, M. | 65 |
| Mattison, E. | 352 |
| McAvoy, B. | 556 |
| McGowan, R. | 53 |
| Miljkovic, M. | 404 |
| Milliren, B. | 334 |
| Miyayama, T. | 327 |
| Mizan, M. | 53 |
| Montress, G. | 263 |
| Morley, P. | 205 |
| Moulton, G. | 389 |
| Myers, A. | 472 |
| Nakazato, M. | 65 |
| Newberg, I. | 453 |
| Nicholas III, H. | 357 |
| Oates, D. | 545 |
| Okamoto, Y. | 95 |
| Okano, S. | 327 |
| Parker, T. | 239,263,540 |
| Parshad, R. | 93 |
| Parzen, B. | 348 |
| Peach, R. | 38 |
| Peppler, T. | 465,485 |
| Philippot, E. | 138 |
| Popovic, V. | 456 |
| Popovic-Milovanovic, V. | 456 |
| Powell, W. | 472 |
| Primas, L. | 478 |
| Roberts, G. | 211 |
| Rogers, R. | 301 |
| Samueli, H. | 357 |
| Sato, S. | 288 |
| Schroeder, R. | 490 |
| Schwartzel, J. | 19,85,138 |
| Sebastian, M. | 208 |
| Sekimoto, H. | 95 |
| Sharma, A. | 93 |
| Sharpe, D. | 284 |
| Sheehey, R. | 276 |
| Shenhar, A. | 485 |
| Shick, D. | 230 |
| Simpson, E. | 116 |
| Sinha, B. | 288 |
| Skudera, Jr., W. | 252 |
| Smythe, R. | 73 |
| Spencer, R. | 205 |
| Stein, S. | 447 |
| Steinbeck, J. | 545 |
| Stern, A. | 519 |
| Stevens, D. | 116,276 |
| Sudo, Y. | 288 |
| Sydnor, R. | 478 |
| Tang, M. | 14 |
| Theobald, G. | 496 |
| Thompson, P. | 272 |
| Thurmond, G. | 453 |
| Tiersten, H. | 230 |
| Tremblay, P. | 505 |
| Trifunovic, G. | 404 |
| Vanek, M. | 432 |
| Vanier, J. | 419 |
| Vasiljevic, D. | 342 |
| Vig, J. | 397,419 |
| Wagner, G. | 556 |
| Walls, F. | 279,432 |
| Ward, R. | 442 |
| Warner, Jr., A. | 202 |
| Watts, M. | 442 |
| Wheeler, P. | 472 |
| Wiggins, R. | 442 |
| Winkler, G. | 419 |
| Yannoni, N. | 419 |
| Yen, H. | 453 |
| Yong, Y. | 29,397 |
| Yuen, C. | 378 |
| Zarfaty, Y. | 519 |
| Zarka, A. | 19,85,138,208 |
| Zeroug, S. | 112 |
| Zheng, Y. | 85 |

## SPECIFICATIONS AND STANDARDS GERMANE TO FREQUENCY CONTROL

### INSTITUTE OF ELECTRICAL AND ELECTRONIC ENGINEERS

Order from: IEEE Service Center
445 Hoes Lane
Piscataway, NJ 08854
(201)981-0060

176-1987 IEEE Standard on Piezoelectricity

177-1966 Piezoelectric Vibrators, Definitions and Methods of Measurements for (ANSI C83.17-1970)

180-1986 IEEE Standard Definitions of Primary Ferroelectric Terms

319-1971 Piezomagnetic Nomenclature

### DEPARTMENT OF DEFENSE

Order from: Naval Publication & Form Center
5801 Tabor Avenue
Philadelphia, PA 19120
(215)697-2000

MIL-C-3098 Crystal Unit, Quartz, Gen Spec for
MIL-C-24523 (SHIPS), Chronometer, Quartz Crystal
MIL-F-15733 Filters and Capacitors, Radio Interference, General Specification for
MIL-F-18327 Filters, High Pass, Low Pass, Band Pass Suppression and Dual Functioning, Gen Spec for
MIL-F-28861 Filters and Capacitors, Radio Frequency/ Electromagnetic Interference Suppression, Gen Spec for
MIL-F-28811 Frequency Standard, Cesium Beam Tube
MIL-H-10056 Holders (Encl), Crystal, Gen Spec for
MIL-O-55310 Oscillators, Crystal, Gen Spec for
MIL-O-39021 Oven, Crystal, Gen Spec for
MIL-S-49433 Surface Acoustic Wave Devices, Bandpass Filter
MIL-S-49433(ER) Surface Acoustic Wave Devices, Gen Spec for
MIL-STD-683 Crystal Units, Quartz/Holders, Crystal
MIL-STD-188-115 Interoperability and Performance Standards for Communications, Timing and Synchronization Subsystems
MIL-STD-1395 Filters and Networks, Selection and Use of
MIL-W-46374D Watch wrist: General Purpose
MIL-W-87967 Watch wrist: Digital

### GENERAL SERVICES ADMINISTRATION

Order from: Naval Publication & Form Center **or**
General Services Administration Business Service Centers in major U.S. cities

FED-STD-1002 Time and Frequency Reference Information in Telecommunication Systems

### ELECTRONIC INDUSTRIES ASSOCIATION

Order from: Electronic Industries Association
2001 Eye Street, NW
Washington, DC 20006
(202)457-4900

(a) Holders and Sockets

RS-192-A, Holder Outlines and Pin Connections for Quartz Crystal Units (Standard Dimensions for Holder types).

RS-367, Dimensional and Electrical Characteristics Defining Receiver Type Sockets (including crystal sockets).

RS-417, Crystal Outlines (Standard dimensions and pin connections for current quartz crystal units-1974).

(b) Production Tests

RS-186-E, (All Sections), Standard Test Methods for Electronic Component Parts

EIA 512, Standard Methods for Measurement of Equivalent Electrical Parameters of Quartz Crystal Units, 1 kHz to 1 GHz, 1985.

EIA/IS-17-A, Assessment of Outgoing Nonconforming Levels in Parts Per Million (PPM).

EIA/IS-18, Lot Acceptance Procedure for Verifying Compliance with the Specified Quality Level (SQL) in PPM

(c) Applicatin Information

EIA Components Bulletin No. CB6-A, Guide for the Use of Quartz Crystal Units for Frequency Control, Oct 1987

(d) EIA-477, Cultured Quartz (Apr. 81)

EIA-477-1, Quartz Crystal Test Methods (May 1985)

### INTERNATIONAL ELECTROTECHNICAL COMMISSION (IEC)

Order from: American National Standards Inst. (ANSI)
1430 Broadway
New York, NY 10018
(212)354-3300

### IEC PUBLICATIONS

122: - Quartz crystal units for frequency control and selection.
122-1 (1976) Part 1: Standard values and test conditions Amendment No. 1 (1983).
122-2 (1983) Part 2: Guide to the use of quartz crystal units for frequency control and selection.
122-3 (1977) Part 3: Standard outlines and pin connections.
Amendment No. 1 (1984).
122-3A(1979) First supplement.
122-3B(1980) Second supplement.
122-3C(1981) Third supplement.
283 (1968) Methods for the measurement of frequency and equivalent resistance of unwanted resonances of filter crystal units.
302 (1969) Standard definitions and methods of measurement for piezoelectric vibrators operating over the frequency range up to 30 MHz.
314 (1970) Temperature control devices for quartz crystal units.
Amendment No. 1 (1979).
314A (1971) First supplement.
368: - Piezoelectric filters.
368-1 (1982) Part 1: General information, standard values and test conditions.
368A (1973) First supplement.
Amendment No. 1 (1977).
Amendment No. 2 (1982).
368B (1975) Second supplement.
368-2: - Part 2: Guide to the use of piezoelectric filters
368-2-1(1988) Section One - Quartz crystal filters.
368-3 (1979) Part 3: Standard outlines.
368-3A (1981) First supplement.
444: - Measurement of quartz crystal unit parameters by zero phase technique in a $\pi$-network.
444-1 (1986) Part 1: Basic method for the measurement

of resonance frequency and resonance resistance of quartz crystal units by zero phase technique in a $\pi$-network.

444-2 (1980) Part 2: Phase offset method for measurement of motional capacitance of quartz crystal units.

444-3 (1986) Part 3: Basic method for the measurement of two-terminal parameters of quartz crystal units up to 200 MHz by phase technique in a $\pi$-network with compensation of the parallel capacitance $C_0$.

444-4 (1988) Part 4: Method for the measurement of the load resonance frequency $f_L$, load resonance resistance $R_L$ and the calculation of other derived values of quartz crystal units, up to 30 MHz.

483 (1976) Guide to dynamic measurements of piezoelectric ceramics with high electromechanical coupling.

642 (1979) Piezoelectric ceramic resonators and resonator units for frequency control and selection. Chapter I: Standard values and conditions. Chapter II: Measuring and test conditions.

679: - Quartz crystal controlled oscillators.

679-1 (1980) Part 1: General information, test conditions and methods.

679-2 (1981) Part 2: Guide to the use of quartz crystal controlled oscillators.
Amendment No. 1 (1985).

689 (1980) Measurements and test methods for 32 kHz quartz crystal units for wrist watches and standard values.

758 (1983) Synthetic quartz crystal.
Chapter I: Specification for synthetic quartz crystal.
Chapter II: Guide to the use of synthetic quartz crystal.
Amendment No. 1 (1984).

862: - Surface acoustic wave (SAW) filters.

862-1-1(1985) Part 1: General information, test conditions and methods.

862-3 (1986) Part 3: Standard outlines.

## PROCEEDINGS

## ANNUAL FREQUENCY CONTROL SYMPOSIA

| NO. | YEAR | DOCUMENT NUMBER | OBTAIN FROM* | COST |
|---|---|---|---|---|
| 10 | 1956 | AD298322 | NTIS | $41.50 |
| 11 | 1957 | AD298323 | " | 44.50 |
| 12 | 1958 | AD298324 | " | 46.00 |
| 13 | 1959 | AD298325 | " | 49.00 |
| 14 | 1960 | AD246500 | " | 32.50 |
| 15 | 1961 | AD265455 | " | 28.00 |
| 16 | 1962 | PB162343 | " | 35.50 |
| 17 | 1963 | AD423381 | " | 43.00 |
| 18 | 1964 | AD450341 | " | 43.00 |
| 19 | 1965 | AD471229 | " | 47.50 |
| 20 | 1966 | AD800523 | " | 47.50 |
| 21 | 1967 | AD659792 | " | 41.50 |
| 22 | 1968 | AD844911 | " | 44.50 |
| 23 | 1969 | AD746209 | " | 25.00 |
| 24 | 1970 | AD764210 | " | 28.00 |
| 25 | 1971 | AD746211 | " | 28.00 |
| 26 | 1972 | AD771043 | " | 26.50 |
| 27 | 1973 | AD771042 | " | 34.00 |
| 28 | 1974 | ADA011113 | " | 31.00 |
| 29 | 1975 | ADA017466 | " | 34.00 |
| 30 | 1976 | ADA046089 | " | 40.00 |
| 31 | 1977 | ADA088221 | " | 44.50 |
| 32 | 1978 | | EIA | 20.00 |
| 33 | 1979 | | " | 20.00 |
| 34 | 1980 | | " | 20.00 |
| 35 | 1981 | | " | 20.00 |
| 36 | 1982 | ADA130811 | NTIS | 41.50 |
| 37 | 1983 | 83CH1957-0 | IEEE | 64.00 |
| 38 | 1984 | 84CH2062-8 | " | 61.00 |
| 39 | 1985 | 85CH2186-5 | " | 70.00 |
| 40 | 1986 | 86CH2330-9 | " | 70.00 |
| 41 | 1987 | 87CH2427-3 | " | 64.00 |
| 42 | 1988 | 88CH2588-2 | " | 70.00 |

* NTIS - National Technical Information Service
Sills Building
5285 Port Royal Road
Springfield, VA 22161

* EIA - Annual Frequency Control Symposium
c/o Electronic Industries Association
2001 Eye Street
Washington, DC 20006

*IEEE - Institute of Electrical & Electronics Engineers
445 Hoes Lane
Piscataway, NJ 08854 (201) 981-1393

Remittance must be enclosed with all orders. Prices are subject to change without prior notice.

A subject and author index for the Proceedings of the 10th through the 38th Symposia appears as a supplement to the 38th Proceedings volume.